Edición Biblioteca de IA en español

Inteligencia artificial para el diseño de nuevos materiales y la propulsión de cohetes

Potenciales de IA, laboratorios autónomos y aprendizaje automático informado por la física (PIML)

Kim Kyung-jin, abogado

kimkj.com · Biblioteca de IA · 2026

Inteligencia artificial para el diseño de nuevos materiales y la propulsión de cohetes

Potenciales de IA, laboratorios autónomos y aprendizaje automático informado por la física (PIML)

Kim Kyung-jin, abogado

Se ha estructurado en diez capítulos que abarcan desde los potenciales de aprendizaje automático que aprenden el movimiento de un solo átomo, los modelos generativos que definen primero las propiedades deseadas y generan inversamente estructuras cristalinas, los laboratorios autónomos donde los robots repiten experimentos por sí mismos, las aleaciones de alta entropía y los compuestos cerámicos para proteger las zonas de alta temperatura de los cohetes, la monitorización de defectos y la garantía de vida útil por fatiga en la impresión 3D de metales, el aprendizaje automático informado por la física que incorpora leyes físicas en la función de pérdida para calcular la combustión, hasta el diseño de canales de refrigeración, el diagnóstico en tiempo real y el control mediante aprendizaje por refuerzo de motores cohete de propelente líquido. Se ha redactado sin fórmulas matemáticas para lectores generales interesados en la ciencia y la tecnología, incluye fuentes documentales verificadas en cada capítulo, y constituye una edición revisada que incorpora las críticas rigurosas, al estilo de un director de tesis, de dos revisores de inteligencia artificial.

Repositorio público de materiales de investigación

Este libro fue organizado y producido mediante inteligencia artificial. Se trata de una recopilación y síntesis de materiales en los principales idiomas del mundo. NotebookLM no limita el idioma de las preguntas, con independencia de la lengua en que estén redactadas las fuentes. El lector puede formular preguntas en su propio idioma al mismo repositorio de investigación utilizado para preparar este libro. Esto se aplica aun cuando parte de los materiales se encuentre en coreano, inglés u otros idiomas. Al abrir la configuración en NotebookLM y establecer el idioma de salida deseado, las respuestas y los resultados generados se presentarán en dicho idioma.

Repositorio público de NotebookLM utilizado en la preparación de este libro (AI for Materials Science and Rocket Engine Engineering, 129 documentos):
https://notebook.google.com/notebook/ac39e3ce-77a9-4e78-9aaf-b52c45c2934d

La edición original en coreano de este libro se encuentra en la Biblioteca de IA de kimkj.com:
https://kimkj.com/ai-library/?mod=document&uid=12453

Índice

Haga clic en el título de cada capítulo para ir al texto correspondiente. Para leer desde el principio, seleccione el capítulo 1.

Capítulo 1. Fundamentos de la informática de materiales y potenciales interatómicos de aprendizaje automático (MLIP)

Introducción

Todos los objetos que usamos están hechos de materiales. Las cucharas se hacen de hierro y los vasos se hacen de vidrio. Las toberas de los motores de cohete también se fabrican con metales y cerámicas especiales. Si el material cambia, las propiedades del objeto también cambian.

Las propiedades de un material comienzan en partículas diminutas e invisibles a simple vista. Esas partículas son los átomos. El átomo es la unidad básica que compone la materia. Según cómo se unan y se muevan los átomos, el hierro puede volverse duro o blando.

Quienes construyen cohetes desean conocer de antemano el movimiento de estos átomos. En el interior de un motor de cohete, la temperatura sube hasta miles de grados. La presión también es sumamente alta. Es necesario saber con antelación si el metal puede resistir sin fundirse en tales condiciones. Fabricar piezas reales para probarlas cuesta mucho dinero y tiempo. Por eso, los científicos mueven los átomos dentro de una computadora. Esto es la simulación atómica.

Para mover átomos con una computadora se necesitan reglas. Esto se debe a que es preciso saber en qué dirección y cuánta fuerza recibe cada átomo. Durante mucho tiempo, el método para crear estas reglas con precisión y el método para crearlas con rapidez fueron distintos entre sí. El método exacto era lento y el método rápido era inexacto. El nuevo camino que derriba el muro entre ambos es el potencial interatómico de aprendizaje automático (Machine Learning Interatomic Potentials, MLIP). En adelante, a esta tecnología la llamaremos MLIP.

Este capítulo explica paso a paso qué es el MLIP y cómo ha evolucionado. Primero, examina dos métodos tradicionales de cálculo atómico. A continuación, muestra cómo el aprendizaje automático tendió un puente entre ambos. Sigue el orden cronológico desde los métodos iniciales hasta los más recientes. Por último, señala qué precauciones deben tomarse al utilizar esta tecnología en la investigación de nuevos materiales y materiales para cohetes. No se utilizan fórmulas matemáticas. Se explica con palabras que permiten imaginar ilustraciones.

El término informática de materiales (materials informatics) también aparece aquí por primera vez. Esta expresión significa informática aplicada a los materiales. Es una disciplina que recopila gran cantidad de datos sobre materiales y los analiza por computadora para descubrir nuevos materiales. Dentro de ella, el MLIP es una herramienta fundamental encargada de los cálculos a nivel atómico.

Aclaramos de antemano por qué esta historia se vincula con los cohetes. El desarrollo de cohetes es difícil de separar de la búsqueda de nuevos materiales. Si existen metales más ligeros y capaces de soportar temperaturas más altas, el rendimiento del cohete mejora. Fabricar y probar tales materiales uno por uno en el laboratorio lleva varios años. Si primero se filtran los candidatos mediante cálculos atómicos en la computadora, se reduce este tiempo. El MLIP es la clave para hacer que ese cálculo sea rápido y preciso. Tras leer este capítulo, se comprenderá el marco general de cómo nacen los nuevos materiales dentro de una computadora.

1. Introducción y antecedentes: combinación de la teoría del funcional de la densidad (DFT) y el aprendizaje automático

Lo que se aborda en esta sección

Para tratar átomos con una computadora, es necesario conocer la energía y las fuerzas entre ellos. Existían principalmente dos formas de obtener estos valores. Una es el método exacto que calcula incluso el movimiento de los electrones. La otra es el método rápido que los obtiene mediante fórmulas sencillas predeterminadas. Esta sección explica qué hace bien y qué no puede hacer cada uno de estos dos métodos.

Asimismo, aborda el antiguo dilema entre ambos métodos: el método exacto es demasiado lento y el método rápido comete errores con frecuencia. Para resolver este problema surgió el aprendizaje automático. Esta sección traza primero ese panorama general.

Algunas palabras de esta sección se seguirán usando más adelante. Son energía, fuerza y superficie de energía potencial. Comprender su significado ahora facilitará la lectura de las siguientes secciones. No hay fórmulas difíciles. Se explica únicamente con palabras que permiten visualizar imágenes mentales.

Los dos valores fundamentales en los cálculos que tratan con átomos son la energía y la fuerza. La energía indica cuán estable es una disposición atómica. Cuanto más baja es la energía, más relajado o estable es ese estado de disposición. La fuerza indica en qué dirección es empujado cada átomo. Al conocer la fuerza, se puede calcular hacia dónde se moverá el átomo en el siguiente instante.

En realidad, la energía y la fuerza no son valores independientes. La fuerza proviene de cuán abruptamente cambia la energía según la posición. Si comparamos la energía con el relieve del terreno, la fuerza es la pendiente de esa cuesta. Una pelota rueda hacia donde la pendiente desciende. Los átomos también son empujados hacia donde la energía disminuye. Por lo tanto, si se conoce el mapa de energía, la fuerza se obtiene a partir de la pendiente de ese mapa. Esta relación volverá a ser importante más adelante.

La teoría del funcional de la densidad (Density Functional Theory, DFT) es un método para calcular la energía y las fuerzas mediante la distribución de electrones. En adelante la llamaremos DFT. Los electrones son partículas diminutas con carga eléctrica que orbitan alrededor de los átomos. Dónde y cuántos electrones se concentran determina los enlaces químicos. La DFT calcula esta distribución electrónica basándose en los principios de la mecánica cuántica. La mecánica cuántica son las leyes de la física que rigen el mundo microscópico. Por ello, la DFT describe con precisión los enlaces químicos.

Los valores calculados por la DFT tampoco son una verdad absoluta. La DFT trata la repulsión y atracción entre electrones mediante fórmulas de aproximación. Según cómo se elijan estas fórmulas de aproximación, los resultados varían ligeramente. Por eso, el valor de la DFT es un valor de referencia obtenido bajo una configuración determinada, más que un valor verdadero exacto. El aprendizaje automático aprende este valor de referencia. Cuando más adelante nos refiramos a la DFT como la respuesta correcta, debe entenderse en este sentido.

El problema es la velocidad. En la DFT, la cantidad de cálculo aumenta con rapidez a medida que crece el número de átomos. Si el número de átomos se duplica, el tiempo de cálculo se multiplica mucho más que el doble. Incluso usando grandes supercomputadoras, la cantidad de átomos que se puede manejar suele ser de unos pocos cientos. El tiempo simulable también se limita a un instante sumamente breve. No resulta adecuado para observar durante mucho tiempo grandes bloques de material, como ocurre con los materiales para cohetes.

En el extremo opuesto se encuentran los campos de fuerza empíricos (Empirical Force Field). Este método calcula las fuerzas entre átomos mediante fórmulas sencillas predeterminadas. No calcula los electrones uno por uno. En su lugar, determina las fuerzas utilizando solo unos pocos valores, como la distancia interatómica. Como el cálculo es ligero, puede manejar millones de átomos. También permite observar periodos de tiempo mucho más prolongados.

Los campos de fuerza empíricos son rápidos, pero tienen desventajas. Un campo de fuerza que asume enlaces fijos no describe bien los momentos en que los enlaces se rompen y se forman de nuevo. Los campos de fuerza reactivos o para metales desarrollados posteriormente logran abordar en cierta medida los cambios de enlace y los defectos en rangos limitados. Aun así, resulta difícil mantener la precisión en un amplio abanico de condiciones. Las transiciones de fase, donde los átomos cambian drásticamente de posición, también son complicadas. En entornos calientes y extremos como los motores de cohete, estos fenómenos ocurren con frecuencia. Por ello, con campos de fuerza simples los errores se vuelven grandes.

Aquí podemos comparar la diferencia entre ambos métodos con una situación de la vida cotidiana. La DFT es como calificar minuciosamente la tarea de un solo alumno: es precisa, pero calificar la tarea de todos los estudiantes de la escuela lleva mucho tiempo. El campo de fuerza empírico es como revisar rápidamente solo las respuestas para ver si coinciden: es rápido, pero comete errores en los problemas difíciles. Durante mucho tiempo, los científicos tuvieron que elegir entre precisión o rapidez.

Los potenciales interatómicos de aprendizaje automático son un intento de derribar el muro de esta disyuntiva. Un potencial es como un mapa que indica cuánta energía corresponde a cada disposición atómica. A este mapa se le llama superficie de energía potencial. El MLIP aprende mediante aprendizaje automático los valores de referencia calculados por la DFT. El aprendizaje automático es un método de cálculo que descubre reglas por sí mismo a partir de numerosos ejemplos. Una vez que aprende suficientes ejemplos, predice con rapidez la energía y las fuerzas de nuevas disposiciones atómicas.

Expliquemos con un poco más de detalle cómo aprende el MLIP. Primero, se calculan con la DFT la energía y las fuerzas de diversas disposiciones atómicas. El resultado de estos cálculos equivale a un cuaderno de ejercicios con las respuestas resueltas. El modelo de aprendizaje automático examina este cuaderno y asimila las reglas. Desarrolla por sí mismo una intuición sobre por qué determinada disposición tiene cierta energía. Al concluir el entrenamiento, resuelve rápidamente incluso disposiciones nuevas que no estaban en el cuaderno.

El objetivo de este método es claro: lograr cálculos tan precisos como la DFT y tan rápidos como los campos de fuerza empíricos. Cuando se alcance este objetivo, se podrán probar materiales para cohetes a gran escala y durante largos periodos dentro de la computadora. Se podrá observar a escala atómica cómo vibran los metales a altas temperaturas y cómo se propagan las grietas. Será posible filtrar los materiales candidatos antes de las pruebas reales. El resto de este capítulo relata las diversas tecnologías destinadas a lograr este propósito.

Hay un aspecto que conviene aclarar de antemano. El MLIP no reemplaza por completo a la DFT. Esto se debe a que necesita los valores de referencia generados por la DFT para poder aprender. Ambos métodos no compiten, sino que forman una pareja. La DFT crea los valores de referencia y el MLIP propaga esos valores con rapidez. Esta cooperación hace posibles cálculos de gran escala.

2. Modelos clásicos de aprendizaje automático basados en descriptores

Lo que se aborda en esta sección

Para aprender la energía atómica con aprendizaje automático, primero es necesario convertir la geometría circundante del átomo en números. Esto se debe a que las computadoras procesan números y no dibujos. A esta lista de números se la denomina descriptor (Descriptor). Esta sección examina tres métodos de descriptores utilizados en las primeras etapas del MLIP.

Los tres métodos poseen características distintas: uno utiliza redes neuronales, otro compara la similitud entre dos entornos y el tercero aprende incluso los descriptores mediante aprendizaje profundo (deep learning). Los tres obtuvieron buenos resultados, pero también mostraron limitaciones. Esas limitaciones dieron paso a la siguiente generación de tecnologías.

Un buen descriptor debe cumplir una regla: el mismo material debe representarse siempre con los mismos números. Aunque giremos un escritorio, sigue siendo un escritorio. Si rotamos o trasladamos una estructura atómica en su conjunto, sus posiciones espaciales cambian. Sin embargo, el material sigue siendo el mismo. El descriptor no debe verse alterado por estos cambios. A esta propiedad se la denomina invariancia (Invariance).

Si no se preserva la invariancia, surgen problemas. El modelo confunde el mismo material como si fueran varios materiales distintos. En consecuencia, tiene que aprender lo mismo varias veces por separado. El aprendizaje se vuelve lento y las predicciones, inestables. Las primeras investigaciones sobre MLIP dedicaron grandes esfuerzos a incorporar manualmente esta invariancia dentro de los descriptores.

2.1 Redes neuronales que aprenden dividiendo la energía por átomo (BPNN)

En 2007, Behler y Parrinello propusieron una idea fundamental: descomponer la energía total en la suma de las energías de cada átomo individual. A este método se le llama descomposición de la energía atómica. Cada átomo observa únicamente su entorno local y produce un pequeño valor de energía. Al sumar los valores de todos los átomos, se obtiene la energía de la estructura completa. A esta red neuronal se la conoce como red neuronal de Behler-Parrinello (Behler-Parrinello Neural Network, BPNN).

Una red neuronal es una estructura computacional diseñada a semejanza de las conexiones entre células cerebrales humanas. Recibe entradas y produce respuestas a través de múltiples etapas. Al entrenarse con numerosos ejemplos, aprende las reglas entre las entradas y las respuestas. La BPNN recibe como entrada la geometría del entorno de un átomo y produce como respuesta la energía de dicho átomo.

En este proceso, la herramienta que convierte el entorno atómico en números son las funciones de simetría centradas en el átomo (Atom-Centered Symmetry Functions, ACSF). Esta función coloca un átomo en el centro y examina los átomos vecinos dentro de una distancia determinada. Registra numéricamente la distancia hasta los vecinos y los ángulos que estos forman entre sí. Se divide en una parte que recoge solo distancias y otra que abarca también ángulos. Los números generados de este modo no cambian aunque la estructura rote o se traslade.

Este método tiene también gran relevancia para la investigación de materiales para cohetes. Cuando un metal se calienta, los átomos vibran intensamente en sus posiciones. Las distancias y los ángulos respecto a los vecinos cambian continuamente. Las ACSF capturan estas variaciones en números. Por ello, constituyeron el primer paso para aprender el comportamiento de los materiales en función de la temperatura.

Sin embargo, sus limitaciones eran evidentes. A medida que aumenta la variedad de elementos químicos, la cantidad de números requeridos crece con rapidez. Las aleaciones para cohetes a menudo combinan múltiples elementos. Si hay cinco tipos de elementos, las combinaciones de pares de elementos aumentan notablemente. Las combinaciones de tríos de átomos que definen ángulos crecen a un ritmo aún más acelerado. A este problema, en el que la carga de cálculo aumenta drásticamente a medida que crecen las variedades, se le llama la maldición de la dimensionalidad.

Otra limitación reside en que los descriptores debían diseñarse manualmente. Es una persona quien decide qué distancias y ángulos registrar y con qué nivel de detalle. Si esta elección es incorrecta, se pierde información fundamental. Crear un buen descriptor era en sí mismo una tarea compleja. Los esfuerzos para mitigar esta carga dieron lugar a los métodos siguientes.

A pesar de ello, el legado de este método es muy significativo. La idea de dividir la energía total átomo por átomo sigue vigente en la actualidad. La mayoría de los modelos más avanzados continúan aplicando este principio de descomposición. Consiste en que cada átomo calcula su valor observando únicamente su entorno local, para luego sumar todas las contribuciones. Gracias a este enfoque, los cálculos pueden dividirse y procesarse eficientemente aun con un gran número de átomos. La propuesta de Behler y Parrinello se convirtió en la piedra angular del MLIP.

2.2 Potenciales que miden la similitud entre dos entornos (GAP)

En 2010, Bartók, Csányi y sus colaboradores eligieron un camino diferente: en lugar de redes neuronales, adoptaron un método que compara directamente cuán parecidos son dos entornos. A este método se le denomina potencial de aproximación gaussiana (Gaussian Approximation Potential, GAP). Cuando se encuentra con un nuevo entorno atómico, lo compara con los entornos previamente aprendidos. Si existen muchos entornos similares, calcula la respuesta tomando como referencia sus energías.

La herramienta clave para esta comparación es la superposición suave de posiciones atómicas (Smooth Overlap of Atomic Positions, SOAP). SOAP describe el entorno de un átomo no como puntos rígidos, sino como una nube suave. Es como colocar una esfera difusa en cada átomo vecino. Si las nubes de dos entornos atómicos se superponen en gran medida, se consideran entornos similares. Este método fue diseñado para ser invariante ante rotaciones.

Hay una razón para representarlo como una nube suave. Los átomos siempre vibran ligeramente. Si la posición se considera solo como un punto rígido, incluso una pequeña vibración hace que los valores fluctúen bruscamente. Si se representa como una nube, las pequeñas vibraciones se asimilan suavemente. Por ello, SOAP resultó muy eficaz para tratar defectos en cristales o materiales desordenados.

Los potenciales de aproximación gaussiana lograron una alta precisión incluso con pocos datos. Esto resulta de gran ayuda cuando los cálculos de DFT son costosos. Se utilizaron ampliamente en el estudio de defectos cristalinos, estados líquidos y materiales amorfos desordenados. En los materiales para cohetes, los defectos y el desorden también son temas importantes. Por esta razón, este método se convirtió en un estándar citado durante mucho tiempo.

La limitación se hace evidente cuando los datos aumentan. Este método compara un entorno nuevo con todos los entornos aprendidos. Cuantos más datos de entrenamiento haya, más elementos habrá para comparar. Esto incrementa la carga tanto de cálculo como de almacenamiento. La cantidad de cálculo resulta excesiva para manejar grandes volúmenes de datos que superan los cientos de miles, como ocurre hoy en día.

Por tanto, este método cedió su lugar a otras arquitecturas en problemas de gran escala. Para aprender con rapidez a partir de grandes volúmenes de datos, las redes neuronales de aprendizaje profundo resultaron más adecuadas. Esto se debe a que el aprendizaje profundo es un método que gana eficacia a medida que aumentan los datos. No obstante, el estándar de precisión establecido por los potenciales de aproximación gaussiana sigue vigente en la actualidad. Se utiliza con frecuencia como referencia comparativa al medir el rendimiento de nuevos modelos.

La idea introducida por SOAP también perdura. Se trata del concepto de considerar el entorno atómico como una densidad suave. Esta idea se integró en varios modelos posteriores. Tratar las posiciones como una distribución suave en lugar de puntos rígidos estabiliza los cálculos, ya que los valores no fluctúan ante pequeñas vibraciones. En cálculos a altas temperaturas donde los átomos vibran intensamente, como en los materiales para cohetes, esta estabilidad resulta de gran ayuda.

2.3 Aprendizaje profundo que aprende incluso los descriptores por sí mismo (DeePMD)

Zhang, Car y sus colaboradores introdujeron plenamente el aprendizaje profundo en los cálculos atómicos en 2018. Este método se denomina dinámica molecular de potencial profundo (Deep Potential Molecular Dynamics, DeePMD). La dinámica molecular es una simulación que sigue el movimiento de los átomos a lo largo del tiempo. DeePMD calcula las fuerzas que se utilizarán en esta simulación mediante aprendizaje profundo.

En los métodos anteriores, los seres humanos diseñaban los descriptores. DeePMD redujo enormemente esta carga. La red neuronal transforma por sí misma las coordenadas del entorno atómico en una representación adecuada. Ya no es necesario que una persona elija distancias y ángulos uno por uno. Al mismo tiempo, preservó la propiedad de ser invariante ante rotaciones y traslaciones.

Aquí conviene explicar con un poco más de detalle el término aprendizaje profundo. El aprendizaje profundo se refiere a redes neuronales apiladas en múltiples capas. Al apilar capas profundas, se pueden aprender reglas complejas. Las redes neuronales poco profundas solo aprenden reglas simples. Las redes neuronales profundas abarcan incluso las intrincadas relaciones entre átomos. A cambio, requieren una gran cantidad de datos y cómputo para aprender. El término 'profundo' proviene del hecho de que las capas son profundas.

La fortaleza de este método reside en su gran escala. En el aprendizaje profundo, la carga de cálculo no aumenta drásticamente incluso con muchos átomos. Por ello, DeePMD puede manejar el movimiento de millones de átomos. Se ha utilizado en diversos materiales como agua, metales y semiconductores. Demostró su poder al observar grandes sistemas durante períodos prolongados.

DeePMD también tiene un significado importante en la investigación de materiales para cohetes. Para observar cómo cambia un bloque de metal a altas temperaturas, se requieren muchos átomos. Si solo se observa un fragmento pequeño, se pierde el comportamiento global. DeePMD puede manejar fragmentos grandes, por lo que resultaba adecuado para este tipo de investigaciones.

Aun así, persistía la limitación compartida por todos los métodos iniciales basados en descriptores. Estos métodos solo observan por separado a los vecinos situados dentro de una distancia determinada. Para captar la influencia de átomos lejanos, es necesario aumentar la distancia de observación. Al hacerlo, el número de vecinos se incrementa en gran medida y la carga computacional aumenta. Resultaba estructuralmente difícil abarcar las interacciones de largo alcance.

Otra puerta que abrió DeePMD es la combinación con grandes herramientas de cómputo. Este método condujo a investigaciones que manejan miles de millones de átomos en grandes supercomputadoras.

También se utiliza para calcular a gran escala aleaciones en las que se mezclan múltiples elementos. La investigación de aleaciones de alta temperatura para cohetes requiere tales cálculos a gran escala, ya que en sistemas extensos es posible observar comportamientos que se pasarían por alto al mirar solo fragmentos pequeños.

Esta limitación dio lugar a una nueva idea: superar el método de colocar un solo átomo en el centro y observar solo su entorno. ¿Qué pasaría si los átomos intercambiaran información entre sí? Si la información se intercambia varias veces, la influencia de átomos distantes se transmite de forma sucesiva. Esta idea conduce a las redes neuronales de grafos de la siguiente sección.

3. Redes neuronales de grafos (GNN) equivariantes de última generación

Esta sección aborda el método que se ha convertido en la corriente principal de los MLIP en la actualidad: las redes neuronales de grafos (Graph Neural Network, GNN). Este método considera los átomos como nodos y las relaciones de vecindad entre átomos como líneas. Aquí, las líneas no representan enlaces químicos en sí mismos, sino la unión de vecinos que se encuentran dentro de una distancia determinada. Estas líneas pueden cambiar durante el cálculo. Es similar a un mapa de metro que dibuja las estaciones como puntos y las conexiones entre ellas como líneas. La energía se calcula intercambiando información a lo largo de los puntos y las líneas.

La palabra clave de esta sección es la equivariancia (Equivariance). Es una propiedad que forma pareja con la invariancia de la sección anterior. Primero se explica qué es la equivariancia y por qué es necesaria en los cálculos atómicos. A continuación, se examinan en orden cronológico, desde los iniciales hasta los más recientes, los modelos representativos que incorporan esta propiedad. Finalmente, se abordan los modelos grandes preentrenados con amplios conjuntos de datos.

Las redes neuronales de grafos calculan mediante el intercambio de información. Cada átomo envía un mensaje a sus vecinos. Luego, reúne los mensajes recibidos de los vecinos y actualiza su propio estado. Este proceso se repite varias veces. Con un intercambio, conoce el átomo contiguo. Con dos intercambios, conoce incluso al vecino de su vecino. Cuanto más se repite, más se transmite la influencia de átomos más distantes. Este método se denomina paso de mensajes (Message Passing).

Esta estructura resuelve en cierta medida las limitaciones de la sección anterior. No es necesario forzar un aumento de la distancia de observación, ya que al intercambiar mensajes varias veces, la información más allá de los vecinos inmediatos se transmite sucesivamente. Sin embargo, si el número de capas es fijo, el alcance de la información también queda limitado. Efectos como la electrostática o las fuerzas de dispersión a muy larga distancia son difíciles de abarcar únicamente con este método. Tales efectos deben complementarse con términos de largo alcance diseñados por separado. Aun así, las redes neuronales de grafos constituyeron un gran avance al incorporar de manera eficiente interacciones de amplio alcance.

3.1 Simetría que se mantiene al rotar y reflejar

En el mundo atómico existen simetrías que se conservan. Aunque una estructura se traslade o rote por completo, las leyes de la física permanecen idénticas. Al conjunto que reúne traslaciones, rotaciones y reflexiones se le llama E(3). Entre ellas, al conjunto que reúne únicamente las rotaciones se le llama SO(3). Los nombres parecen complejos, pero su significado es simple: es una colección de traslaciones, rotaciones y reflexiones como en un espejo. Rotar y reflejar son movimientos distintos. La rotación gira la forma tal como está, mientras que la reflexión crea una imagen especular con la izquierda y la derecha invertidas.

La invariancia vista anteriormente es la propiedad de que los valores no cambian ante tales movimientos. La energía es una cantidad invariante. Aunque se rote la estructura, la energía de esa estructura debe ser la misma. Sin embargo, no todos los valores son invariantes. La fuerza es una cantidad que tiene dirección. Si se rota la estructura, la dirección de la fuerza también debe rotar junto con ella.

Esta propiedad es precisamente la equivariancia. La equivariancia es la propiedad de que, si la entrada rota, la salida también rota de la misma manera. La fuerza es una cantidad equivariante. Si la estructura se rota hacia la derecha, la fuerza ejercida sobre cada átomo también rota hacia la derecha. Solo al respetar conjuntamente la invariancia y la equivariancia, los cálculos atómicos concuerdan con las leyes de la física.

Se puede distinguir entre cantidad invariante y cantidad equivariante en una sola frase: si el valor permanece igual al rotar, es una cantidad invariante; si el valor rota conjuntamente al rotar, es una cantidad equivariante. La energía es un escalar sin dirección, por lo que es una cantidad invariante. Valores como la temperatura o la densidad tampoco tienen dirección y, por tanto, son invariantes. Valores con dirección, como la fuerza o la velocidad, son cantidades equivariantes. El modelo debe reconocer a qué categoría pertenece cada valor para tratarlo adecuadamente.

Aquí retomamos la explicación que se había pospuesto en la sección anterior. Se mencionó que la fuerza es el gradiente del mapa de energía. Un buen modelo no predice la fuerza de manera aislada e independiente. Primero calcula la energía y luego obtiene la fuerza como el gradiente de ese mapa de energía. Hay una razón para respetar este orden. Si la fuerza se produce de manera arbitraria, puede generarse una fuerza que no conserve la energía. Si se realizan cálculos prolongados con tales fuerzas, surge o desaparece energía de forma espontánea. Como resultado, la simulación se distorsiona o se desmorona gradualmente. Al obtener la fuerza a partir del gradiente de la energía, no se producen estas discrepancias. En física, las fuerzas derivadas de esta manera a partir de la energía se denominan fuerzas conservativas.

Incorporar estas dos propiedades en el modelo desde el principio ofrece una gran ventaja. El modelo no necesita aprender el mismo material por separado desde múltiples orientaciones. Una regla aprendida una vez se aplica automáticamente a todas las direcciones. Por ello, aprende de forma estable incluso con pocos datos. Dado que los cálculos de DFT son costosos, este ahorro de datos representa una gran fortaleza.

En los materiales para cohetes ocurren muchos fenómenos donde la dirección es fundamental. Las grietas se propagan en direcciones específicas. Las deformaciones por deslizamiento a lo largo de planos cristalinos también tienen dirección. Las reacciones de oxidación ocurren de manera diferente según la orientación de la superficie. Si no se maneja con precisión la dirección de las fuerzas, estos fenómenos se representan erróneamente. Los modelos que respetan la equivariancia resultan ventajosos en este aspecto.

Traslademos esta simetría a una imagen más sencilla. Supongamos que se toman varias fotografías de una misma estructura de bloques de juguete. La foto tomada de frente y la foto tomada de lado parecen diferentes. Sin embargo, la estructura de bloques en sí es la misma. Un ser humano comprende esto de inmediato. Pero para una computadora, ambas fotos son entradas distintas. Un modelo que no incorpora simetría puede considerar ambas fotos como materiales diferentes. Un modelo que incorpora simetría considera ambas fotos como el mismo material. Esta diferencia determina en gran medida la eficiencia del aprendizaje.

En resumen, la simetría no es un adorno, sino la estructura fundamental. Un modelo que respeta la simetría ahorra datos y estabiliza las predicciones. Un modelo que viola la simetría puede arrojar

respuestas físicamente anómalas. Por ello, la investigación reciente en MLIP se concentra en cómo incorporar adecuadamente esta propiedad.

3.2 De observar solo distancias a intercambiar también orientaciones

SchNet es una red neuronal de grafos temprana que apareció en 2018. Este modelo intercambia únicamente las distancias entre átomos como mensajes. La distancia es un valor sin dirección. Por tanto, SchNet es una red neuronal de grafos invariante que maneja exclusivamente información invariante. Su estructura es limpia y su cálculo es rápido.

El método que utiliza solo distancias tiene una debilidad: resulta difícil captar de forma directa enlaces donde la dirección es importante. Aunque la distancia sea la misma, si el ángulo es diferente, el enlace es distinto. Si se pasa por alto esta diferencia, se generan errores. DimeNet complementó este aspecto e incluyó también los ángulos entre átomos en los mensajes. De este modo continuaron los esfuerzos por incorporar información direccional.

Batzner, Kozinsky y sus colaboradores dieron un gran paso en 2022 con el modelo llamado NequIP. Este modelo va más allá de distancias y ángulos e intercambia la propia dirección como mensajes. Transmite y combina directamente valores con dirección. Esto es el paso de mensajes equivariante. Refleja con mucha mayor fidelidad la dirección de las fuerzas y la geometría tridimensional alrededor de los átomos.

La gran ventaja de NequIP es la eficiencia de datos. Dado que incorpora la información direccional conforme a los principios fundamentales, aprende bien incluso con pocos ejemplos. El artículo publicado señala que logró una alta precisión con una cantidad de datos de entrenamiento mucho menor que la de los modelos existentes. En investigaciones donde cada cálculo de DFT es costoso, este ahorro representa un beneficio tangible.

Conviene recalcar una vez más por qué es importante la eficiencia de datos. Cada cálculo de DFT requiere un tiempo de cómputo considerable. Si se necesitan decenas de miles de cálculos para el entrenamiento, el costo computacional se dispara. NequIP logra una precisión similar con un número de casos mucho menor. De este modo, con el mismo presupuesto se pueden abordar materiales más complejos. Los nuevos materiales para cohetes a menudo presentan escasez de datos. Un modelo con alta eficiencia de datos resulta sumamente útil en tales situaciones.

Este avance tiene una relevancia directa en los cálculos de materiales para cohetes. Las grietas, los defectos y las transiciones de fase están intrínsecamente ligados a la dirección. Los modelos equivariantes describen con mayor precisión estos fenómenos direccionales. Permiten aprender el comportamiento en entornos extremos incluso con escasos datos de DFT. Por ello, tras NequIP, el enfoque equivariante se aproximó al estándar en este campo.

En una época similar a la de NequIP, también continuaron las investigaciones orientadas a mejorar la eficiencia. Se trató de intentos por manejar la información direccional aligerando al mismo tiempo el cómputo. Un estudio reciente presentó un modelo equivariante de mayor eficiencia que se publicó en una revista académica en 2025. Es una dirección que reduce la carga computacional mientras preserva la precisión. Con la acumulación de estos esfuerzos, el enfoque equivariante se volvió verdaderamente viable para la investigación práctica.

Sin embargo, esto tiene un precio. Combinar la información direccional según los principios físicos requiere operaciones matemáticas pesadas. A esta operación se le llama producto tensorial. A cambio de una alta precisión, la carga de cálculo y de memoria es grande. Esta carga volverá a plantearse como un problema más adelante. Esta historia continúa en la sección 5.

3.3 Cómo capturar las relaciones de múltiples átomos a la vez (MACE)

Batatia, Csányi y sus colaboradores presentaron MACE en 2022. Este modelo combinó las redes neuronales de grafos equivariantes con una teoría física clásica llamada expansión de muchos cuerpos. «Muchos cuerpos» se refiere a los efectos que múltiples átomos crean de forma conjunta. En un material, no solo importa la relación entre dos átomos individuales. La disposición que forman tres o cuatro átomos juntos también modifica las propiedades.

Las redes neuronales de grafos anteriores intercambiaban mensajes varias veces para capturar estos efectos de muchos cuerpos. Si el número de intercambios es elevado, la carga de cálculo aumenta y el aprendizaje se ralentiza. MACE encontró un camino diferente. Multiplica entre sí la información de los vecinos recopilada una vez, generando los efectos de múltiples átomos al mismo tiempo. Por ello, con solo un reducido número de pasos de paso de mensajes, logra capturar los efectos de cuatro o más átomos.

La ventaja de este diseño radica en lograr tanto precisión como velocidad. Al reducir el paso de mensajes, el cálculo se vuelve más ligero. Al mismo tiempo, conserva la precisión sin perder los efectos de muchos cuerpos. Esta ventaja se magnifica en cristales grandes o en dinámicas moleculares de larga duración. Por esta razón, MACE se utiliza como un sólido modelo de referencia en numerosas investigaciones.

MACE también se emplea en la investigación de materiales para cohetes. Recientemente, un equipo de investigación trabajó con una aleación de alta entropía compuesta por aluminio, titanio, cromo, molibdeno y wolframio. Esta aleación es un material que mezcla varios elementos en proporciones similares. El equipo creó el potencial interatómico de esta aleación utilizando MACE y examinó su estabilidad de fase en función de la temperatura. Este estudio se publicó en una revista de física en 2026. Es un ejemplo del uso de MACE en el diseño de aleaciones para altas temperaturas. Sin embargo, lo que demostró este estudio fue el cálculo de la estabilidad de fase; no validó las grietas ni la vida útil de los componentes reales a partir de esta única investigación.

También examinamos sus limitaciones. MACE también utiliza productos tensoriales al combinar información direccional. Por lo tanto, si se incrementa la precisión direccional llamada momento angular, la carga de cálculo aumenta con rapidez. En sistemas muy grandes o en simulaciones de muy larga duración, esta carga sigue siendo un problema. Las investigaciones para reducir este inconveniente se expondrán más adelante.

Traslademos el efecto de muchos cuerpos a un ejemplo sencillo. En un aula, si solo observamos a dos estudiantes, apenas sabremos si se llevan bien. Cuando se juntan tres estudiantes, el ambiente cambia. Cuando se reúnen cuatro, surge otra dinámica diferente. Lo mismo ocurre con los átomos. La relación entre solo dos átomos no basta para abarcar todas las propiedades de un material. La disposición creada conjuntamente por múltiples átomos modifica sus propiedades. MACE captura estas relaciones entre múltiples átomos en pocos pasos.

Otro desafío son las situaciones fuera del rango aprendido. Por muy preciso que sea MACE, puede equivocarse ante configuraciones desconocidas que no formaron parte del entrenamiento. El entorno de los cohetes siempre es extremo y presenta muchos sucesos imprevistos. Por eso, cuanto más preciso sea un modelo, más necesita un mecanismo para alertar por sí mismo cuando desconoce algo. Este mecanismo es el aprendizaje activo de la sección 4.

3.4 Grandes modelos preentrenados en una amplia gama de materiales (CHGNet, MatterSim)

Los modelos anteriores suelen entrenarse para un único material o un rango estrecho. Cuando se enfrentan a un material nuevo, deben reentrenarse con los datos correspondientes. La tendencia orientada a reducir esta carga son los potenciales fundacionales. Un potencial fundacional es un modelo grande que ha aprendido previamente datos de una amplísima variedad de materiales. Es similar a un cocinero que domina diversos platos y sabe desenvolverse bien ante un ingrediente nuevo.

CHGNet es un potencial fundacional representativo presentado en 2023. Este modelo fue entrenado con una gran base de datos computacional llamada Materials Project. El Materials Project es un recurso público que recopila resultados de cálculos DFT de innumerables materiales. Además de esto, CHGNet aprende pistas sobre los estados de carga y magnéticos. Por ello, rastrea eficazmente la migración de litio en materiales de baterías o los cambios de fase en óxidos de metales de transición.

Hay una razón por la que CHGNet aprende estados de carga y magnéticos. En los materiales de baterías, el litio entra y sale, transfiriendo carga. En los óxidos de metales de transición, el estado magnético de los átomos cambia las propiedades. Si se pasan por alto estos cambios, el comportamiento del material se modelará incorrectamente. CHGNet incorpora estas pistas en la representación del entorno atómico para abarcar una variedad más amplia de materiales. Sin embargo, no es un modelo electroquímico que rastree detalladamente la carga como una magnitud conservada. Es más adecuado entenderlo como una representación aprendida que capta indicios del estado magnético. Como los óxidos y los metales de transición se utilizan con frecuencia en materiales para cohetes, esta capacidad resulta útil.

MatterSim es otro potencial fundacional presentado en 2024. Este modelo apunta a un rango de temperatura y presión extremadamente amplio. El artículo publicado señala que cubre desde 0 Kelvin hasta 5000 Kelvin, y desde 0 gigapascales hasta 1000 gigapascales. El Kelvin es la unidad de temperatura absoluta y el gigapascal es una unidad de presión. Dentro de este rango también se incluyen las condiciones de altas temperaturas de los motores cohete. Sin embargo, que el rango de entrenamiento coincida es distinto a predecir con exactitud en la práctica. 1000 gigapascales es un valor mucho más alto que la presión en la cámara de combustión de un cohete. Las situaciones propias de los cohetes, como los gases de combustión o las superficies oxidadas, deben verificarse por separado. Por ello, este modelo atrae la atención como punto de partida para la investigación de materiales en condiciones extremas. El artículo publicado afirma que este modelo ahorró una gran cantidad de datos. Gracias a haber aprendido una amplia gama de condiciones de antemano, responde con rapidez ante nuevos materiales.

Recientemente, también ha aparecido un nuevo potencial fundacional que eleva simultáneamente la precisión y la velocidad: el modelo GRACE. Este modelo se construyó bajo el marco de la expansión de conglomerados atómicos de grafos. Los investigadores entrenaron este modelo con una parte de las bases de datos de materiales más grandes disponibles. El estudio publicado indica que este modelo estableció una nueva línea de equilibrio entre precisión y eficiencia. Esta investigación se publicó en una revista de ciencia computacional de materiales en 2026. Demostró que, al ajustar adecuadamente un modelo grande, se puede adaptar a tareas específicas mientras se mantiene la precisión.

Los potenciales fundacionales están transformando la forma de investigar. Antes, se creaba un modelo desde cero para cada material. Ahora, se utiliza primero un modelo grande y solo se ajustan las partes necesarias. A este método de hacer ligeras adaptaciones se le llama ajuste fino. Incluso al encontrarse con una nueva aleación, se puede responder con rapidez utilizando pocos datos.

También están creciendo los esfuerzos por comparar de manera justa estos grandes modelos. Con la aparición de diversos potenciales fundacionales, se volvió difícil determinar cuál era mejor. Por ello, surgieron espacios de evaluación pública que miden el rendimiento bajo las mismas condiciones. Se trata de plataformas de evaluación comparativa creadas por diversos equipos de investigación. Estos espacios comparan la precisión, la velocidad y la estabilidad a la vez. Los usuarios pueden elegir el modelo adecuado para sus propios problemas.

Aun así, hay aspectos que requieren precaución. Los potenciales fundacionales no aciertan automáticamente en todas las situaciones. En aleaciones o reacciones superficiales que eran poco frecuentes en los datos de entrenamiento, los errores pueden aumentar. Por lo tanto, cuanto mejor sea el modelo, más necesario es acompañarlo de una validación. Las estructuras desconocidas deben verificarse recalculándolas mediante DFT. Este procedimiento de verificación es el tema de la siguiente sección.

4. Bucle de aprendizaje activo (Active Learning) y evaluación de errores e incertidumbre

Por excelente que sea un modelo, puede cometer errores en situaciones que no ha aprendido. El problema surge cuando el modelo da una respuesta con seguridad a pesar de estar equivocado. En tal caso, resulta difícil para las personas detectar el error. Esta sección aborda los métodos para que el modelo alerte por sí mismo en los momentos en que desconoce algo.

Ese método es el aprendizaje activo (Active Learning). El aprendizaje activo es un procedimiento cíclico en el que el modelo selecciona las estructuras en las que carece de confianza y las recalcula con exactitud. Analizaremos cómo funciona este ciclo y cómo se cuantifica numéricamente la falta de confianza. También examinaremos por qué este procedimiento actúa como un dispositivo de seguridad en entornos extremos y desconocidos, como los materiales para cohetes.

Traslademos la interpolación y la extrapolación a un ejemplo de la vida cotidiana. Si conocemos las temperaturas de los últimos años, podemos estimar la temperatura de hoy. Estimar dentro del rango de lo ya experimentado es la interpolación. Sin embargo, si sobreviene una ola de calor sin precedentes, la estimación fallará. Salir hacia un rango nunca antes experimentado es la extrapolación. De igual modo, los modelos funcionan bien dentro del rango que han experimentado, pero vacilan fuera de él.

El ámbito donde el modelo realiza predicciones seguras es el interior de los datos de entrenamiento. A esta región se la llama región de interpolación. Dentro del espacio delimitado por los datos de entrenamiento, las predicciones son confiables. No obstante, al salir de allí, la situación cambia. A este exterior se le llama región de extrapolación. En la región de extrapolación, el modelo puede ofrecer respuestas aparentemente verosímiles pero incorrectas.

Los errores en la región de extrapolación son peligrosos. Durante una simulación, los átomos pueden superponerse entre sí o la estructura puede colapsar de manera anómala. Si esto ocurre de forma silenciosa, la totalidad de los resultados pierde su credibilidad. El entorno de los cohetes siempre es extremo, por lo que la extrapolación ocurre con frecuencia. Esto se debe a que las altas temperaturas, los impactos fuertes, la oxidación y las grietas son sucesos poco comunes en los datos de entrenamiento. Por ello, se necesita imperativamente un mecanismo que detecte estos instantes.

El bucle de aprendizaje activo recorre cuatro etapas. Primero, se entrena el modelo con un conjunto reducido de datos DFT. A continuación, se calculan con rapidez los movimientos atómicos utilizando este modelo. Si durante el cálculo aparece una estructura sobre la cual el modelo duda, se marca. Se seleccionan únicamente esas estructuras y se recalculan con precisión mediante DFT. Los nuevos resultados se incorporan a los datos y se ajusta el modelo. Al repetir estas etapas, el modelo aprende

situaciones cada vez más amplias. Es similar a revisar de nuevo el mapa en cada bifurcación cuando se recorre un camino desconocido.

El núcleo de este bucle es medir numéricamente la falta de confianza. A este valor numérico se le llama incertidumbre (Uncertainty). Existen diversos métodos para cuantificar la incertidumbre. En las siguientes subsecciones examinaremos los métodos representativos.

4.1 Medir numéricamente los momentos en que el modelo desconoce (UQ)

El primer método para medir la incertidumbre consiste en utilizar el grado de extrapolación. Un modelo llamado potencial tensorial de momentos (Moment Tensor Potential, MTP) utiliza este enfoque. Este modelo mide numéricamente en qué medida la disposición atómica actual difiere de los datos de entrenamiento. A este número se le denomina grado de extrapolación. Si el grado de extrapolación es bajo, se trata de una configuración familiar y la predicción es segura. Si el grado de extrapolación es alto, es una configuración desconocida y debe recalcularse.

El segundo método consiste en contrastar las respuestas de múltiples modelos. Se crean varios modelos entrenados de forma ligeramente distinta con los mismos datos. A este conjunto se le llama ensamble. Si las respuestas de estos modelos para una nueva estructura son similares entre sí, generalmente son confiables. Si las respuestas discrepan ampliamente, esa estructura representa una señal de peligro. Sin embargo, los modelos que han aprendido de los mismos datos también pueden equivocarse juntos ante entradas desconocidas. Por lo tanto, la concordancia del ensamble es solo un indicador de referencia y no una garantía absoluta. Es como hacer la misma pregunta a varias personas: si las respuestas difieren, se considera un problema difícil.

El tercer método consiste en hacer que el propio modelo genere el margen de error. Investigaciones recientes han avanzado en esta dirección. Un equipo de investigación aplicó el método denominado aprendizaje profundo basado en evidencia a los potenciales interatómicos. Este método genera el valor predicho junto con el grado de confianza en dicho valor. El cálculo se vuelve más liviano porque no es necesario ejecutar múltiples modelos por separado. Este estudio se publicó en una revista académica en 2025.

También surgen investigaciones que buscan integrar la incertidumbre de manera más profunda en el aprendizaje activo. Un método hace que la propia simulación busque las zonas de incertidumbre. Se induce el movimiento atómico hacia las configuraciones en las que el modelo tiene poca confianza. De este modo, se recopilan con mayor rapidez estructuras con un alto valor de aprendizaje. Es una forma eficiente de abarcar situaciones desconocidas.

Este procedimiento constituye el pilar que salvaguarda la seguridad en la investigación de materiales para cohetes. En entornos extremos, siempre surgen configuraciones imprevistas. Si no se dispone de un mecanismo de incertidumbre, existe el riesgo de asumir como verdaderos resultados erróneos. Con un mecanismo de incertidumbre, el cálculo se detiene y se verifica en los momentos de riesgo. En esa misma medida, los resultados se vuelven confiables.

Los métodos para gestionar la incertidumbre continúan evolucionando. Recientemente, existen investigaciones que abordan incluso los casos en que el modelo está mal diseñado. Si un modelo falla desde sus supuestos básicos, la incertidumbre también resultará errónea. Un estudio que aborda este problema se publicó en una revista académica en 2025. Esto demuestra lo complejo que resulta medir la incertidumbre. Aun así, los avances en esta dirección sustentan la seguridad en el cálculo de materiales para condiciones extremas.

Otra ventaja de este procedimiento cíclico es el ahorro de datos. Los cálculos DFT son costosos. Por lo tanto, calcular estructuras arbitrarias sin criterio genera un gran desperdicio. El aprendizaje activo selecciona y calcula únicamente las estructuras con alto valor de aprendizaje. Con el mismo presupuesto, se recopilan datos más útiles. De ese modo, se obtiene un modelo superior con mayor rapidez.

Este método puede compararse con el estudio para un examen. Volver a resolver problemas que ya se conocen es una pérdida de tiempo. Si solo se seleccionan y resuelven las preguntas erradas, la habilidad mejora rápidamente. El aprendizaje activo también selecciona y subsana únicamente los puntos débiles del modelo. Es una forma de lograr grandes avances con un cómputo reducido. En la investigación de materiales para cohetes, el presupuesto computacional siempre es ajustado. Por ello, este ahorro resulta fundamental en el desarrollo real.

Sin embargo, medir la incertidumbre con precisión es en sí mismo una tarea difícil. Algunos métodos implican una gran carga computacional y otros pierden precisión. Si el modelo está mal diseñado desde el principio, incluso la incertidumbre calculada puede ser errónea. Recientemente continúan las investigaciones para abordar este problema. La cuantificación de la incertidumbre no es aún una tecnología completa, sino un campo en desarrollo.

5. Redes neuronales geométricas que reducen el producto tensorial: eficiencia computacional y suavidad de las curvas

Los modelos equivariantes son precisos, pero cargan con la desventaja de requerir una gran cantidad de cálculos. Una causa importante de esa carga es el producto tensorial mencionado anteriormente. Esta sección aborda por qué el producto tensorial representa una carga tan pesada y cuáles son las investigaciones más recientes para reducirlo o eliminarlo.

Otro aspecto importante aquí es la suavidad de las curvas. El resultado del cálculo atómico es la superficie de energía potencial, que es un mapa de energía. Si esta superficie es irregular, los cálculos de las fuerzas oscilan bruscamente y la simulación se rompe. Esta sección explica el proceso mediante el cual los nuevos métodos intentan lograr tanto velocidad como suavidad.

El producto tensorial es una operación matemática que combina valores direccionales de acuerdo con principios físicos. Los modelos equivariantes manejan con precisión la información direccional mediante esta operación. El problema radica en que, a medida que se incrementa la precisión direccional, la carga de esta operación aumenta con rapidez. El tiempo de cálculo y la memoria requerida se disparan simultáneamente. En sistemas grandes o en simulaciones de larga duración, esta carga se convierte en un obstáculo.

Describamos el producto tensorial en términos más sencillos. Al combinar información que contiene dirección, las reglas son complejas. No basta con sumar o multiplicar simplemente dos valores direccionales. Es necesario emparejarlos y combinarlos de acuerdo con las leyes de la física. El producto tensorial realiza este complejo emparejamiento con precisión. A cambio de dicha exactitud, el número de combinaciones aumenta rápidamente según la precisión de la dirección. Por ello, incluso un ligero aumento en la precisión incrementa la carga de cálculo.

Surgen investigaciones en diversas direcciones para superar este obstáculo. Una de ellas consiste en crear programas especializados que calculen el producto tensorial con mayor rapidez. Un estudio reciente aceleró significativamente los cálculos de dinámica molecular mediante este enfoque. Dicho estudio se publicó en una revista científica en 2026. Es una vía que optimiza los cálculos mientras mantiene el producto tensorial en sí.

5.1 Nuevos diseños que prescinden por completo del producto tensorial

Otra dirección consiste en no utilizar en absoluto el producto tensorial. Un estudio denominado Geodite ha optado por este camino. El título de este trabajo es potencial interatómico equivariante sin producto tensorial. Reemplaza la pesada operación de combinar direcciones por otro método. Al mismo tiempo, preserva rigurosamente la simetría rotacional de acuerdo con los principios fundamentales.

Examinemos brevemente el principio por el cual este método reemplaza el producto tensorial. En lugar de la pesada operación de emparejamiento, establece un sistema de coordenadas locales. Es decir, define un marco de referencia direccional alrededor de cada átomo. Al procesar la información sobre este marco, es posible conservar la dirección mediante cálculos ligeros. Mantiene intacta la simetría y a la vez aligera la carga.

Los logros señalados por este estudio son dos. Uno es la velocidad. El resumen publicado indica que es de tres a cinco veces más rápido que los modelos de precisión similar. El otro es el alcance de la precisión. Logró una precisión comparable a la de los modelos precedentes en la predicción de estabilidad de materiales, conductividad térmica, vibraciones de red y dinámica molecular a escala de nanosegundos. En otras palabras, incrementó notablemente la velocidad sin sacrificar la exactitud. Este trabajo es una investigación preliminar que no ha pasado por revisión por pares. Por lo tanto, se considera una prometedora tecnología candidata y no un estándar establecido.

Eliminar el producto tensorial no solo mejora la velocidad. También contribuye a la suavidad de las curvas. Cuando se trunca la precisión direccional en el producto tensorial, pueden aparecer ondulaciones microscópicas en la superficie. Estas ondulaciones provocan saltos erráticos en el cálculo de las fuerzas. Los diseños que no utilizan el producto tensorial están concebidos para evitar tales ondulaciones. La curva se mantiene suave incluso en los intervalos donde los enlaces se forman y se rompen.

Esta suavidad es crucial en los cálculos de materiales para cohetes. A altas temperaturas, los átomos vibran intensamente y los enlaces se rompen con frecuencia. Si la curva es rugosa, las fuerzas se desvían de manera abrupta en esos intervalos y la simulación colapsa. Una curva suave estabiliza los cálculos incluso en condiciones extremas. Por ello, la suavidad es una propiedad tan importante como la precisión.

Existen otras investigaciones con objetivos similares. Un estudio despliega la información direccional en una cuadrícula esférica y luego la combina mediante una red neuronal ligera. Este método también es más rápido que el producto tensorial y preserva la capacidad expresiva. Asimismo, han surgido estudios que modifican suavemente la distancia de corte según la situación. Todos estos trabajos se encuentran en una fase preliminar de investigación. Diversos equipos están probando diferentes caminos hacia el mismo objetivo.

Examinemos un poco más en detalle por qué la suavidad es tan importante. La fuerza proviene del gradiente de la superficie de energía. Si la superficie es suave, el gradiente también cambia de manera gradual. Si hay pequeñas ondulaciones en la superficie, el gradiente salta repentinamente. Una fuerza errática empuja los átomos en direcciones incorrectas y arruina la simulación. Por eso, una superficie suave es tan necesaria como una energía precisa. Aunque el error aparente sea pequeño, si la superficie es rugosa, los problemas se agravan en cálculos prolongados.

El significado de esta tendencia es claro: reducir la carga computacional manteniendo la precisión de los modelos equivariantes. Cuando el cálculo se vuelve más ligero, es posible observar sistemas más grandes durante periodos más prolongados. Esto ayuda de forma directa a monitorizar durante largo

tiempo materiales grandes, como los componentes de cohetes. No obstante, dado que las pruebas empíricas aún se están acumulando, es necesario seguir de cerca su validación.

6. Modelado de densidad electrónica y poda estructurada para el funcionamiento en equipos ligeros

Los modelos grandes son precisos, pero conllevan una gran carga y exigen equipos de cálculo de gran envergadura. Esta sección aborda métodos para aligerar los modelos grandes preservando al mismo tiempo las leyes de la física. Examinamos dos ideas: una es la poda, que consiste en recortar las vías de cálculo menos importantes; la otra es un método para incorporar al modelo indicios de la distribución de los electrones. El núcleo de esta sección es hacer el modelo más ligero sin realizar cortes indiscriminados. En los cálculos atómicos existen muchas propiedades físicas que deben preservarse: la simetría, la dirección de las fuerzas y la coherencia entre la energía y las fuerzas son algunas de ellas. Se explica cómo reducir el modelo preservando estas propiedades.

La poda (Pruning) es un método para reducir las vías menos importantes en una red neuronal. Las vías que transportan información dentro de una red neuronal se denominan canales. No todos los canales tienen la misma importancia. Si se reducen los canales menos utilizados, el cálculo se vuelve más ligero. Es similar a sacar del equipaje las cosas que apenas se usan para aligerar la maleta.

El problema es que, si se corta cualquier canal al azar, las leyes físicas se ven comprometidas. En los modelos equivariantes, los canales están estrechamente vinculados con la dirección. Si se cortan de forma dispersa canales pertenecientes al mismo grupo direccional, la dirección de la fuerza se distorsiona. Conviene aclarar aquí un malentendido: si el modelo calcula primero la energía y obtiene la fuerza a partir de su gradiente, la relación entre energía y fuerza se preserva intacta aunque se reduzcan los canales. La poda no rompe directamente esta relación. En cambio, pueden deteriorarse la precisión en el manejo de las direcciones y la exactitud predictiva. Por ello, la poda debe realizarse con sumo cuidado.

6.1 Poda estructurada que preserva la coherencia física

La poda estructurada (Structured Pruning) es un método que recorta los canales en unidades de grupos direccionales. En lugar de cortar canales aislados individualmente, los maneja por bloques completos. De este modo, el equilibrio direccional no se rompe. También se gestiona para mantener la proporción entre canales escalares y canales vectoriales. Los escalares son valores sin dirección y los vectores son valores con dirección. Mantener esta proporción es indispensable para que la simetría y la dirección de las fuerzas permanezcan intactas.

Ilustremos por qué se debe recortar por bloques completos. La aguja de una brújula solo tiene sentido si señala el norte, el sur, el este y el oeste de forma conjunta. Si se deja solo el eje norte-sur y se corta el eje este-oeste, la dirección se distorsiona. Lo mismo ocurre con los canales direccionales en los modelos equivariantes. Los valores dentro de un mismo grupo deben permanecer juntos para indicar la dirección de manera correcta. Si solo se corta una parte de forma dispersa, la dirección de la fuerza se desvía. Por ello, al podar, se debe tomar necesariamente el grupo como unidad.

La decisión de qué canales cortar se determina según su importancia. Se mide cuánto contribuye cada canal al resultado final y se reducen primero los canales con menor contribución. También en este caso se toma como unidad el grupo direccional. Se evalúan conjuntamente los valores dentro de un mismo grupo. De esta manera, se eliminan únicamente los elementos superfluos mientras se preserva la estructura fundamental que transporta la información.

Recientemente, un equipo de investigación aplicó este método a un potencial fundacional. El título de este estudio es modelo atómico equivariante SO(3) compacto creado mediante poda estructurada. El resumen publicado indica que los parámetros se redujeron varias veces. Los parámetros son el número de valores que aprende el modelo. La cantidad de cálculo requerida para el preentrenamiento también se redujo notablemente. Al mismo tiempo, señala que, tras el ajuste fino, los errores de energía y fuerza incluso disminuyeron. Se trata de un estudio preliminar publicado en 2026. Dado que no ha pasado por revisión por pares, no se considera un resultado definitivo.

El objetivo perseguido por este estudio es ejecutar cálculos precisos incluso en equipos ligeros. Los modelos grandes suelen requerir múltiples equipos de alto rendimiento. Si se reduce adecuadamente el modelo, es posible realizar cálculos similares con equipos más pequeños. Al disminuir la barrera de entrada para la investigación, más investigadores pueden participar en el cálculo de materiales. No obstante, aún debe comprobarse en el futuro hasta qué tipo de equipos es viable.

Esta dirección también tiene un gran valor en la investigación de materiales para cohetes. Para simular entornos extremos durante largos periodos, es necesario repetir los cálculos muchas veces. Si el modelo es ligero, se pueden ejecutar más experimentos con el mismo equipo. Esto ayuda a filtrar con rapidez los materiales candidatos. La clave radica en ganar velocidad manteniendo la precisión.

Existe otro método con un objetivo similar a la poda: transferir el conocimiento de un modelo grande a un modelo pequeño. Este método se denomina destilación de conocimiento. El modelo grande bien entrenado actúa como maestro del modelo pequeño. El modelo pequeño aprende imitando las respuestas del maestro. Así, a pesar de su tamaño reducido, alcanza un rendimiento cercano al del modelo grande. Tanto la poda como la destilación de conocimiento son vías para aligerar y trasladar modelos grandes.

Lo que se debe tener en cuenta es no omitir la validación física. Los modelos ligeros obtenidos mediante la poda también deben verificarse rigurosamente. Se deben comprobar conjuntamente la coherencia entre energía y fuerza, la dirección de las fuerzas, la simetría y la indicación de incertidumbre. Solo cuando se preservan estas propiedades se puede confiar en un modelo ligero. Un modelo que solo es rápido resulta peligroso en cálculos bajo condiciones extremas.

6.2 Métodos para añadir indicios de densidad electrónica

Al reducir un modelo, existe el riesgo de que también disminuya su capacidad expresiva. Una idea para compensar esta pérdida es incorporar la densidad electrónica. La densidad electrónica es un mapa que indica cómo se distribuyen los electrones en el espacio. Es la magnitud física fundamental que trata originalmente la teoría del funcional de la densidad (DFT). La propuesta consiste en introducir este indicio directamente en el modelo para suplir la falta de capacidad expresiva. Este método de combinar informaciones de distinta naturaleza se denomina enfoque híbrido.

Repasemos brevemente los tipos de enlaces químicos. El enlace iónico es aquel en el que los átomos se unen al transferir y recibir electrones; la sal común está formada por este tipo de enlace. El enlace metálico es aquel en el que los electrones fluyen libremente entre múltiples átomos; gracias a este enlace, los metales conducen bien la electricidad. El enlace covalente es aquel en el que los átomos comparten electrones entre sí. De este modo, la distribución y disposición de los electrones varían según el tipo de enlace.

La información de la densidad electrónica ayuda a distinguir los tipos de enlaces. Dentro de los materiales existen diversos tipos de enlaces: los enlaces iónicos, metálicos y covalentes difieren entre sí. Al observar dónde y en qué cantidad se concentran los electrones, es posible reconocer estas

diferencias. Incorporar este indicio en el modelo permite captar con mayor precisión la naturaleza de los enlaces.

Se considera que este método también resulta útil para recuperar los efectos de largo alcance. La influencia de los electrones se extiende más allá de los vecinos inmediatos. Al aligerar y reducir el modelo, es fácil pasar por alto estos efectos a larga distancia. La expectativa es que, al utilizar adecuadamente los indicios de densidad electrónica, se puedan capturar en cierta medida estas influencias lejanas. Sin embargo, incorporar la densidad electrónica no resuelve de forma automática las fuerzas electrostáticas y de dispersión a larga distancia. Qué densidad utilizar, de qué manera incorporarla y cómo validarla son cuestiones sobre las que aún se están acumulando investigaciones. Por ello, es más apropiado entender esta dirección como una propuesta y no como una solución definitiva.

En los materiales para cohetes suele haber una mezcla compleja de distintos tipos de enlaces. Se utilizan conjuntamente metales y cerámicas, y se forman películas de óxido. Para tratar estos materiales, es necesario distinguir con precisión la naturaleza de los enlaces. El método de incorporar la densidad electrónica facilita esta diferenciación y puede potenciar los cálculos de materiales en condiciones extremas.

También se abre la posibilidad de realizar cálculos a gran escala. Recientemente, un equipo de investigación simuló una aleación multicomponente a una escala de un millón de átomos. Emplearon un potencial interatómico capaz de manejar múltiples elementos para abordar sistemas de gran tamaño. Este trabajo es una investigación preliminar publicada en 2026. Las aleaciones de alta entropía son candidatas fundamentales como materiales para altas temperaturas en cohetes. Cuanto mayor sea el sistema que se pueda modelar, más fielmente se podrá observar el comportamiento real de los componentes.

Reflexionemos sobre lo expuesto en esta sección a la luz de la realidad del desarrollo de cohetes. En muchas empresas e institutos de investigación de cohetes, los recursos computacionales no son abundantes. Si los modelos grandes solo pueden ejecutarse en múltiples equipos de alta gama, la barrera de entrada es elevada. Cuando los modelos se vuelven más ligeros, más equipos pueden acceder al cálculo atómico. Esto permite descartar con rapidez materiales candidatos y reducir el número de muestras para pruebas físicas reales. Este ciclo iterativo entre cálculo y experimentación acorta el tiempo de desarrollo.

Las dos ideas de esta sección apuntan al mismo objetivo: aligerar el cálculo preservando la precisión. La poda elimina lo superfluo y los indicios de densidad electrónica suplen la falta de capacidad expresiva. Al utilizar ambas conjuntamente, es posible construir modelos prácticos y útiles incluso con equipos pequeños. Esta dirección se encuentra aún en desarrollo y en una etapa de acumulación de validaciones.

Resumen

Este capítulo fue la historia de las reglas para manipular átomos mediante computadoras. Esta regla es el potencial interatómico. Desde hace mucho tiempo, los científicos tuvieron que elegir entre un método preciso y un método rápido. La DFT es precisa pero lenta, y los campos de fuerza empíricos son rápidos pero inexactos. El potencial interatómico de aprendizaje automático es una herramienta para derribar el muro entre ambos.

Repasamos brevemente este recorrido. Al principio, los humanos convertían el entorno de los átomos en números. Luego, el modelo aprendió esos números por sí mismo. Después, se hizo que los átomos intercambiaran información entre sí. Ahora, han surgido incluso grandes modelos que han aprendido

previamente una amplia variedad de materiales. En cada paso, se redujo la intervención humana y se incrementó el poder de los datos. Esta corriente es el eje principal del avance de los MLIP.

Los primeros MLIP convertían la forma del entorno atómico en números diseñados por humanos. En este método de descriptores se encontraban BPNN, los potenciales de aproximación gaussiana y DeePMD. Estos lograron grandes avances, pero también mostraron limitaciones. A medida que aumentaba el número de elementos, la carga computacional crecía y resultaba difícil capturar los efectos de largo alcance.

La siguiente generación son las redes neuronales de grafos equivariantes. Este método considera los átomos como puntos y las relaciones de vecindad como líneas para intercambiar información. Al incorporar en el modelo las simetrías correspondientes a rotaciones y traslaciones, se ahorraron datos. Al pasar de SchNet a NequIP y MACE, la precisión y la eficiencia aumentaron juntas. Los potenciales fundacionales como CHGNet, MatterSim y GRACE se consolidaron como grandes modelos que han aprendido previamente una amplia variedad de materiales.

Incluso los modelos precisos tienen debilidades. Pueden equivocarse en situaciones desconocidas que no han aprendido. Por eso se necesitan el aprendizaje activo y la cuantificación de la incertidumbre para que el modelo avise por sí mismo cuando no sabe algo. El entorno de los cohetes siempre es extremo y presenta muchos eventos desconocidos. Solo con estos mecanismos de seguridad se puede confiar en los cálculos de materiales en condiciones extremas.

Las investigaciones recientes persiguen al mismo tiempo la velocidad, la suavidad y la ligereza. Reducen o eliminan los productos tensoriales para acelerar los cálculos. Reducen el modelo mediante poda estructural y complementan la capacidad expresiva con pistas de densidad electrónica. Gran parte de estos métodos son todavía investigaciones iniciales que no han pasado por la revisión por pares. Son prometedores, pero aún es pronto para considerarlos estándares consolidados.

Resumimos en cinco puntos lo que sería bueno que el lector general recuerde al leer este capítulo. En los cálculos atómicos existió un prolongado dilema entre precisión y rapidez. El aprendizaje automático tendió un puente para mitigar este dilema. Preservar la simetría es la clave de un buen modelo. Por muy bueno que sea un modelo, solo es seguro si avisa por sí mismo cuando no sabe algo. Y gran parte de las tecnologías más recientes son aún investigaciones iniciales en proceso de verificación.

Estos cinco puntos se aplican igualmente a los capítulos siguientes. Crear estructuras cristalinas, diseñar aleaciones y manipular cerámicas son tareas que se apoyan por completo en estos cálculos atómicos. Sin cálculos atómicos precisos, el diseño construido sobre ellos también se tambalea. Por ello, el contenido de este capítulo sirve como piedra angular de todo el libro.

Dejamos también una perspectiva equilibrada. Los MLIP son herramientas poderosas, pero no son una solución para todo. Si la calidad de los datos de entrenamiento es deficiente, las predicciones también empeoran. Si se descuida la validación, uno se deja engañar por errores verosímiles. Por eso, esta tecnología no reemplaza el juicio humano ni la experimentación. Los cálculos reducen los candidatos y la confirmación final queda a cargo de los experimentos. Esta colaboración es el camino para llevar nuevos materiales al mundo de forma segura.

El objetivo de todas estas tecnologías converge en uno solo. Consiste en probar materiales para cohetes y nuevos materiales dentro de la computadora a gran escala y durante largos periodos, de manera precisa y segura. Filtrar los candidatos antes de fabricar componentes reales ahorra dinero y tiempo. A partir del siguiente capítulo, pasaremos a la historia de cómo crear directamente nuevos materiales sobre la base de estos sólidos cálculos atómicos.

Fuentes y referencias relacionadas

Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. Physical Review, 140, A1133-A1138. https://doi.org/10.1103/PhysRev.140.A1133

Behler, J., & Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. Physical Review Letters, 98, 146401. https://doi.org/10.1103/PhysRevLett.98.146401

Bartók, A. P., Payne, M. C., Kondor, R., & Csányi, G. (2010). Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. Physical Review Letters, 104, 136403. https://doi.org/10.1103/PhysRevLett.104.136403

Bartók, A. P., Kondor, R., & Csányi, G. (2013). On representing chemical environments. Physical Review B, 87, 184115. https://doi.org/10.1103/PhysRevB.87.184115

Zhang, L., Han, J., Wang, H., Car, R., & E, W. (2018). Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. Physical Review Letters, 120, 143001. https://doi.org/10.1103/PhysRevLett.120.143001

Schütt, K. T., Sauceda, H. E., Kindermans, P. J., Tkatchenko, A., & Müller, K. R. (2018). SchNet: A deep learning architecture for molecules and materials. Journal of Chemical Physics, 148, 241722. https://doi.org/10.1063/1.5019779

Gasteiger, J., Groß, J., & Günnemann, S. (2020). Directional message passing for molecular graphs (DimeNet). International Conference on Learning Representations. https://arxiv.org/abs/2003.03123

Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., & Kozinsky, B. (2022). E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials (NequIP). Nature Communications, 13, 2453. https://doi.org/10.1038/s41467-022-29939-5

Batatia, I., Kovács, D. P., Simm, G. N. C., Ortner, C., & Csányi, G. (2022). MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. Advances in Neural Information Processing Systems, 35, 11423-11436. https://arxiv.org/abs/2206.07697

Batatia, I., Batzner, S., Kovács, D. P., Musaelian, A., Simm, G. N. C., Drautz, R., Ortner, C., Kozinsky, B., & Csányi, G. (2025). The design space of E(3)-equivariant atom-centred interatomic potentials. Nature Machine Intelligence, 7, 56-67. https://doi.org/10.1038/s42256-024-00956-x

Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., & Ceder, G. (2023). CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. Nature Machine Intelligence, 5, 1031-1041. https://doi.org/10.1038/s42256-023-00716-3

Yang, H., Hu, C., Zhou, Y., Liu, X., Shi, Y., Li, J., et al. (2024). MatterSim: A deep learning atomistic model across elements, temperatures and pressures. arXiv:2405.04967. https://doi.org/10.48550/arXiv.2405.04967

Deringer, V. L., Caro, M. A., & Csányi, G. (2021). Machine learning interatomic potentials as emerging tools for materials science. Advanced Materials, 33, 1902765. https://doi.org/10.1002/adma.201902765

Shapeev, A. V. (2016). Moment tensor potentials: A class of systematically improvable interatomic potentials. Multiscale Modeling and Simulation, 14, 1153-1173. https://doi.org/10.1137/15M1054183

Podryabinkin, E. V., & Shapeev, A. V. (2017). Active learning of linearly parametrized interatomic potentials. Computational Materials Science, 140, 171-180. https://doi.org/10.1016/j.commatsci.2017.08.031

Lysogorskiy, Y., Bochkarev, A., & Drautz, R. (2026). Graph atomic cluster expansion for foundational machine learning interatomic potentials (GRACE). npj Computational Materials, 12, 114. https://doi.org/10.1038/s41524-026-01979-1

Reschützegger, T., Aykent, S., Perin, G. J., Nunes, B. H., Cipcigan, F., Ferreira, R. N. B., Steiner, M., & Thiemann, F. L. (2026). Equivariant interatomic potentials without tensor products (Geodite). arXiv:2601.15492. Preimpresión sin revisión por pares. https://arxiv.org/abs/2601.15492

Wang, C., Hu, S., Tan, G., & Jia, W. (2026). Compact SO(3) equivariant atomistic foundation models via structural pruning. arXiv:2605.08885. Preimpresión sin revisión por pares. https://arxiv.org/abs/2605.08885

Tan, C. W., Descoteaux, M. L., Kotak, M., de Miranda Nascimento, G., Kavanagh, S. R., Zichi, L., Wang, M., Saluja, A., Hu, Y. R., Smidt, T. E., Johansson, A., Witt, W. C., Kozinsky, B., & Musaelian, A. (2026). High-performance training and inference for deep equivariant interatomic potentials. Digital Discovery, 5, 1558-1567. https://doi.org/10.1039/d5dd00423c

Miklaucic, N., Wei, L., Dong, R., Fu, N., Omee, S. S., Li, Q., Dey, S., Fung, V., & Hu, J. (2025). Facet: Highly efficient E(3)-equivariant networks for interatomic potentials. arXiv:2509.08418. Preimpresión sin revisión por pares. https://arxiv.org/abs/2509.08418

Han, K., Cong, H., Deng, B., & Barati Farimani, A. (2026). Smooth dynamic cutoffs for machine learning interatomic potentials. arXiv:2601.21147. Preimpresión sin revisión por pares. https://arxiv.org/abs/2601.21147

Xu, H., Cui, T., Tang, C., Ma, J., Zhou, D., Li, Y., Gao, X., Gong, X., Ouyang, W., Zhang, S., & Su, M. (2025). Evidential deep learning for interatomic potentials. Nature Communications, 16, 937. https://doi.org/10.1038/s41467-025-67663-y

Kulichenko, M., Barros, K., Lubbers, N., Li, Y. W., Messerly, R., Tretiak, S., Smith, J. S., & Nebgen, B. (2023). Uncertainty-driven dynamics for active learning of interatomic potentials. Nature Computational Science, 3, 230-239. https://doi.org/10.1038/s43588-023-00406-5

Perez, D., Subramanyam, A. P. A., Maliyov, I., & Swinburne, T. D. (2025). Uncertainty quantification for misspecified machine learned interatomic potentials. npj Computational Materials, 11, 263. https://doi.org/10.1038/s41524-025-01758-4

Yang, Z., Wang, X., Li, Y., Lv, Q., Chen, C. Y.-C., & Shen, L. (2025). Efficient equivariant model for machine learning interatomic potentials. npj Computational Materials, 11, 49. https://doi.org/10.1038/s41524-025-01535-3

Shuang, F., Ying, P., Liu, K., Wei, Z., Liu, F., Fan, Z., Jiang, M., & Dey, P. (2026). Benchmarking chemically scalable machine-learning interatomic potentials for large-scale simulations of multicomponent alloys. arXiv:2604.01642. Preimpresión sin revisión por pares. https://arxiv.org/abs/2604.01642

Arnold, J. E., Woodgate, C. D., Hafizi, R., Harris, M. J., Mottura, A., García Fuentes, G., & Gurrutxaga-Lerma, B. (2026). Thermodynamics and phase stability of the AlTiCrMoW high-entropy alloy simulated using machine-learned interatomic potentials. Physical Review Materials, 10, 073605. https://doi.org/10.1103/5tz1-zg9z

Chiang, Y., Kreiman, T., Zhang, C., Kuner, M. C., Weaver, E. J., Amin, I., Park, H., Lim, Y., Kim, J., Chrzan, D., Walsh, A., Blau, S. M., & Asta, M. (2025). MLIP Arena: Advancing fairness and transparency in machine learning interatomic potentials via an open, accessible benchmark platform. Advances in Neural Information Processing Systems, 38. https://arxiv.org/abs/2509.20630

Capítulo 2. Predicción de estructuras cristalinas (CSP) y diseño inverso de materiales basados en IA generativa y modelos de difusión (Diffusion Models)

Introducción

Cuando fabricamos un objeto, por lo general comenzamos eligiendo el material. Supongamos que vamos a hacer una olla. Primero elegimos el hierro o el aluminio. Luego definimos el grosor y la forma. Por último, la ponemos al fuego para comprobar qué propiedades presenta.

El método tradicional para buscar nuevos materiales es similar a esto. Primero se elige una sustancia ya conocida. Luego se miden sus propiedades mediante computación o experimentos. Si son buenas, se conservan; si son malas, se descartan. Este método es fiable. Sin embargo, como hay que comprobar una por una, la velocidad es lenta.

El problema radica en que el mundo de los materiales es demasiado vasto. Hay muchos tipos de elementos. El número de átomos también puede variar. Las posiciones donde se colocan los átomos y el tamaño de la red también se pueden cambiar. Si se combina todo esto, el número de posibilidades aumenta como las estrellas en el cielo. Es imposible que un ser humano llegue al final probándolas una a una.

Este capítulo trata sobre invertir la dirección del pensamiento. No se empieza eligiendo los materiales, sino definiendo primero las propiedades deseadas. Se establece que se desea un electrolito sólido con buena conductividad eléctrica. Se establece que se desea un catalizador capaz de resistir en la tobera caliente de un cohete. Entonces, la inteligencia artificial dibuja por sí misma una estructura cristalina que se ajuste a esas propiedades. A este enfoque se le denomina diseño inverso de materiales (Inverse Design).

La importancia del diseño inverso se puede comparar con la cocina. El método tradicional consiste en cocinar con los ingredientes que hay en el refrigerador. Dado que los ingredientes están fijados, la comida que se puede preparar también está limitada. El diseño inverso consiste en decidir primero qué comida se quiere comer. Luego se buscan hacia atrás los ingredientes necesarios para ese plato. Incluso se consiguen ingredientes nuevos que no estaban en el refrigerador. De este modo, sin quedar atrapado por los ingredientes de la nevera, uno se acerca a la comida que imaginaba.

Aquí, el protagonista es la inteligencia artificial generativa (Generative AI). Es fácil entenderlo si pensamos en una IA que dibuja imágenes. Comienza a partir de un conjunto de puntos borrosos. Elimina el ruido poco a poco. Al final, aparece el rostro de una persona o un paisaje. El principio del diseño inverso de materiales es el mismo. Comienza a partir de un grupo difuso de átomos. Mientras elimina el ruido, da forma a un cristal regular. A este método de eliminación de ruido se le llama modelo de difusión (Diffusion Model).

En este capítulo, examinamos paso a paso cómo los modelos de difusión crean cristales. Vemos cómo respetan las reglas de simetría. Analizamos la técnica de coincidencia de flujo (Flow Matching) para acelerar los cálculos. También examinamos nuevas ideas como el inpainting para rellenar huecos y los átomos espejismo (Mirage atoms). Vemos cómo guiar la generación hacia mejores estructuras mediante el aprendizaje por refuerzo. Por último, revisamos casos en los que este método se aplicó realmente a electrolitos sólidos para baterías, catalizadores y aleaciones aeroespaciales. Explicamos paso a paso por qué los cohetes y los nuevos materiales necesitan esta tecnología.

1. Del método de elegir materiales primero al diseño inverso a partir de las propiedades

Lo que se aborda en esta sección es cómo ha cambiado la dirección de la búsqueda de materiales. El método antiguo elige primero el material y comprueba sus propiedades después. El nuevo método define primero las propiedades y crea el material después. Examinamos por qué es necesaria esta transición y qué significado tiene para los cohetes y los nuevos materiales.

Al método tradicional se le denomina cribado directo de propiedades (Forward Property Screening). El término 'directo' significa avanzar desde el material hacia las propiedades. El cribado es la tarea de filtrar múltiples candidatos. Los investigadores extraen candidatos de bases de datos de cristales ya conocidas. Existen dos repositorios representativos: el Materials Project y la Base de Datos de Estructuras Cristalinas Inorgánicas (ICSD).

Tras extraer los candidatos, se calculan sus propiedades. En este punto se utiliza un cálculo físico llamado teoría del funcional de la densidad (Density Functional Theory, DFT). Este cálculo resuelve el movimiento de los electrones dentro de los átomos para obtener la energía y las propiedades. Los resultados son precisos. Sin embargo, la cantidad de cálculo es enorme. Las computadoras deben funcionar durante mucho tiempo. Por ello, el número de candidatos que se pueden examinar al día es reducido.

La limitación fundamental de este método radica en que está atrapado en una jaula estrecha. El examen solo se realiza sobre sustancias ya conocidas. Los materiales nuevos que no están en la base de datos ni siquiera entran como candidatos desde el principio. Aunque existan sustancias desconocidas con propiedades sobresalientes, quedan fuera de alcance. Así, el método tradicional tiene poca capacidad para crear cosas nuevas.

El diseño inverso de materiales rompe esta jaula. Invierte el orden. Primero se establecen las propiedades objetivo. Se define que debe tener una alta conductividad iónica. Se define que debe ser termodinámicamente estable. A continuación, la inteligencia artificial dibuja directamente una estructura que satisfaga esas condiciones. Incluso genera estructuras que no existían en las bases de datos. El centro de gravedad de la búsqueda de materiales pasa del examen a la creación.

Veamos un ejemplo de cuán grande es el número de combinaciones posibles. En la naturaleza se pueden utilizar alrededor de 90 tipos de elementos. Supongamos que colocamos solo diez átomos dentro de una caja. Elegimos qué elemento poner en cada posición. La posición de los átomos también se puede variar ligeramente. Considerando únicamente esta combinación, el número resultante es astronómico. Surgen más candidatos que granos de arena en la playa. Si un ser humano tuviera que calcularlos uno por uno, no terminaría ni en toda una vida.

El tiempo necesario para cada cálculo tampoco es desdeñable. Los cálculos DFT tardan más a medida que aumenta el número de átomos. Si el número de átomos se duplica, el cálculo se multiplica varias veces. Para resolver una sola estructura grande, la computadora debe funcionar durante varios días. Por eso no se puede calcular una gran cantidad de candidatos a ciegas. Hay que seleccionar bien los candidatos con antelación. La inteligencia artificial proporciona la sabiduría necesaria para discernir qué candidatos calcular primero.

Dibujar un cristal implica definir tres cosas al mismo tiempo. Primero se determina la red (Lattice). La red es la forma y el tamaño de la caja que contiene los átomos. A continuación, se determinan las posiciones de los átomos: en qué lugar dentro de la caja se ubica cada átomo. Por último, se determina el tipo de átomo: qué elemento ocupa esa posición. Estos tres factores deben coincidir entre sí para

convertirse en un cristal auténtico. Si tan solo uno de ellos desentona, los átomos colisionan o la estructura se derrumba.

Resulta difícil manejar estos tres elementos a la vez. La red consiste en valores continuos con longitudes y ángulos. Las posiciones atómicas también son valores continuos. Sin embargo, los tipos de átomos son valores discretos. No existe un elemento que sea como 'medio cloro' después del cloro. De esta manera, hay que procesar conjuntamente valores con propiedades tan diferentes. Los modelos de difusión logran esta difícil tarea: restauran a partir del ruido tanto los valores continuos como los valores discretos de forma simultánea.

Expliquemos un poco más el principio de los modelos de difusión. El entrenamiento tiene dos direcciones. En primer lugar, en el proceso directo, se añade ruido poco a poco a un cristal limpio. Si se añade repetidamente, el cristal se transforma en un conjunto desordenado de puntos. La inteligencia artificial aprende a revertir este proceso. Practica prediciendo el cristal original a partir del estado ruidoso. Una vez completado este entrenamiento, es capaz de dar forma a nuevos cristales a partir de puntos desordenados.

La fuerza de este método reside en su capacidad para aprender bien a partir de los datos. Se muestran al modelo decenas de miles de cristales conocidos en el mundo. El modelo desarrolla una intuición sobre qué disposiciones son plausibles. Aprende las distancias que los átomos deben mantener entre sí. Aprende qué elementos suelen ubicarse en qué posiciones. Con esa intuición aprendida, propone nuevos cristales. Incluso sin que los humanos redacten cada regla una por una, el modelo las asimila por sí mismo.

El diseño inverso no se limita únicamente a los cristales regulares. También existen materiales con átomos desordenados, como el vidrio. A estos materiales se les llama amorfos (Amorphous). Un estudio publicado en 2025 generó estructuras amorfas utilizando modelos de difusión. Las creó miles de veces más rápido que los métodos de cálculo tradicionales. Produjo diversas estructuras variando la composición y las condiciones del proceso. Esto demuestra que el alcance del diseño inverso se extiende también más allá de los cristales.

Los modelos de diseño inverso también aprenden reglas químicas. Cada átomo tiene un número de enlaces que puede formar con sus vecinos. A este número se le llama valencia (Valence). Si se viola la valencia, surgen estructuras imposibles. Un estudio publicado en 2026 incorporó estas reglas de valencia en el proceso de generación. Como resultado, seleccionó únicamente sustancias viables en la realidad. Continúa la tendencia de integrar el sentido común de la física y la química en los modelos.

Esta transición tiene un gran significado para los cohetes y los nuevos materiales. Los cohetes exigen materiales capaces de soportar entornos extremos. La cámara de combustión es abrasadora. La tobera sufre la erosión de gases a alta velocidad. Las baterías deben ser ligeras y seguras. Los materiales que satisfacen todas estas condiciones son escasos en la naturaleza. Por ello se necesita el diseño inverso para dibujar desde cero lo que no existe. La inteligencia artificial recorre en nuestro lugar el espacio de materiales que los humanos nunca han explorado.

2. Modelos generativos que respetan las reglas de simetría cristalina

Lo que se aborda en esta sección son las reglas de simetría de los cristales. Un cristal no es un montón de átomos arrojados al azar en cualquier lugar. Es un orden donde los átomos se repiten según reglas precisas. Si la inteligencia artificial desconoce estas reglas, genera cristales falsos. Examinamos cómo incorporar la simetría dentro del modelo dividiéndolo en dos generaciones de modelos.

Para comprender los cristales, primero debemos entender el concepto de simetría. La simetría es la propiedad por la cual la forma permanece idéntica independientemente de ciertas transformaciones. Si al doblar un papel por la mitad ambos lados coinciden, existe simetría bilateral. En los cristales, estas simetrías se superponen en múltiples capas: permanecen iguales al desplazarse lateralmente, al rotarse o al reflejarse.

La clasificación que reúne todas estas reglas de simetría se denomina grupo espacial (Space Group). En los cristales tridimensionales existen exactamente 230 grupos espaciales. Cada cristal pertenece a uno de estos 230 tipos. Esta clasificación se aplica a los cristales periódicos tridimensionales ideales. Los materiales amorfos como el vidrio o los cuasicristales quedan fuera de este marco. Conocer el grupo espacial equivale a conocer todas las reglas de repetición y rotación de ese cristal.

Dentro de un grupo espacial existen posiciones determinadas donde pueden situarse los átomos. A este conjunto de posiciones se le llama posiciones de Wyckoff (Wyckoff Position). Una posición de Wyckoff no es una sola ubicación, sino un grupo de posiciones vinculadas por simetría. Si se coloca un átomo en una posición, las reglas de simetría sitúan átomos idénticos en las posiciones emparejadas. Es similar a colocar una pelota frente a un espejo, donde también aparece una pelota reflejada. Por ello, definir una sola posición determina las posiciones de múltiples átomos a la vez. El número de posiciones asociadas a este grupo se denomina multiplicidad (Multiplicity). Cuando un modelo generativo adopta esta regla, puede evitar que los átomos se superpongan entre sí.

El poder del grupo espacial radica en dibujar la totalidad a partir de poca información. En un cristal de alta simetría, solo es necesario definir unos pocos átomos. El resto lo completan automáticamente las reglas de simetría. Es fácil visualizarlo si pensamos en un copo de nieve. Las formas que se extienden en seis direcciones son idénticas. Si se dibuja una sola dirección, las otras cinco se obtienen mediante rotación. Los cristales también determinan su totalidad a partir de un fragmento tan pequeño.

Si el modelo generativo desconoce estas reglas, surgen problemas. Supongamos que la inteligencia artificial coloca los átomos uno a uno libremente. La simetría se rompe con solo un pequeño ruido de cálculo. A los cristales que casi carecen de simetría se les llama P1. P1 es un grupo espacial formal en toda regla. De hecho, existen sustancias reales que pertenecen a P1. Por tanto, no se puede calificar de falso un cristal únicamente por ser P1. El problema se presenta cuando un cristal que debería tener alta simetría colapsa en P1. En ese caso, es necesario examinar por separado la distancia interatómica, la energía y la estabilidad.

Por ello, las investigaciones más recientes intentan incorporar previamente las reglas de simetría dentro del modelo. En lugar de colocar los átomos uno por uno, siguen un orden diferente. Primero crean una pequeña unidad básica que contiene toda la simetría. A continuación, aplican las reglas de simetría para desplegar el cristal completo. De esta manera, aumenta la probabilidad de que el producto generado preserve la simetría.

Un método excelente para manejar la simetría es la propiedad llamada equivariancia (Equivariance). La equivariancia es la propiedad por la cual, si se rota la entrada, la salida rota conjuntamente. Es fácil de entender con el ejemplo de un guante. Si se le da la vuelta a un guante para la mano derecha, se convierte en un guante para la mano izquierda. La mano y el guante se invierten juntos. Si un modelo de cristal posee equivariancia, sus propiedades se mantienen idénticas aunque se rote la estructura. La simetría se preserva de forma natural sin necesidad de forzarla.

El uso de redes neuronales dotadas de esta equivariancia facilita el aprendizaje. Esto se debe a que el modelo comprende la simetría por sí mismo. No es necesario mostrarle la misma estructura desde múltiples ángulos. Aprende eficazmente incluso con pocos datos. Por eso, los modelos de generación

cristalina de última generación adoptan la equivariancia como una herramienta fundamental. Esta propiedad también resulta clave en los modelos de coincidencia de flujo que veremos más adelante.

Una revisión exhaustiva de modelos generativos de cristales publicada en 2026 sintetiza bien esta tendencia. La revisión señala que la capacidad de preservar la simetría es un requisito fundamental para un buen modelo generativo de cristales. Sostiene que los modelos que respetan la simetría generan un mayor número de candidatos estables. Esta revisión compara en paralelo los modelos de difusión y los modelos de coincidencia de flujo.

2.1 El punto de partida de la difusión para la generación de cristales (CDVAE)

La investigación sobre la difusión aplicada a la generación de cristales comenzó a partir de un modelo pionero. El nombre de este modelo es CDVAE. Examinamos paso a paso cuáles fueron sus aciertos y cuáles sus limitaciones.

CDVAE es la abreviatura de autoencoder variacional de difusión de cristales (Crystal Diffusion Variational Autoencoder). Este modelo se publicó en 2022 y abrió las puertas a este campo. Estableció el primer marco de difusión que procesa simultáneamente la red cristalina, las posiciones atómicas y los tipos de átomos.

El gran mérito de CDVAE radica en haber considerado las condiciones de contorno periódicas (Periodic Boundary Condition). Las condiciones de contorno periódicas significan que el cristal se repite indefinidamente con la misma forma. Es fácil si se piensa en el patrón de un papel tapiz. Una sola pieza se conecta arriba, abajo, a la izquierda y a la derecha para cubrir toda la pared. En los cristales, las unidades pequeñas también se repiten para formar una masa sólida.

Este modelo desplaza poco a poco las posiciones atómicas mediante un método llamado dinámica de Langevin (Langevin Dynamics). En este proceso, la dirección en la que se empujan los átomos es el gradiente de probabilidad aprendido por el modelo. Esto difiere de las fuerzas reales entre átomos o de la energía libre. Por lo tanto, este proceso por sí solo no garantiza una estructura estable. Tras establecer la estructura básica mediante la generación, esta se refina nuevamente con DFT. Es necesario distinguir estas dos etapas.

Sin embargo, CDVAE presenta claras limitaciones. No puede fijar directamente la simetría del grupo espacial. Al mover los átomos libremente, la simetría se rompe con facilidad. Como resultado, la estructura generada colapsaba a menudo en un estado P1 donde la simetría desaparecía. En los cristales donde la simetría es indispensable, se requería una verificación adicional por separado.

Otra limitación es que el número de átomos debe determinarse de antemano. Este modelo predice primero la cantidad de átomos y luego genera la estructura adaptada a ese número. Cuando la cantidad está fija, la flexibilidad disminuye. Resulta difícil manejar situaciones en las que faltan átomos o se añaden otros. Esta limitación se resuelve más adelante mediante el método de átomos espejismo.

Estas deficiencias no son defectos, sino peldaños. La tarea planteada por CDVAE se convirtió en el objetivo de los modelos de la siguiente generación. ¿Cómo preservar la simetría? ¿Cómo manejar libremente el número de átomos? Estas dos preguntas guiaron las investigaciones posteriores. Conocer las limitaciones de los modelos iniciales es el punto de partida del progreso.

2.2 Modelos de difusión condicional conscientes de la simetría

Los modelos de la siguiente generación intentaron respetar las reglas de simetría desde el principio. Para evitar confusiones, explicaremos por separado dos investigaciones con nombres similares. Las cifras de rendimiento se abordan con cautela, solo dentro de los límites verificados.

La idea central de los modelos conscientes de la simetría es simple. No colocan los átomos libremente uno por uno. Primero generan una unidad asimétrica (Asymmetric Unit). La unidad asimétrica es el conjunto mínimo de átomos que puede generar todo el cristal mediante reglas de simetría. A esta pequeña semilla se le aplican las operaciones de simetría del grupo espacial. Al pasar por esta operación, la semilla se despliega en el cristal completo.

La ventaja de este método reside en que la simetría se preserva de forma automática. Al haber desplegado la semilla con reglas de simetría, el resultado ya lleva impregnada la simetría. No es necesario ajustar los átomos por separado. Es fácil si se piensa en el recorte de papel doblado. Si se dobla una vez y se corta, al desdoblarlo ambos lados quedan idénticos. Los modelos conscientes de la simetría recortan los cristales de esta manera.

Aquí es necesario distinguir dos estudios con nombres parecidos. Uno es SymmCD. Es un modelo de difusión para la generación de cristales que preserva la simetría. Fue presentado por Levy et al. en 2025 en la Conferencia Internacional sobre Representaciones del Aprendizaje (ICLR). Es un artículo formal revisado por pares. El número de identificación de este artículo es arXiv:2502.03638.

El otro es un estudio independiente que aborda las posiciones de Wyckoff. Es un preimpreso (preprint) publicado por Ishii et al. en 2026. Un preimpreso es una investigación inicial que aún no ha pasado por la revisión por pares. Este artículo propone un método llamado WyckoffDiff-Adaptor. Su número de identificación es arXiv:2601.08115. Aunque ambos artículos tienen objetivos similares, difieren en métodos y autores. Este libro presenta ambos estudios por separado sin mezclarlos.

Entre CDVAE y los modelos conscientes de la simetría también existe un puente intermedio. Uno de ellos es DiffCSP. Este modelo genera simultáneamente la red y las coordenadas fraccionarias mediante difusión. Las coordenadas fraccionarias son posiciones relativas considerando el tamaño de la caja como 1. DiffCSP preservó la simetría mejor que CDVAE. Allanó el camino hacia los modelos de generaciones posteriores que incorporan directamente la simetría. De este modo, los modelos avanzan superando deficiencias generación tras generación.

Se debe tener cuidado al tratar con cifras de rendimiento. Aunque existen cifras que señalan una tasa de preservación de la simetría superior al 98 por ciento, esto no está confirmado en el texto original. Por lo tanto, este libro no utiliza números específicos como si fueran un rendimiento definitivo. En su lugar, se entiende como una tendencia. Los modelos conscientes de la simetría preservan la simetría mejor que los modelos anteriores. Como resultado, generan más candidatos cercanos a los cristales reales.

Los modelos conscientes de la simetría también manejan adecuadamente la generación condicional. La generación condicional (Conditional Generation) es un método en el que se proporcionan las propiedades deseadas como condiciones. Supongamos que se especifica un valor llamado banda prohibida (bandgap). La banda prohibida es el umbral de energía que determina las propiedades de un semiconductor. El modelo selecciona y genera únicamente los cristales que se ajustan a esa banda prohibida. Al establecer tales condiciones, solo se reúnen candidatos que cumplen el objetivo. Gracias a ello, el desperdicio de generar candidatos inútiles se reduce considerablemente.

Un preimpreso publicado en 2026 llevó esta generación condicional aún más lejos. Generó simultáneamente la estructura cristalina y las propiedades electrónicas. Al aprender también las propiedades electrónicas, se ajustó mejor a las propiedades objetivo. En las pruebas donde se fijaron como condiciones la banda prohibida y la energía de formación, la tasa de éxito aumentó. La proporción de candidatos estables y novedosos también se incrementó. La tendencia de abordar conjuntamente la simetría y las condiciones continúa avanzando.

Examinemos por qué esta tecnología es necesaria para los nuevos materiales. La simetría está profundamente ligada a las propiedades de la materia. Una estructura espuria con la simetría rota da lugar a predicciones erróneas de sus propiedades. Las aleaciones para la industria aeroespacial o los materiales para baterías solo tienen sentido si conducen a una síntesis real. Los modelos generativos que preservan la simetría reducen los candidatos inviables. En esa misma medida, aumenta la calidad de los candidatos que pasan a la fase experimental.

3. Flow matching para crear cristales más rápido

Los modelos de difusión son precisos, pero resultan lentos debido a su gran cantidad de pasos. El flow matching reduce los pasos para aumentar la velocidad. También veremos cómo se maneja el espacio curvo en el que habitan los cristales.

Primero se debe comprender por qué los modelos de difusión son lentos. Los modelos de difusión crean la estructura eliminando el ruido poco a poco. Solo borran una pequeña cantidad a la vez. Por ello, se requieren de cientos a miles de pasos hasta completarla. Cuando hay muchos pasos, el cálculo tarda mucho tiempo. Toma tiempo generar un solo candidato. En la búsqueda de materiales, donde es necesario extraer una gran cantidad de candidatos, este tiempo representa una carga.

El emparejamiento de flujos (Flow Matching) aborda este problema desde otro ángulo. En lugar de eliminar el ruido poco a poco, aprende un atajo. Aprende el camino directo desde el estado inicial hasta el estado objetivo. Es fácil si se piensa en la corriente de un río. Una vez fijado el curso del agua, las hojas flotan siguiendo ese cauce. El flow matching aprende precisamente este cauce por donde fluye el agua.

Cuando este cauce se expresa mediante una fórmula matemática, se convierte en una ecuación diferencial ordinaria. Una ecuación diferencial ordinaria (Ordinary Differential Equation) es una ecuación que describe la variación de un valor. Muestra cómo cambia un valor con el tiempo. El flow matching sigue esta ecuación desde el punto de inicio hasta el punto final. A diferencia de la difusión, que fluctúa aleatoriamente en cada paso, esta trayectoria no lo hace. Una vez fijado el punto inicial, sigue una ruta determinada hasta el final. Por eso el camino es directo y requiere pocos pasos. Sin embargo, el punto de inicio en sí sigue seleccionándose al azar. El hecho de que la trayectoria esté determinada no reduce automáticamente las distorsiones en los ángulos y el volumen de la red.

La diferencia entre la difusión y el flow matching puede compararse con el acto de descender de una montaña. La difusión desciende tanteando paso a paso en medio de la niebla. Es prudente, pero lenta. El flow matching desciende siguiendo un sendero de montaña previamente aprendido. Como conoce el camino, avanza con pasos firmes y rápidos. Ambos métodos llegan al destino al pie de la montaña, pero el tiempo requerido es diferente.

Al trabajar con cristales, surge la dificultad de que el espacio no es plano. La red tiene longitudes y ángulos. Las coordenadas atómicas son posiciones dentro de una caja que se repite. Este tipo de espacio no es plano como una hoja de papel, sino curvo como una esfera. Un espacio curvo se denomina variedad de Riemann (Riemannian Manifold).

Si se trata un espacio curvo como si fuera un papel plano, se producen errores. Es lo mismo que ocurre cuando un globo terráqueo se despliega en un mapa plano y las áreas se distorsionan. Las regiones polares se dibujan más grandes de lo que son en realidad. Lo mismo ocurre con los cristales. Los ángulos y el volumen de la red se deforman. Por esta razón, el flow matching utiliza desde el principio una regla adaptada al espacio curvo. Esta regla se denomina métrica de Riemann (Riemannian Metric). Al tratar el espacio curvo tal como es, la distorsión disminuye.

3.1 Búsqueda de atajos sobre espacios curvos

Examinaremos a través de modelos representativos cómo el flow matching maneja el espacio curvo. Nos centraremos en artículos de investigación verificados.

Encontrar un atajo en un espacio curvo se asemeja a volar sobre la Tierra. El atajo para ir de Seúl a Nueva York no es una línea recta sobre un mapa. Esto se debe a que la Tierra es redonda. El atajo real es una curva que sigue la superficie terrestre. Esta curva se denomina geodésica (Geodesic). La generación de cristales también se desplaza siguiendo geodésicas sobre el espacio curvo.

El primer modelo que aplicó con éxito esta idea a los cristales fue FlowMM. FlowMM es un modelo que extiende el flow matching de Riemann para adaptarlo a los cristales. Considera conjuntamente las simetrías características de los cristales. Maneja traslaciones, rotaciones, permutaciones de átomos y condiciones de contorno periódicas. Este modelo redujo drásticamente el número de pasos para encontrar materiales estables. El artículo publicado informa que halló candidatos estables con aproximadamente tres veces menos pasos que los métodos existentes.

FlowMM realiza dos tareas. Una consiste en predecir la estructura estable cuando la composición está fijada. La otra consiste en proponer simultáneamente una nueva composición y su estructura. La primera tarea se denomina predicción de la estructura cristalina (Crystal Structure Prediction). La segunda es el verdadero descubrimiento de nuevos materiales. El flow matching realiza ambas tareas con rapidez. Por ello, se emplea ampliamente en diversos ámbitos de la búsqueda de materiales.

Otro caso verificado es CrystalFlow. Esta investigación se publicó en 2025 en Nature Communications. Es un artículo formal revisado por pares. CrystalFlow genera materiales cristalinos mediante un modelo generativo basado en flujos. Genera simultáneamente la red y la disposición atómica. Demostró que el flow matching facilita el entrenamiento en comparación con los modelos de difusión.

El flow matching también se aplica a cristales orgánicos. Un preimpreso publicado en 2026 utilizó el emparejamiento de flujos de celda unitaria (Unit Cell Flow Matching). Con este método se predijeron con rapidez estructuras de cristales orgánicos. Este estudio no es un modelo exclusivo para cristales inorgánicos. Por lo tanto, se debe tener cuidado de no generalizarlo como un resultado para materiales inorgánicos. No obstante, demuestra que el flow matching se utiliza ampliamente en diversos tipos de cristales.

Veamos por qué la velocidad es importante en relación con los cohetes. El diseño inverso de materiales requiere generar y filtrar numerosos candidatos. Aunque se extraigan diez mil, es posible que solo unos pocos sean útiles. Si el tiempo necesario para generar un solo candidato es prolongado, toda la búsqueda se bloquea. El flow matching resuelve este cuello de botella. Genera más candidatos en el mismo tiempo. Cuanto más exigentes sean las condiciones, como en los materiales aeroespaciales, más valor cobra esta velocidad.

¿Cuánto reduce el flow matching el número de pasos? El artículo de FlowMM informa que redujo aproximadamente tres veces los pasos necesarios para encontrar candidatos estables. Recorre el camino que tomaban los modelos de difusión con menos pasos. El tiempo para generar un solo candidato se acorta en esa misma proporción. Cifras desmedidas como reducciones de cien veces o pasar de un día a unos pocos minutos carecen de fundamento. Con solo considerar los valores reportados, cuanto mayor sea el número de candidatos generados, más tiempo ahorrado se acumula.

El flow matching también enfrenta desafíos pendientes. Para aprender un atajo, es necesario conocer bien la distribución objetivo. Si los datos de entrenamiento son insuficientes, se aprende un camino erróneo. El cálculo en espacios curvos sigue siendo complejo. Por ello, el flow matching no reemplaza a

todos los modelos de difusión. Según el problema, se elige entre ambos métodos. Si prima la precisión, se utiliza la difusión; si prima la velocidad, se utiliza el flow matching.

4. Modificación parcial y autodeterminación del número de átomos

No siempre es necesario diseñar cristales desde cero. También existe la vía de modificar solo una parte de una estructura existente. El inpainting, que rellena vacancias y defectos, es uno de esos métodos. También existe el concepto de átomos espejismo para abordar problemas en los que el número de átomos no está fijado.

Primero se debe entender el término inpainting (Inpainting). El inpainting es una técnica que rellena partes vacías de manera natural. Se observa a menudo en la edición fotográfica. Si se borra un poste de luz en una fotografía, ese espacio queda vacío. La inteligencia artificial observa el cielo circundante y rellena el espacio vacío. Lo mismo ocurre en las estructuras cristalinas. En los cristales, rellena de forma natural las vacancias atómicas y los sitios con defectos.

Diseñar un cristal desde cero y el inpainting son procesos distintos. La generación desde cero comienza sin colocar ningún átomo. El inpainting comienza a partir de una estructura ya existente. Solo deja vacía una parte y rediseña esa sección. El resto se mantiene intacto. Ambos métodos tienen usos diferentes. Cuando se busca un material completamente nuevo, se utiliza la generación desde cero. Cuando se desea mejorar un buen material, se utiliza el inpainting.

Hay una razón por la que esta tecnología es necesaria en la investigación de materiales. En la práctica, rara vez se dibuja un cristal entero desde cero. A menudo ya se dispone de un buen cristal matriz (Host Matrix). A este se le añaden o quitan pequeñas cantidades de otros elementos para modificar sus propiedades. A esta sustitución de algunos átomos por otros se le llama dopaje (Doping). Al estado en el que falta un átomo o queda atrapado en otra posición se le denomina defecto (Defect).

El dopaje es un método que la industria de los semiconductores ha utilizado durante mucho tiempo. El silicio puro no conduce bien la electricidad. Si se le añade una cantidad diminuta de fósforo o boro, sus propiedades eléctricas cambian drásticamente. Los chips de los teléfonos inteligentes funcionan bajo este principio. El inpainting de cristales confía este diseño de dopaje a la inteligencia artificial. El modelo sugiere qué elemento introducir y en qué posición.

Este pequeño cambio transforma enormemente las propiedades. Se parece a cómo cambia el sabor de la sal al mezclarle tan solo una pizca de otra sustancia. Si se sustituye un solo átomo en el cristal, la conductividad eléctrica cambia. El rendimiento catalítico cambia. Por eso, es necesario diseñar con precisión qué poner y dónde. El inpainting ayuda en este diseño de precisión. Mantiene intacta la matriz y redibuja únicamente las posiciones que se van a cambiar.

Existe una razón por la que el inpainting es más práctico que la generación desde cero. Encontrar un buen cristal matriz lleva mucho tiempo por sí mismo. Si ya se cuenta con una matriz verificada, es más rápido mejorar sobre ella. Es como remodelar una sola habitación en lugar de construir una casa sólida desde los cimientos. Como los cimientos no se tambalean, el riesgo es menor. Por ello, los investigadores prácticos suelen recurrir con frecuencia al método de inpainting.

Para llenar los espacios vacíos, se deben observar atentamente los átomos circundantes. Si se introduce cualquier elemento al azar, la estructura se derrumba. El tamaño debe coincidir con el de los átomos vecinos. El equilibrio eléctrico también debe encajar. La inteligencia artificial aprende las reglas de la matriz y satisface estas condiciones. Mediante un método llamado enmascaramiento (Masking), oculta únicamente las posiciones a modificar. Redibuja solo las partes ocultas y mantiene fijas las demás.

4.1 Difusión de átomos espejismo (MiAD) que determina por sí misma el número de átomos

Como método para manejar libremente el número de átomos, existe un modelo de preprint llamado MiAD. Examinamos por qué resulta útil el concepto de átomos espejismo.

MiAD es la abreviatura de difusión de átomos espejismo (Mirage Atom Diffusion). Es un preprint publicado a finales de 2025. Se trata de una investigación preliminar antes de pasar por la revisión por pares. Su título oficial contiene el significado de un modelo de difusión que genera nuevos cristales desde cero. Su número de identificación es arXiv:2511.14426.

La idea central de este modelo es el átomo espejismo. Mirage significa espejismo. Un átomo espejismo es una posición virtual que puede convertirse en un átomo real o desaparecer. El modelo primero llena holgadamente la caja con posiciones virtuales. Luego, durante el proceso de difusión, determina si cada posición es un átomo real o un espacio vacío. En las posiciones que quedan como reales se asienta un elemento, y las que no, desaparecen.

La ventaja de este método radica en que no fija de antemano el número de átomos. Los modelos anteriores creaban la estructura después de fijar primero la cantidad de átomos. MiAD determina el número por sí mismo activando y desactivando posiciones virtuales. Lo que este artículo demostró realmente es la generación que determina por sí misma el número de átomos. No es un estudio que haya verificado la concentración de defectos, las posiciones de dopaje o la no estequiometría. Sin embargo, existe la posibilidad de aplicar este método a materiales con un número variable de átomos.

A los materiales cuyas proporciones de elementos no encajan en números enteros exactos se les llama compuestos no estequiométricos. Estos materiales son comunes en los materiales reales. El espacio vacío donde falta un solo átomo altera las propiedades. También se forman cúmulos de defectos donde se agrupan varios de ellos. Asimismo, existen tipos intersticiales donde los átomos se intercalan en espacios sobrantes. Este desorden microscópico determina la conductividad y la resistencia. Un método de generación que maneje libremente el número de átomos se convierte en una clave para representar dicho desorden.

MiAD reportó resultados en pruebas de referencia (benchmark). Utilizó el conjunto de pruebas ampliamente conocido como MP-20. Al indicador que representa la proporción de cristales estables, únicos y novedosos se le llama tasa SUN. El artículo reportó cierto rendimiento en este indicador. Sin embargo, este valor es una cifra en fase de preprint. Por tanto, debe tomarse como un resultado preliminar y no como un rendimiento definitivo.

Examinamos el significado que esta tecnología tiene para los cohetes y los nuevos materiales. Los materiales aeroespaciales reales no son perfectamente regulares. Siempre existen defectos en los que faltan algunos átomos o cambian de posición. Estos defectos determinan la resistencia y la conductividad. Un modelo que maneje libremente el número de átomos representa mejor estas estructuras reales. De ese modo, produce candidatos más cercanos a los resultados de laboratorio. No obstante, dado que este método se encuentra aún en fase de preprint, su capacidad de diseño de defectos debe verificarse más a fondo.

5. Guiar hacia buenos candidatos mediante aprendizaje por refuerzo

Los modelos generativos producen muchos candidatos. Sin embargo, no todos son útiles. Con el aprendizaje por refuerzo, se orienta la dirección hacia los buenos candidatos. Examinamos el principio de cómo guía el aprendizaje por refuerzo y sus mecanismos de seguridad.

El aprendizaje por refuerzo (Reinforcement Learning) es un aprendizaje que modifica las acciones en función de las recompensas. Resulta fácil si se piensa en el entrenamiento de un perro. Si lo hace bien, se le da un premio. Si lo hace mal, no se le da. El perro va aumentando las conductas que le permiten obtener premios. La inteligencia artificial también incrementa las acciones que reciben altas recompensas.

En la generación de materiales, se otorgan altas recompensas a las buenas estructuras. Se asignan puntuaciones elevadas a los cristales estables y con buenas propiedades. A las estructuras con colisiones atómicas o inestables se les dan puntuaciones bajas. El modelo cambia su hábito de generación hacia puntuaciones más altas. Por ello, a medida que se repiten los ensayos, aumenta la cantidad de candidatos viables.

Los modelos generativos y el aprendizaje por refuerzo tienen roles distintos. El modelo generativo esparce ampliamente candidatos verosímiles. El aprendizaje por refuerzo conduce esos candidatos hacia el objetivo. Es como la diferencia entre la mano que siembra las semillas y la que abre los canales de agua. Ambas manos deben actuar juntas para que broten buenos frutos. Ni la generación por sí sola ni el aprendizaje por refuerzo por sí solo bastan. Al entrelazar ambos, nos acercamos a las propiedades deseadas.

Existe una razón por la que este método es necesario. Los modelos generativos crean bien candidatos verosímiles. Sin embargo, la verosimilitud difiere de la estabilidad real. Aunque la apariencia sea buena, muchas estructuras se desmoronan termodinámicamente. Por eso, la generación por sí sola no es suficiente. Se deben sopesar conjuntamente la estabilidad, la novedad y la sintetizabilidad. El aprendizaje por refuerzo incorpora estos tres factores en la recompensa para guiar al modelo.

La sintetizabilidad es un objetivo difícil de abordar. Un material que es estable sobre el papel puede no producirse en el laboratorio. El proceso de fabricación puede ser demasiado complejo o pueden faltar las materias primas. Por ello, no basta con fijarse solo en la estabilidad. También se debe examinar si realmente se puede fabricar. La recompensa del aprendizaje por refuerzo intenta incorporar incluso estas condiciones reales. Solo así surgen candidatos capaces de pasar al laboratorio.

Existen diversos métodos para incorporar el aprendizaje por refuerzo a la generación de materiales. Como ejemplo representativo, se encuentra la técnica denominada optimización de política proximal (Proximal Policy Optimization). Esta técnica modifica la política solo poco a poco cada vez. Esto se debe a que un cambio brusco y grande desestabiliza el aprendizaje. Avanza con cautela y de forma gradual para aumentar las recompensas. Por eso mejora el rendimiento de manera estable. Esta técnica también se utiliza ampliamente para perfeccionar modelos de lenguaje.

Un preprint publicado en 2025 ilustra bien esta dirección. Realizó un ajuste fino del modelo de difusión mediante aprendizaje por refuerzo, elevando la tasa de éxito del diseño inverso. Produjo más estructuras acordes con las propiedades deseadas. También aumentó la proporción que cumplía las condiciones de estabilidad, unicidad y novedad. Esta investigación se encuentra en una etapa preliminar. Por lo tanto, para su aplicación industrial, se requieren cálculos DFT adicionales y síntesis experimental.

5.1 Mecanismos de seguridad e indicadores de rendimiento del aprendizaje por refuerzo

En el aprendizaje por refuerzo existen dos mecanismos de seguridad. Uno es el dispositivo que evita que el modelo se desvíe demasiado. El otro es el dispositivo que almacena los buenos candidatos para reutilizarlos. También observamos los indicadores que miden el rendimiento.

En el aprendizaje por refuerzo existe un peligro oculto. Si solo se persigue la recompensa, el sistema se inclina hacia estructuras anómalas. Queda atrapado en una región estrecha que otorga altas puntuaciones. Pierde la diversidad de candidatos y repite una única respuesta. A este fenómeno se le llama colapso de modo (Mode Collapse). Cuando ocurre el colapso, la exploración se detiene.

El mecanismo que previene esto es la regularización KL (KL Regularization). La divergencia KL es una vara de medir cuán diferentes son dos distribuciones. Este dispositivo sostiene al modelo ajustado para que no se aleje en exceso de la distribución aprendida originalmente. Se parece a estar atado a una banda elástica. Avanza hacia nuevas direcciones, pero si se desvía en demasía de la raíz, se tensa hacia atrás. Así, preserva la diversidad mientras busca buenos candidatos.

Otro mecanismo es la repetición de experiencias (Experience Replay). Guarda los buenos candidatos sin descartarlos. Extrae de nuevo los candidatos almacenados y los utiliza en el aprendizaje. Se parece a reunir en un cuaderno de notas las preguntas bien resueltas de un examen para volver a practicarlas. De esta manera, no olvida las buenas estructuras encontradas con dificultad. El aprendizaje continúa de forma más estable.

Un indicador utilizado frecuentemente para medir el rendimiento es la tasa SUN. SUN significa cristales estables (Stable), únicos (Unique) y novedosos (Novel). Es la proporción de candidatos que satisfacen las tres condiciones. Estable significa que no se descompone fácilmente en otras sustancias. Único significa que los productos generados no se solapan entre sí. Novedoso significa que se trata de sustancias que no estaban presentes en las bases de datos.

Para medir la estabilidad, se observa el valor denominado energía sobre el casco convexo (Energy above Hull). Este valor se calcula mediante DFT en el cero absoluto. Cuanto menor sea el valor, más estable es computacionalmente. Si el valor es cercano a 0, se mantiene bien dentro del cálculo. Sin embargo, la temperatura y la presión reales, el agua, el oxígeno y la velocidad de reacción no están contenidos aquí. También existen errores de cálculo. Por lo tanto, este valor debe interpretarse como un indicador que señala candidatos estables en el cálculo. La recompensa del aprendizaje por refuerzo utiliza este valor como un ingrediente crucial.

La recompensa coloca múltiples objetivos juntos en la balanza. Si solo se persigue la estabilidad, surgen materiales con propiedades mediocres. Si solo se persiguen las propiedades, surgen materiales difíciles de sintetizar. Por ello, se asigna un peso a la estabilidad, a las propiedades y a la novedad respectivamente. Mediante estos pesos se ajusta qué objetivo priorizar. Este ajuste de pesos es tarea del diseñador. Si se equilibra la balanza de acuerdo con los objetivos, los candidatos convergen hacia la dirección deseada.

Examinamos el significado que estos mecanismos de seguridad tienen para el diseño de materiales para cohetes. Los materiales aeroespaciales solo tienen sentido si se sintetizan realmente. Una estructura que solo es buena sobre el papel no sirve para nada. El aprendizaje por refuerzo incorpora la sintetizabilidad en la recompensa. La regularización KL y la repetición de experiencias mantienen la exploración estable. Como resultado, aumenta la proporción de candidatos viables para pasar al laboratorio.

6. Casos de aplicación práctica de la canalización de diseño inverso

¿Cómo se han utilizado las tecnologías anteriores en la investigación real? Observamos electrolitos sólidos para baterías, catalizadores y aleaciones aeroespaciales. Nos basamos en investigaciones recientes y verificadas. Examinamos cómo se engranan la generación, el cálculo y la experimentación.

Para utilizar en la práctica las tecnologías tratadas anteriormente, es necesario entrelazar diversas herramientas. Los modelos de difusión y el emparejamiento de flujo crean candidatos. La conciencia de simetría y el inpainting elevan la calidad de los candidatos. El aprendizaje por refuerzo guía los candidatos hacia el objetivo. Estas herramientas cobran fuerza cuando se conectan en un único flujo. A este flujo se le llama canalización de diseño inverso (Pipeline). Pipeline significa conectar múltiples procesos mediante tuberías.

El diseño inverso en la práctica suele seguir tres etapas. Primero, en la etapa de generación, la inteligencia artificial dibuja estructuras candidatas acordes con las propiedades objetivo. A continuación, en la etapa de verificación computacional, la teoría del funcional de la densidad (DFT) y el potencial interatómico de aprendizaje automático (MLIP) calculan la estabilidad y las propiedades. Por último, en la etapa de síntesis experimental, un laboratorio autónomo o los investigadores las fabrican y comprueban en la realidad.

Estas tres etapas no fluyen en una sola dirección. Los resultados obtenidos en los experimentos se devuelven al modelo generativo. Con esta retroalimentación, el modelo dibuja mejor los siguientes candidatos. A este método donde circula la retroalimentación se le llama aprendizaje activo (Active Learning). Un estudio publicado en 2025 mejoró sustancialmente la eficiencia del diseño inverso mediante el aprendizaje activo. Repitiendo la generación de candidatos y la verificación, se acercó a las propiedades deseadas.

Examinamos un poco más la segunda etapa, la verificación computacional. Los candidatos creados por el modelo generativo son todavía dibujos sobre el papel. Mediante cálculos se evalúa si estos dibujos son verdaderamente estables. La DFT es precisa pero lenta. Por ello, se utiliza conjuntamente el potencial interatómico de aprendizaje automático tratado en el capítulo anterior. Este método ofrece una precisión cercana a la DFT con mucha mayor rapidez. Filtra velozmente numerosos candidatos. De entre ellos, solo una minoría prometedora se vuelve a verificar con DFT.

Esta colaboración determina la velocidad de la exploración. Por muy rápida que sea la generación, si la verificación es lenta, no sirve de nada. Si la verificación se convierte en un cuello de botella, los candidatos solo se acumulan. La generación rápida y la verificación rápida deben engranarse. Por eso, las investigaciones más recientes conectan las tres etapas en un único flujo automatizado. El laboratorio autónomo que trataremos en el siguiente capítulo se encarga del final de este flujo.

6.1 Diseño de electrolitos sólidos de haluro para baterías de estado sólido

El primer caso es el diseño de electrolitos sólidos para baterías. Vemos por qué este material es importante y cómo ayuda la IA. Examinamos la cuestión centrándonos en cifras verificadas.

Una batería de estado completamente sólido (All-Solid-State Battery) es una batería que utiliza un electrolito sólido. En las baterías convencionales, el electrolito es líquido. El electrolito líquido conlleva el riesgo de inflamarse. El electrolito sólido reduce este riesgo. No obstante, que sea sólido no significa que siempre sea seguro. También puede reaccionar con el litio metálico, la humedad o los electrodos. Los basados en sulfuros y algunos basados en haluros generan productos de descomposición nocivos o presentan resistencia y grietas en la interfaz. Aun así, el interés es enorme debido a la gran ventaja de reducir el riesgo de fugas e ignición.

Las propiedades requeridas para un electrolito sólido son exigentes. Los iones de litio deben desplazarse con rapidez a través del sólido. La facilidad con la que se desplazan los iones se denomina conductividad iónica (Ionic Conductivity). Cuanto mayor sea este valor, mejor. Al mismo tiempo, debe soportar voltajes elevados. No debe ser quebradizo y debe adherirse bien bajo presión. Resulta difícil satisfacer todas estas condiciones.

Los electrolitos sólidos de haluro (Halide Solid Electrolyte) se consideran candidatos prometedores. Los haluros se refieren a elementos como el cloro, el bromo y el yodo. Esta familia es relativamente estable a altos voltajes. También presenta un buen equilibrio entre ductilidad y conductividad. Por ello, la investigación es activa. Un material representativo es un compuesto formado por litio, itrio y cloro.

La inteligencia artificial ayuda al diseño de estos materiales de diversas formas. Un estudio publicado en 2025 combinó el aprendizaje automático con cálculos de DFT. Predijo múltiples propiedades a la vez para filtrar buenos candidatos de haluros. Predijo conjuntamente la conductividad iónica, el módulo elástico y la ventana de estabilidad electroquímica. Este enfoque, que evalúa varias propiedades simultáneamente, aumentó la practicidad.

Otro estudio publicado en 2026 utilizó un método para extraer reglas a partir de datos. Identificó las características químicas clave que determinan la conductividad iónica. Con base en estas características, diseñó una nueva composición. El artículo informó que uno de los materiales diseñados mostró una conductividad iónica de aproximadamente 12 milisiemens por centímetro a temperatura ambiente. Este valor es bastante alto para esta familia de materiales. Muestra el flujo que va desde la extracción de reglas a partir de datos hasta el diseño directo.

Es difícil calcular cómo se desplazan los iones dentro de un sólido. Los iones de litio se mueven saltando de un sitio a otro. Cuanto más baja sea la barrera para este salto, mayor será la conductividad. Para calcular esta barrera, es necesario seguir el movimiento de los átomos durante un tiempo prolongado. Aquí también se utiliza el potencial interatómico de aprendizaje automático (MLIP). Un estudio de 2025 informó que este método es muy útil para el cálculo de conductores iónicos sólidos. Se proponen candidatos mediante generación y se verifica la migración iónica con este método.

Examinemos la importancia que tienen estos materiales en la industria aeroespacial y en los cohetes. Los satélites y las sondas espaciales dependen en gran medida de las baterías. Durante los períodos sin luz solar, sobreviven con la electricidad almacenada. Si la batería es inestable, toda la misión corre peligro. Las baterías de estado sólido, seguras y ligeras, reducen este riesgo. El diseño inverso mediante IA ayuda a encontrar estos materiales de batería con mayor rapidez.

6.2 Diseño para encontrar catalizadores de división del agua

El siguiente caso es el diseño de catalizadores. Veamos qué es un catalizador y cómo la IA encuentra catalizadores eficaces. Examinaremos casos recientes en colaboración con laboratorios robóticos.

Un catalizador (Catalyst) es una sustancia que ayuda a acelerar una reacción química. Apenas sufre cambios por sí mismo, al tiempo que aumenta la velocidad de reacción. Un catalizador es indispensable en los dispositivos que producen hidrógeno y oxígeno a partir del agua. Este dispositivo se denomina electrolizador de agua. Con un buen catalizador, se puede obtener más hidrógeno con la misma cantidad de electricidad.

El rendimiento del catalizador depende de la disposición de los átomos en su superficie. La reacción ocurre en la superficie del catalizador. Las sustancias reactivas se adhieren a la superficie, se transforman y se desprenden. El proceso por el cual una sustancia se adhiere a la superficie se denomina adsorción (Adsorption). Si se adhiere con demasiada fuerza, es un problema: no se desprende con facilidad y la reacción se bloquea. Si se adhiere con demasiada debilidad, también es un problema: se desprende antes de que ocurra la reacción. Un buen catalizador retiene las sustancias con la fuerza adecuada.

La fuerza de adsorción está determinada por la disposición de los átomos en la superficie. Incluso una ligera variación en la distancia entre los átomos cambia dicha fuerza. Por eso es crucial cómo se

configura la superficie. La inteligencia artificial ajusta la fuerza de adsorción modificando las posiciones de los átomos superficiales. Encuentra la disposición que retiene adecuadamente los intermediarios de reacción. Esta disposición es el secreto de un buen catalizador.

La reacción para producir oxígeno es compleja. Esta reacción se denomina reacción de evolución de oxígeno (Oxygen Evolution Reaction). Para que ocurra esta reacción, se debe aplicar un voltaje superior al teóricamente necesario. A ese voltaje adicional se le llama sobrepotencial (Overpotential). El sobrepotencial es un concepto distinto de la energía de activación, que es la barrera de velocidad de reacción. Cuanto menor sea el sobrepotencial, mejor será el catalizador. Un sobrepotencial bajo significa un menor consumo de electricidad. Por lo tanto, reducir el sobrepotencial es un objetivo primordial en el diseño de catalizadores.

Un estudio publicado en 2026 diseñó catalizadores de forma inversa mediante IA generativa. Esta investigación abordó catalizadores de alta entropía (High-Entropy Catalyst) que combinan múltiples elementos. Integró información espectroscópica, modelos generativos y equipos experimentales robóticos en un solo sistema. El robot sintetizó candidatos, midió su rendimiento y devolvió los resultados al modelo. Este ciclo de retroalimentación se ejecutó con gran rapidez.

Las cifras reportadas por este estudio demuestran su practicidad. Redujo drásticamente el tiempo necesario para fabricar y medir una muestra. Una tarea que tomaba unas 20 horas se redujo a 78 minutos. Además, disminuyó el sobrepotencial del mejor candidato encontrado en 32 milivoltios adicionales. Este caso representa una investigación académica que a la vez demuestra la automatización robótica. Se puede apreciar cuánto aumenta la velocidad cuando los modelos generativos y los experimentos se sincronizan.

Esta tendencia transforma el trabajo humano. En el pasado, un investigador pasaba todo el día preparando una sola muestra. Ahora, los robots realizan ese trabajo día y noche. Los humanos reflexionan sobre qué objetivos establecer y juzgan qué resultados tienen significado. El trabajo se desplaza del uso de las manos al uso del pensamiento. En cierto sentido, el papel humano no desaparece, sino que asciende a un nivel superior. En el próximo capítulo se abordará este laboratorio autónomo en detalle.

Existe una razón por la cual los catalizadores de alta entropía que combinan múltiples elementos atraen tanta atención. Al mezclar varios elementos, se generan diversos sitios activos en la superficie. Esto aumenta la probabilidad de que aparezca un sitio óptimo para la reacción. Sin embargo, el número de combinaciones es tan inmenso que resulta difícil para los humanos probarlas todas. El modelo generativo explora este vasto espacio en su lugar. Reduce la cantidad de experimentos acotando las composiciones prometedoras. El robot sintetiza y verifica rápidamente esos candidatos.

Examinemos la importancia que tienen los catalizadores en los cohetes y la exploración espacial. En misiones espaciales prolongadas, es necesario producir agua y aire de forma autónoma. Los catalizadores se utilizan para dividir el agua y obtener oxígeno. También se emplean para generar combustible. Un buen catalizador produce más oxígeno y combustible con la misma electricidad. Como resultado, la carga útil que debe transportar la nave espacial se aligera. El diseño inverso mediante IA acelera la llegada de tales catalizadores.

El hidrógeno también se utiliza como combustible para cohetes. El hidrógeno líquido genera un potente empuje. A medida que avanza la tecnología para obtener hidrógeno dividiendo el agua, la producción de combustible se vuelve más sencilla. En este proceso, el catalizador también tiene la clave de la reacción. Si se fabrica un buen catalizador con elementos económicos, los costos disminuyen. El objetivo es reducir el uso de platino, que es escaso y costoso. La IA generativa demuestra su potencia en la búsqueda de estos catalizadores no basados en metales preciosos.

6.3 Diseño inverso de aleaciones y superaleaciones aeroespaciales

El último caso es el diseño de aleaciones para motores de cohetes. Veremos por qué son necesarias nuevas aleaciones y cómo los modelos generativos descubren composiciones de aleaciones. Nos basaremos en investigaciones recientes y validadas.

Los motores de cohetes deben soportar un calor extremo. El interior de la cámara de combustión alcanza temperaturas de miles de grados. La tobera sufre la erosión de los gases a alta velocidad. Durante décadas, las superaleaciones (Superalloy) a base de níquel han ocupado este lugar. Una superaleación es una aleación que no pierde su resistencia incluso a altas temperaturas. Sin embargo, las superaleaciones de níquel también tienen límites de temperatura. En entornos aún más calientes, se funden o se ablandan.

Como alternativa, han surgido las aleaciones de alta entropía (High-Entropy Alloy). En las aleaciones convencionales se añade una pequeña cantidad a un elemento principal. En cambio, una aleación de alta entropía mezcla múltiples elementos en proporciones similares. Al mezclarse de este modo, la estructura se vuelve sorprendentemente estable. Aquellas fabricadas con elementos de alto punto de fusión se denominan aleaciones de alta entropía refractarias (Refractory HEA). Estas aleaciones se consideran candidatas para reemplazar a las superaleaciones de níquel en entornos que superan los 1100 grados.

La razón por la cual las aleaciones de alta entropía son estables radica en el desorden. Cuando varios elementos se mezclan de manera homogénea, aumentan las posibles configuraciones de disposición. Este gran número de configuraciones sostiene y estabiliza la estructura. Esta propiedad se denomina alto desorden configuracional. En cierto modo, el desorden preserva el orden. Gracias a esta propiedad, resisten bien incluso en entornos extremos. Las toberas y las cámaras de combustión de los cohetes son lugares representativos que exigen dicha resistencia.

El problema radica en que el número de combinaciones vuelve a ser astronómicamente alto. Incluso si se mezclan solo cinco elementos, existen innumerables combinaciones de proporciones. Es difícil sintetizarlas y probarlas una por una. Aquí es donde intervienen los modelos generativos. Si se definen las propiedades objetivo, el modelo busca la composición correspondiente de forma inversa. Se introducen los valores de resistencia y tenacidad como condiciones, y el modelo sugiere las proporciones de elementos capaces de alcanzar dichos valores.

Diversos estudios publicados en 2026 muestran esta dirección. Un estudio resumió varios métodos generativos para crear composiciones de aleaciones, examinando conjuntamente autoencoders variacionales, redes generativas antagónicas y modelos de difusión. Otro estudio recopiló literatura y datos para acotar candidatos de superaleaciones de alta entropía para altas temperaturas. Al seleccionar composiciones que cumplen con las propiedades objetivo, redujo los candidatos a evaluar. Ambos estudios se encuentran en la fase de acotar candidatos, no en la etapa de fabricación de componentes reales.

Las propiedades de una aleación no están determinadas únicamente por su composición. Incluso con la misma composición, varían según el método de fabricación. La forma en que se aplica el calor y se enfría altera la microestructura. La microestructura es el tamaño y la disposición de los granos cristalinos. Por ello, el diseño de aleaciones debe considerar tanto la composición como el proceso. Los modelos generativos también avanzan en la dirección de abordar ambos aspectos conjuntamente. En los siguientes capítulos se profundizará en esta cuestión de los procesos.

Sin embargo, hay aspectos que deben considerarse con cautela. Las aleaciones de alta entropía refractarias todavía enfrentan obstáculos que deben superarse. Conservan la tendencia a la fragilidad a

temperatura ambiente y su superficie puede deteriorarse al reaccionar con el oxígeno a altas temperaturas. Por ello, existe una distancia entre las promesas del laboratorio y los componentes reales. Aunque los modelos generativos ayuden a reducir los candidatos, estas dos barreras deben superarse mediante la experimentación. Controlar la fragilidad a temperatura ambiente y la oxidación a alta temperatura es el siguiente desafío.

6.4 Comparación de modelos generativos y la importancia de la validación

¿Cómo se comparan los diversos modelos generativos? Veamos con qué métricas se evalúan. Examinemos por qué la validación es más importante que los resultados en sí y aclaremos por qué debemos ser cautelosos con las cifras no verificadas.

Al comparar modelos generativos, se evalúan varios criterios en conjunto. Uno de ellos es la estabilidad, que verifica si el cristal generado se mantiene por sí mismo. También se evalúa la novedad, analizando si se trata de un material que no existía en las bases de datos. Se mide asimismo la validez de la distancia atómica, comprobando que los átomos no estén excesivamente juntos. Se examina la preservación de la simetría para ver si se respeta la simetría cristalina. Por último, se mide la velocidad, evaluando el tiempo que toma generar un candidato.

Un modelo no es bueno solo por destacar en una única métrica. Si la velocidad es alta pero la estabilidad es baja, su utilidad es reducida. Aunque conserve bien la simetría, si carece de novedad, su significado se diluye. Por ello, se busca un equilibrio entre múltiples indicadores. La métrica prioritaria cambia según el problema. Si se trata de materiales para baterías, la estabilidad y la conductividad tienen prioridad; si urge una exploración rápida, la velocidad se antepone.

Para hacer una comparación justa, se debe utilizar el mismo conjunto de prueba. En el campo de la generación de cristales, existen conjuntos de prueba ampliamente utilizados. MP-20 es un ejemplo representativo, consistente en una colección de unos veinte mil cristales extraídos del Materials Project. También existen conjuntos que recopilan únicamente cristales de perovskita. Todos los modelos se evalúan con el mismo conjunto para poder comparar cuál es superior. Si los conjuntos de prueba son diferentes, las puntuaciones no se pueden comparar directamente.

Aquí hay una precaución que debe tenerse en cuenta. Con frecuencia, las cifras que comparan el rendimiento de varios modelos uno al lado del otro no están confirmadas en los textos originales. Por esta razón, este libro no utiliza números de tablas no verificadas. Las cifras no verificadas pueden inducir a error.

En su lugar, nos orientamos con un resultado verificado. Se trata de un estudio sobre modelos generativos publicado por Microsoft en Nature en 2025. El nombre de este modelo es MatterGen. Este modelo genera materiales inorgánicos en toda la tabla periódica de elementos y también puede ajustarse finamente según las propiedades deseadas. El artículo informó que las estructuras generadas por este modelo eran estables y novedosas el doble de veces en comparación con los modelos anteriores, y estaban diez veces más cerca del mínimo de energía.

Este modelo puede reentrenarse para adaptarse a diversas condiciones. Si se entrena según valores de elasticidad, genera materiales duros; si se entrena según valores magnéticos, genera materiales magnéticos. También puede aprender a utilizar únicamente los elementos deseados. Es un método que adapta un único modelo base a múltiples objetivos. A este método se le llama ajuste fino (Fine-tuning). En lugar de reconstruir la base, simplemente se intercambian los objetivos.

Este estudio avanzó hasta la validación experimental. Estableciendo un módulo de compresibilidad objetivo de 200 gigapascales, se generó una nueva estructura material. Este material se sintetizó

realmente. Sin embargo, lo sintetizado fue una solución sólida con átomos mezclados de forma desordenada. El equipo de investigación midió el módulo de Young mediante nanoindentación y estimó el módulo de compresibilidad utilizando valores obtenidos mediante cálculo. El valor estimado obtenido de esta manera fue de un máximo de 169 gigapascales entre cuatro intentos, con un promedio de 158±11 gigapascales. Es necesario distinguir al leer entre valores medidos y valores convertidos, así como entre la estructura calculada y la estructura sintetizada. Aun así, se trata de una demostración poco común que abarca desde la generación hasta la síntesis.

Hay que tener aún más cuidado al tratar con preprints. Varios modelos presentados en este capítulo se encuentran en etapa de preprint. Un preprint aún no ha pasado por la revisión por pares. Los resultados pueden cambiar más adelante. Por ello, este libro presenta los preprints aclarando que son investigaciones iniciales. Incluso ante cifras sobresalientes, no se dan por definitivas antes de su validación. Esta postura protege al lector de falsas certezas.

Examinemos por qué la validación es más importante que los resultados. Los modelos generativos crean infinitos candidatos. Sin embargo, generarlos no significa que todos sean reales. Hay que filtrarlos mediante cálculos. Hay que confirmarlos mediante experimentos. El diseño inverso de materiales cobra sentido cuando la generación, el cálculo y el experimento funcionan de manera conjunta. Este principio debe ser aún más estricto en campos donde la seguridad está en juego, como los materiales aeroespaciales. Solo después de pasar por la validación computacional y la síntesis experimental se puede confiar en un candidato.

Resumen

Este capítulo trató sobre invertir la dirección en la búsqueda de materiales. El método tradicional primero elegía el material y luego medía sus propiedades. El nuevo método primero define las propiedades y luego diseña el material. Este nuevo enfoque se denomina diseño inverso de materiales. Con solo cambiar de dirección, el poder de búsqueda se transforma por completo. Se abre el camino para concebir materiales que antes no existían.

El protagonista del diseño inverso es la inteligencia artificial generativa. Los modelos de difusión esculpen cristales eliminando el ruido de una nube difusa de átomos. El principio es similar al de la IA de generación de imágenes. Sin embargo, generar cristales es más difícil. Esto se debe a que es necesario ajustar simultáneamente la red cristalina, las posiciones atómicas y los tipos de átomos.

Los cristales se rigen por una regla estricta llamada simetría. Un cristal tridimensional pertenece a uno de los 230 grupos espaciales. El modelo inicial CDVAE solía colapsar en su estructura porque no podía fijar la simetría de manera directa. Los modelos de difusión condicional conscientes de la simetría redujeron este problema creando una pequeña semilla y desplegándola mediante simetría. Es importante mantener la postura de distinguir y analizar por separado estudios con nombres similares sin confundirlos.

También continuaron los esfuerzos para aumentar la velocidad. El emparejamiento de flujos (flow matching) aprende atajos en lugar de eliminar el ruido poco a poco. FlowMM y CrystalFlow son ejemplos representativos. Crean cristales rápidamente siguiendo geodésicas en espacios curvos. Esta velocidad resulta de gran utilidad en la exploración de materiales, donde es necesario extraer una gran cantidad de candidatos.

También ha avanzado la tecnología para modificar solo partes específicas. El inpainting cubre vacantes y defectos. La difusión de átomos espejismo gestiona libremente el número de átomos. Establece posiciones virtuales que pueden o no existir, determinando la cantidad por sí misma. Esto permite representar mejor las estructuras de defectos en materiales reales.

El aprendizaje por refuerzo guía al modelo generativo hacia candidatos prometedores. Incorpora la estabilidad, la novedad y la sintetizabilidad en la recompensa. La regularización KL y la repetición de experiencias mantienen la exploración estable. El rendimiento se mide a través de la tasa SUN y la energía sobre el casco.

Los casos prácticos demuestran que esta tecnología no es solo una teoría sobre el papel. En el diseño de electrolitos sólidos de haluro, el análisis de datos se vinculó directamente con el diseño. En el diseño de catalizadores de alta entropía, la combinación de experimentos robóticos y modelos generativos redujo drásticamente el tiempo. El modelo generativo de Microsoft sintetizó realmente el material diseñado y verificó sus propiedades.

Hay una sola idea central que recorre este capítulo. La inteligencia artificial es un guía que recorre el inmenso espacio de materiales en nuestro lugar. Explora en pocos días caminos que a un ser humano le tomaría toda una vida recorrer. Selecciona y presenta candidatos prometedores. Sin embargo, son el cálculo y el experimento los que deciden si confiar o no en esos candidatos. El guía estrecha el camino y el ser humano abre la última puerta.

Por último, reiteramos el principio de validación. Los modelos generativos producen candidatos de forma infinita. Sin embargo, generarlos no significa que todos sean reales. No se utilizan cifras no confirmadas. Hay que filtrar mediante cálculos y verificar mediante experimentos. El diseño inverso de materiales cobra sentido cuando la generación, el cálculo y el experimento funcionan al unísono. En campos donde la seguridad está en juego, como los cohetes y la exploración espacial, este principio debe ser todavía más riguroso. La estructura trazada por la IA no es la respuesta definitiva, sino un candidato experimental bien seleccionado.

Fuentes y referencias relacionadas

Xie, T., Fu, X., Ganea, O. E., Barzilay, R., & Jaakkola, T. (2022). Crystal diffusion variational autoencoder for periodic material generation. International Conference on Learning Representations. https://arxiv.org/abs/2110.06197

Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., & Liu, Y. (2023). Crystal structure prediction by joint equivariant diffusion. Advances in Neural Information Processing Systems. https://openreview.net/forum?id=DNdN26m2Jk

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. Nature, 624, 80-85. https://doi.org/10.1038/s41586-023-06735-9

Zeni, C., Pinsler, R., Zugner, D., Fowler, A., Horton, M., Fu, X., et al. (2025). A generative model for inorganic materials design. Nature, 639, 624-632. https://doi.org/10.1038/s41586-025-08628-5

Levy, D., Panigrahi, S. S., Kaba, S. O., Zhu, Q., Lee, K. L. K., Galkin, M., Miret, S., & Ravanbakhsh, S. (2025). SymmCD: Symmetry-preserving crystal generation with diffusion models. International Conference on Learning Representations. https://arxiv.org/abs/2502.03638

Ishii, T., Hisama, K., & Shinohara, K. (2026). Symmetry-aware conditional generation of crystal structures using diffusion models. arXiv. https://arxiv.org/abs/2601.08115

Okhotin, A., Nakhodnov, M., Kazeev, N., Lazarev, M., Ustyuzhanin, A. E., & Vetrov, D. (2025). MiAD: Mirage atom diffusion for de novo crystal generation. arXiv. https://arxiv.org/abs/2511.14426

Chen, J., Guo, J., Fako, E., & Schwaller, P. (2025). Accelerating inverse materials design using generative diffusion models with reinforcement learning. arXiv. https://arxiv.org/abs/2511.03112

Miller, B. K., Chen, R. T. Q., Sriram, A., & Wood, B. M. (2024). FlowMM: Generating materials with Riemannian flow matching. Proceedings of Machine Learning Research, 235. https://arxiv.org/abs/2406.04713

Luo, X., Wang, Z., Wang, Q., Shao, X., Lv, J., Wang, L., Wang, Y., & Ma, Y. (2025). CrystalFlow: A flow-based generative model for crystalline materials. Nature Communications, 16, 9267. https://doi.org/10.1038/s41467-025-64364-4

Lo, A., Mucko, L., Cheng, A. H., Cai, A., Price, A. J. A., & Matusik, W. (2026). Fast organic crystal structure prediction with unit cell flow matching. arXiv. https://arxiv.org/abs/2606.03199

Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nicklas, M., & Le, M. (2022). Flow matching for generative modeling. International Conference on Learning Representations. https://openreview.net/forum?id=PqvMRDCJT9t

Hoogeboom, E., Satorras, V. G., Vignac, C., & Welling, M. (2022). Equivariant diffusion for molecule generation in 3D. International Conference on Machine Learning, 8867-8887. https://arxiv.org/abs/2203.17003

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv. https://arxiv.org/abs/1707.06347

Metni, H., Ruple, L., Walters, L. N., Torresi, L., Teufel, J., Schopmans, H., Ceder, G., & Friederich, P. (2026). Generative models for crystalline materials. Advanced Materials, 38(18), e23620. https://doi.org/10.1002/adma.202523620

Mal, S., Ahmed, N., Jami, J., Mishra, S., & Sen, P. (2025). DiffCrysGen: A generative diffusion model for accelerated design of inorganic crystalline materials. arXiv:2510.12329. https://arxiv.org/abs/2510.12329

Okuda, I., Takahara, I., & Mizoguchi, T. (2026). Inverse materials design via joint generation of crystal structures and local electronic descriptors. arXiv:2605.01286. https://arxiv.org/abs/2605.01286

Yang, K., & Schwalbe-Koda, D. (2025). A generative diffusion model for amorphous materials. npj Computational Materials, 12(1), 29. https://doi.org/10.1038/s41524-025-01901-1

Han, X.-Q., Guo, P.-J., Gao, Z.-F., Sun, H., & Lu, Z.-Y. (2025). InvDesFlow-AL: Active learning-based workflow for inverse design of functional materials. npj Computational Materials, 11(1), 364. https://doi.org/10.1038/s41524-025-01830-z

Cheng, M., Luo, W., Tang, H., Yu, B., Cheng, Y., Xie, W., Li, J., Kulik, H. J., & Li, M. (2026). Enhancing materials discovery with valence-constrained design in generative modeling. Nature Computational Science. https://doi.org/10.1038/s43588-026-01037-2

Choong, L. Y. A., Chen, Z., & Ng, M.-F. (2025). Discovery of effective halide solid electrolytes for solid-state rechargeable batteries via machine learning and DFT calculations. ACS Applied Energy Materials, 9(1), 507-520. https://doi.org/10.1021/acsaem.5c03277

Li, D., Lin, Z., Luo, F., Xiao, C., Dai, Z., Ye, Z., Wang, Q., Guo, Y.-G., & Yang, C. (2026). Data-driven chemical feature extraction and material design for halide solid-state electrolytes. eScience, 100590. https://doi.org/10.1016/j.esci.2026.100590

Zhou, D., Yang, R., Jia, Z., Cai, Y., Zhao, L., Guo, L., Ye, G., Wang, S., Chen, L., Liu, D., Smith, P. E. S., Huang, Y., Zhu, Q., & Jiang, J. (2026). A practical inverse design approach for high-entropy catalysts using generative AI. Nature Synthesis, 5(5), 730-739. https://doi.org/10.1038/s44160-025-00983-5

Li, C., Xian, Y., Zhou, Y., Ding, X., Sun, J., & Xue, D. (2026). Generative design for alloys: Harnessing generative models for faster discovery. Advanced Materials, 38(29), e20478. https://doi.org/10.1002/adma.202520478

Rousseau, F., Belmonte, T., Sur, F., & Nominé, A. (2026). Inverse design of high-entropy superalloys using machine learning and generative artificial intelligence. Materials & Design, 266, 116097. https://doi.org/10.1016/j.matdes.2026.116097

Du, H., Hui, J., Zhang, L., & Wang, H. (2025). Universal machine learning interatomic potentials are ready for solid ion conductors. arXiv:2502.09970. https://arxiv.org/abs/2502.09970

Capítulo 3. Laboratorios autónomos y sistemas multiagente científicos

Introducción

Pensemos en la escena de hornear pan en casa. Se decide la cantidad de harina y agua. Se mezcla la masa y se ajusta la temperatura del horno. Se corta el pan horneado y se prueba. Si queda demasiado duro, se añade más agua y se baja la temperatura. La próxima vez sale un pan un poco mejor. De este modo, el proceso de elegir los ingredientes, elaborar, probar y corregir se repite una y otra vez.

La investigación para encontrar nuevos materiales es muy similar. El investigador determina la proporción de los componentes. Mezcla los polvos y los hornea a altas temperaturas. Mide la estructura y las propiedades del material resultante. Si el resultado es insatisfactorio, modifica las condiciones. Este proceso debe repetirse cientos de veces para obtener un único buen material.

El problema es el tiempo. Si una persona realiza esta tarea manualmente, solo puede hacerlo unas pocas veces al día. Por la noche, los experimentos se detienen. Explorar un rango amplio modificando los componentes poco a poco lleva varios años. Las condiciones que deben soportar los materiales utilizados en cohetes son sumamente exigentes. Por ello, la cantidad de sustancias candidatas es tan grande que resulta difícil probarlas todas únicamente con manos humanas.

Un laboratorio autónomo (Self-Driving Laboratory) es un centro de investigación que delega esta repetición a las máquinas. La inteligencia artificial (AI) selecciona las siguientes condiciones que se van a probar. Un robot pesa, mezcla y hornea los polvos. Los equipos de análisis miden los resultados. Dichos resultados regresan a la inteligencia artificial. Los experimentos continúan por sí solos sin necesidad de que una persona dé instrucciones en cada ocasión. Al igual que un robot aspirador que limpia por su cuenta, el laboratorio funciona de forma autónoma durante toda la noche.

A esta estructura en la que los resultados alimentan directamente la siguiente elección se la denomina bucle cerrado (Closed-loop). Un bucle cerrado se refiere a un circuito cerrado como un anillo. El ciclo de experimentación, medición, aprendizaje y siguiente experimento gira sin interrupción. Este capítulo aborda el funcionamiento de estos laboratorios autónomos. Examina cómo la inteligencia artificial elige el siguiente experimento. Observa cómo las máquinas leen los textos de los artículos académicos. Analiza cómo los robots y la inteligencia artificial interactúan de manera segura. Por último, revisa casos reales de laboratorios autónomos en todo el mundo.

También conviene señalar por qué los laboratorios autónomos están cobrando tanto auge en la actualidad. Recientemente, la inteligencia artificial predice las propiedades de los materiales a gran velocidad. Mediante cálculos, genera miles de buenos candidatos al día. Sin embargo, la velocidad para fabricarlos y verificarlos en la práctica sigue siendo lenta. La predicción se ha acelerado, pero la verificación no logra seguirle el ritmo. El laboratorio autónomo es lo que cierra esta brecha. Los robots fabrican y comprueban con rapidez los candidatos propuestos por el cálculo. Sirve, por así decirlo, como un puente que sincroniza la velocidad de predicción con la de verificación.

Muchos de los sistemas presentados en este capítulo aún corresponden a investigaciones iniciales. También se incluyen prepublicaciones recién presentadas en conferencias académicas. Una prepublicación es un estudio preliminar que todavía no ha pasado por la revisión por pares. Por esta razón, más que presumir de cifras de rendimiento, este capítulo examina tanto los principios como las limitaciones. El objetivo de un laboratorio autónomo no es simplemente realizar experimentos rápidos, sino llevar a cabo experimentos confiables.

3.1 ¿Con qué componentes funciona un laboratorio autónomo?

Lo que se aborda en esta sección

Esta sección examina de qué componentes está integrado el laboratorio autónomo en su conjunto. A grandes rasgos, consta de cuatro partes: la cabeza que decide el siguiente experimento, las manos que fabrican físicamente las muestras, los ojos que miden los resultados y la memoria que actualiza lo aprendido. Estas cuatro partes se conectan como un bucle y operan transfiriendo los resultados de un experimento al siguiente.

En primer lugar, se revisa el significado del concepto de bucle cerrado. A continuación, se examina una a una la función que cumple cada parte. Por último, se analiza por qué esta estructura es necesaria en la investigación de materiales para cohetes.

Un laboratorio autónomo funciona mediante el engranaje de cuatro partes. En la capa de planificación, la inteligencia artificial selecciona los componentes y las condiciones del proceso que se probarán a continuación. En este contexto, las condiciones del proceso se refieren a la temperatura de horneado, el tiempo de mantenimiento, la velocidad de calentamiento y la atmósfera gaseosa. En la capa de ejecución, los robots pesan, mezclan y prensan los polvos para formar pastillas, que luego introducen en un horno eléctrico para su cocción. En la capa de medición, herramientas como los equipos de difracción de rayos X miden la estructura del material resultante. En la capa de aprendizaje, los resultados de la medición se introducen en el modelo de inteligencia artificial para actualizar el conocimiento.

El flujo de información es lo que conecta estas cuatro partes. La capa de planificación envía una especificación de trabajo a la capa de ejecución. Dicha especificación es una hoja de instrucciones que indica qué hacer, en qué cantidad y de qué manera. La capa de ejecución envía las muestras preparadas a la capa de medición. La capa de medición transmite valores como la pureza de fase y la constante de red a la capa de aprendizaje. La pureza de fase indica con qué grado de pureza se ha obtenido únicamente el cristal deseado. La capa de aprendizaje utiliza esos valores para perfeccionar el siguiente plan. Cuando la información completa un ciclo de esta manera, concluye un experimento.

El poder del bucle cerrado radica en la rapidez de su repetición. Incluso si un experimento fracasa, los datos quedan registrados. Las condiciones fallidas también enseñan qué es lo que no funciona. La inteligencia artificial aprende tanto de los aciertos como de los fracasos. Gracias a ello, las condiciones siguientes mejoran poco a poco. Los humanos solo supervisan la dirección general una vez cada varios días, mientras que las máquinas se encargan de las iteraciones detalladas.

Veamos por qué esta estructura es necesaria para los materiales de cohetes. Las toberas y las cámaras de combustión de los cohetes soportan condiciones extremas. Las temperaturas se elevan a miles de grados. La presión y las vibraciones también son intensas. Los materiales destinados a estos entornos deben cumplir requisitos sumamente exigentes. Un cambio mínimo en la composición modifica drásticamente las propiedades. Por ello, existen decenas de miles de candidatos que deben evaluarse. Es imposible explorar un espacio tan amplio únicamente con manos humanas. El laboratorio autónomo permite reducir rápidamente este vasto espacio.

En 2026, continuaron las investigaciones destinadas a elevar los laboratorios autónomos a un nuevo nivel. La Real Sociedad de Química británica definió el concepto de Laboratorio Autónomo 2.0. Los laboratorios autónomos de primera generación solo resolvían un problema específico prefijado. La siguiente generación busca manejar con flexibilidad múltiples tipos de equipos y problemas. También se publicó en 2025 un artículo de revisión exhaustivo sobre laboratorios autónomos, el cual analiza tanto el grado de madurez tecnológica como los obstáculos pendientes. Los medios de la Sociedad

Química Estadounidense informan que los laboratorios autónomos están transformando la forma de trabajar de los químicos.

Esta estructura también presenta limitaciones evidentes. Equipar estas instalaciones requiere cuantiosas sumas de dinero. Integrar y conectar robots con equipos de análisis resulta complejo. Incluso pesar los polvos con exactitud no es una tarea sencilla. Detectar automáticamente un experimento fallido también es difícil. Por eso, los laboratorios autónomos actuales solo se adaptan bien a familias de materiales específicas. Todavía no existe un laboratorio universal capaz de manejar cualquier material. Las secciones restantes de este capítulo analizan cómo se van superando estas barreras una a una.

3.1.1 Aprendizaje activo bayesiano para seleccionar el siguiente experimento

Esta subsección aborda la parte correspondiente a la cabeza del laboratorio autónomo. El núcleo radica en cómo la inteligencia artificial elige el siguiente experimento. El método empleado aquí es el aprendizaje activo bayesiano. Se explica en un lenguaje sencillo qué es este método y por qué reduce el número de experimentos.

Una hipótesis es una idea sobre por qué el siguiente experimento resultará favorable. Un buen laboratorio autónomo formula buenas hipótesis. Probar condiciones al azar sin criterio desperdicia tiempo. Es necesario seleccionar y evaluar únicamente aquellas áreas con potencial. El aprendizaje activo bayesiano (Bayesian Active Learning) es el método que ayuda en esta selección. El aprendizaje activo se refiere a la forma en que la máquina selecciona por sí misma los datos de los cuales aprenderá.

Este método construye en primer lugar un modelo sustituto (Surrogate Model). Un modelo sustituto es un modelo computacional pequeño que emula experimentos costosos. Se crea a partir de los resultados experimentales obtenidos hasta el momento. El modelo sustituto predice las propiedades en condiciones aún no probadas. En la predicción se obtienen dos valores simultáneamente: uno es el rendimiento esperado y el otro es el grado de incertidumbre de dicha predicción.

Analizar estos dos valores en conjunto es la clave. Un área con un alto rendimiento esperado vale la pena ser probada. Un área con alta incertidumbre también merece ser evaluada, ya que en regiones desconocidas puede esconderse un material extraordinario inesperado. Si solo se persigue el rendimiento, se termina dando vueltas sobre lo que ya se conoce. Si solo se persigue la incertidumbre, se vaga por zonas inútiles. Un buen método equilibra ambas inclinaciones. A esto se le llama el equilibrio entre exploración y explotación. En este sentido, la exploración consiste en examinar lo desconocido, y la explotación, en profundizar en lo conocido.

La función de adquisición (Acquisition Function) cuantifica este equilibrio en un valor numérico. La función de adquisición asigna una puntuación a cada condición candidata. Si se ordenan de mayor a menor puntuación, la primera de la lista se convierte en el siguiente experimento. Una función de adquisición representativa es la mejora esperada (Expected Improvement). La mejora esperada evalúa cuánto mejorará el siguiente experimento en comparación con el mejor resultado obtenido hasta ahora. Otro método es el límite superior de confianza (Upper Confidence Bound). Este método permite considerar simultáneamente el alto rendimiento y las regiones desconocidas. El laboratorio autónomo determina la siguiente composición, temperatura y tiempo basándose en estas puntuaciones.

Veamos por qué este método es tan importante en los materiales para cohetes. Las cerámicas de ultraalta temperatura y las aleaciones especiales implican un costo elevado en cada experimento. Utilizan materias primas costosas y requieren largos tiempos de horneado. Por lo tanto, reducir el número de experimentos equivale a ahorrar tiempo y dinero. El aprendizaje activo bayesiano selecciona y prueba únicamente las condiciones con potencial de éxito. El número de experimentos se

reduce drásticamente en comparación con probarlo todo a ciegas. Algunos laboratorios autónomos informan haber reducido las pruebas necesarias en decenas de veces.

Su eficacia se demuestra en casos reales. En 2026, un equipo de investigación construyó un laboratorio para el crecimiento autónomo de películas delgadas. Una película delgada se fabrica depositando átomos capa por capa. Este laboratorio leía los patrones de difracción de electrones en tiempo real mediante una cámara. Los patrones de difracción son señales que indican si los átomos se están depositando adecuadamente. La inteligencia artificial seleccionaba las siguientes condiciones observando estos patrones. Como resultado, el número de experimentos requeridos se redujo notablemente. Se informó que se redujo más de treinta veces en comparación con explorar cada condición una por una. Así se encontraron condiciones óptimas con mucha mayor rapidez. Aunque este estudio se encuentra en una etapa inicial, muestra con claridad el camino a seguir.

En 2026, proliferaron las investigaciones para adaptar este método a los laboratorios autónomos. Una prepublicación sintetizó la optimización bayesiana para laboratorios autónomos de materiales, explicando de manera continua desde los algoritmos hasta los flujos experimentales reales. Otra prepublicación propuso un método para reducir de forma autónoma el espacio de búsqueda. Mediante el aprendizaje activo, redujo el espacio de candidatos aproximadamente a la mitad, preservando al mismo tiempo la gran mayoría de las soluciones óptimas. De esta manera, es posible encontrar buenos materiales con mayor rapidez. En 2026 también se publicó una guía práctica de optimización bayesiana para las ciencias experimentales.

Este método también tiene sus limitaciones. Al principio, el modelo sustituto cuenta con pocos datos, por lo que sus predicciones son aproximadas. En los primeros experimentos, puede orientarse en una dirección errónea. Si el espacio composicional es sumamente amplio, la carga de cálculo aumenta considerablemente. Si las mediciones presentan mucho ruido, las predicciones pierden estabilidad. Por esta razón, los humanos no confían a ciegas en las predicciones del modelo sustituto. Las toman como referencia, pero siempre las verifican mediante mediciones reales. Incluso en un laboratorio autónomo, la supervisión humana sigue estando presente.

3.2 Extracción de métodos de síntesis de la literatura científica y su conversión en planes seguros

Lo que se aborda en esta sección

Esta sección aborda dos aspectos: primero, hacer que las máquinas lean los métodos de síntesis descritos en los artículos científicos; segundo, transformar esos métodos en planes ejecutables de forma segura por robots.

Los humanos leen artículos e imaginan mentalmente la secuencia experimental. Las máquinas no pueden hacerlo. Por ello, es necesario convertir las oraciones en un formato que las máquinas puedan procesar. Esta sección examina ese proceso de traducción y la planificación segura.

En los artículos científicos, los métodos de síntesis están redactados en oraciones. En ellos se describe qué polvos mezclar y en qué proporciones, a qué temperatura y durante cuántas horas hornear, y en qué atmósfera de gas llevar a cabo el tratamiento. En la literatura acumulada durante décadas existen millones de métodos de este tipo. Este conocimiento constituye un recurso invaluable para la investigación de nuevos materiales; sin embargo, en su mayor parte se encuentra disperso en lenguaje natural humano.

El problema radica en que estas expresiones varían según cada autor. Algunos artículos se limitan a escribir que se horneó durante toda la noche. Otros solo indican que se trató a alta temperatura. No

especifican con exactitud cuántas horas significa toda la noche ni a cuántos grados equivale alta temperatura. Los humanos llenan estas lagunas gracias a su experiencia, pero las máquinas no pueden hacerlo. Por lo tanto, resulta indispensable convertir estas oraciones dispersas en números precisos.

En la investigación de materiales para cohetes, esta labor es de gran valor. Los métodos de síntesis de cerámicas de ultraalta temperatura y aleaciones especiales están dispersos en artículos antiguos. Para que un ser humano lea y organice todo ello, se requiere mucho tiempo. Si una máquina extrae y organiza automáticamente el texto, se ahorra una gran cantidad de tiempo. El laboratorio autónomo transforma de inmediato los métodos recopilados en instrucciones robóticas, acortando la distancia entre el artículo científico y el experimento real.

Cuando este conocimiento permanece disperso, surgen investigaciones redundantes. Se vuelven a probar condiciones en las que otros investigadores ya fracasaron, o se vuelve a buscar desde cero por desconocer un método que ya tuvo éxito. Reunir los métodos de la literatura en un solo lugar reduce este desperdicio. Dado que para un ser humano resulta muy difícil leer millones de artículos, la labor de las máquinas al extraer y estructurar el texto es sumamente valiosa. Un método bien organizado se convierte en un excelente punto de partida para el laboratorio autónomo.

Aquí intervienen dos etapas. La primera es extraer el método a partir del texto. De esta tarea se encarga un modelo de lenguaje grande (Large Language Model), que es una inteligencia artificial capaz de aprender y procesar los textos humanos. La segunda consiste en transformar el método extraído en un plan seguro, tarea que recae en un lenguaje de planificación. En las dos subsecciones siguientes se examina cada etapa en detalle.

3.2.1 Modelos de lenguaje que formulan planes de síntesis a partir de artículos científicos (MSP-LLM)

Esta subsección aborda cómo extraer métodos de síntesis a partir de artículos científicos. Como caso representativo, se examina la investigación sobre MSP-LLM. Se analiza qué problema intenta resolver este estudio y a través de qué etapas opera.

MSP-LLM es una investigación sobre modelos de lenguaje grandes para la planificación de la síntesis de materiales. Su nombre proviene del título 'Marco unificado de modelos de lenguaje grandes para la planificación completa de síntesis de materiales'. Es un estudio preliminar publicado en 2026. El equipo de investigación está formado por Noh Hee-woong, Na Kyung-soo, Lee Nam-kyeong y Park Chan-young. Este estudio es un artículo preliminar presentado en una conferencia académica. Por lo tanto, no se considera un estándar de campo, sino un intento pionero.

El problema que aborda este estudio es la planificación de la síntesis. La planificación de la síntesis consiste en establecer la secuencia completa para fabricar el material objetivo. Esta tarea se divide en dos subproblemas. Uno es la predicción de precursores. Los precursores son las materias primas que se introducen al inicio para crear el material objetivo. El otro es la predicción de las operaciones de síntesis. Las operaciones de síntesis son la secuencia de acciones tales como mezclar, hornear y enfriar. Solo al resolver ambos problemas juntos se obtiene un plan completo.

MSP-LLM conecta estos dos problemas en un único flujo continuo. Primero, determina a qué categoría pertenece el material objetivo. Esta categoría sirve como criterio de decisión intermedio. A continuación, selecciona los precursores adecuados para esa categoría. Por último, diseña la secuencia de operaciones correspondiente a dichos precursores. De este modo, se crea una cadena que va de la categoría a la materia prima, y de la materia prima a la secuencia. Al conectarlo en una cadena, se obtiene un plan químicamente coherente.

Veamos este flujo con un ejemplo más sencillo. Decidir qué plato cocinar equivale a determinar la categoría del material objetivo; es como decidir que se preparará estofado de kimchi (kimchi jjigae). Luego, decidir qué ingredientes comprar equivale a la predicción de precursores; es como elegir kimchi, carne de cerdo y tofu. Por último, decidir en qué orden cocinar equivale a la predicción de operaciones; es como planificar la secuencia de saltear, verter agua y hervir. MSP-LLM conecta estos tres pasos para completar el plan.

Veamos por qué esta planificación automática es valiosa en los materiales para cohetes. Para fabricar una nueva aleación o cerámica, la selección de materias primas ya resulta difícil. Incluso para un mismo objetivo, si se cambian las materias primas, el resultado varía. Cambiar el orden de horneado también altera el resultado. Si un ser humano tuviera que sopesar todas estas combinaciones, llevaría mucho tiempo. Un planificador automático genera con rapidez un buen punto de partida basándose en el conocimiento de los artículos científicos. El laboratorio autónomo puede comenzar los experimentos a partir de este punto de partida.

Este método también presenta límites claros. Los modelos de lenguaje grandes a veces inventan hechos inexistentes, lo que se conoce como alucinación. Si se experimenta directamente con materias primas o condiciones inventadas, se desperdician recursos. Además, podrían generarse combinaciones peligrosas. Por ello, el plan extraído debe verificarse siempre mediante humanos u otras herramientas de validación. Primero se examina si cumple con las reglas químicas y si es seguro. La planificación automática es una herramienta para asistir a los humanos, no para reemplazarlos.

3.2.2 Prevención de accidentes robóticos mediante lenguajes de planificación (PDDL)

Esta subsección aborda la conversión del método de síntesis extraído en un plan robótico seguro. La herramienta utilizada aquí es el lenguaje de planificación PDDL. Se explica en términos sencillos cómo este lenguaje previene accidentes robóticos.

Un plan experimental no es simplemente una tabla de secuencias. Cada acción tiene condiciones previas que deben cumplirse. El horno eléctrico debe estar vacío para introducir una muestra. El horno eléctrico debe haberse enfriado para retirar la muestra. El pesado del polvo debe haber finalizado para proceder a la mezcla. Si se infringen estas condiciones secuenciales, ocurren accidentes. Resulta peligroso si un robot abre indiscriminadamente la puerta de un horno eléctrico caliente.

PDDL (Planning Domain Definition Language) es un lenguaje para redactar estos planes. Su nombre significa 'lenguaje de definición de dominio de planificación'. Este lenguaje especifica dos elementos para cada acción. Uno son las precondiciones, donde se define qué debe ser verdadero antes de realizar dicha acción. El otro son los efectos de ejecución, donde se describe qué cambia al ejecutar la acción. Por ejemplo, la precondición para la acción de hornear es que el horno eléctrico esté en estado de espera y la temperatura sea baja. Y el efecto de ejecución al concluir esta acción es que la muestra queda tratada térmicamente.

Al definir las condiciones de este modo, la máquina encuentra de forma autónoma una secuencia segura. Las acciones cuyas condiciones no se cumplen no se ejecutan. El comando de aproximación hacia un horno eléctrico caliente queda bloqueado. Una muestra cuyo pesado no haya concluido no avanza a la etapa de mezcla. Estas reglas actúan como una valla de protección que previene de antemano los accidentes del robot. Es el equivalente a trasladar las normas de seguridad del laboratorio seguidas por humanos al lenguaje de las máquinas.

Sin embargo, esta valla no garantiza la seguridad por sí sola. Las precondiciones solo funcionan si un ser humano las redacta una por una de manera precisa. Si se omiten condiciones o se redactan incorrectamente, se producen brechas en la valla. Por lo tanto, PDDL no es un dispositivo que garantice

una ejecución segura por sí mismo, sino una herramienta para que los humanos anticipen los riesgos y los conviertan en reglas. La rigurosidad de las reglas es tan detallada como el criterio del ser humano.

Este plan puede representarse mediante un diagrama. Se coloca cada acción como un nodo y se conectan las relaciones secuenciales con flechas. Este tipo de diagrama se denomina grafo acíclico dirigido. Significa que las flechas fluyen solo en una dirección y nunca regresan sobre sí mismas. Al observar este diagrama, se puede saber qué tareas se pueden realizar en paralelo. Por ejemplo, mientras se hornea una muestra, se pesa el polvo de otra muestra. Al realizar en paralelo tareas que no interfieren entre sí, el laboratorio completo opera mucho más rápido.

En los experimentos con materiales para cohetes, este plan de seguridad es invaluable. Los hornos eléctricos de ultraalta temperatura son extremadamente peligrosos si se manipulan de forma incorrecta. Las materias primas de alta reactividad no deben entrar en contacto con el aire. Múltiples equipos se mueven simultáneamente en un mismo laboratorio. En un laboratorio que opera durante toda la noche sin presencia humana, las reglas para prevenir accidentes son indispensables. Un lenguaje de planificación como PDDL estructura estas reglas de forma rigurosa. La velocidad del laboratorio autónomo solo tiene sentido sobre la base de estos mecanismos de seguridad.

Este enfoque también tiene limitaciones. Resulta difícil describir todas las condiciones del mundo mediante reglas. Pueden surgir situaciones reales que no se hayan previsto por escrito. Si hay demasiadas reglas, aumenta la carga computacional para encontrar un plan. Las averías imprevistas no pueden prevenirse únicamente con reglas. Por ello, un lenguaje de planificación no basta por sí solo. Debe complementarse con otros mecanismos que supervisen el estado de los equipos en tiempo real, los cuales se abordan en la siguiente sección.

3.3 Comunicación estándar que conecta la inteligencia artificial y los equipos de laboratorio (MCP)

Esta sección aborda cómo se comunican la inteligencia artificial y los equipos de laboratorio. Los modelos de lenguaje procesan oraciones. Las balanzas, los hornos eléctricos y los brazos robóticos requieren comandos precisos. Se necesita una capa intermedia que conecte a ambos. Esta sección examina dicha capa intermedia, analizando tanto los protocolos de comunicación estándar como los dispositivos de seguridad implementados sobre ellos.

Los modelos de lenguaje manejan bien el texto humano. Incluso entienden una frase como 'sube la temperatura lentamente'. Sin embargo, el horno eléctrico no puede comprender oraciones. El horno eléctrico debe recibir una señal precisa, como elevar la temperatura a determinada cantidad de grados por minuto. Lo mismo ocurre con las balanzas y los brazos robóticos. Por tanto, se necesita un intérprete que traduzca las oraciones al lenguaje de los equipos.

En el pasado, este intérprete se creaba por separado para cada equipo. Se programaban intérpretes individuales para la balanza, para el horno eléctrico y para el robot. Si se cambiaba un solo equipo, había que reprogramar el intérprete desde cero. Múltiples laboratorios repetían el mismo trabajo, lo que representaba un gran desperdicio. Al no existir un estándar, resultaba difícil compartir y reutilizar el código entre diferentes laboratorios.

Los protocolos de comunicación estándar buscan resolver este problema. Con una norma estándar, los equipos se conectan de la misma manera. Incluso un equipo nuevo se integra de inmediato siempre que cumpla con las reglas. También se vuelve sencillo compartir código entre laboratorios. Las investigaciones recientes sobre laboratorios autónomos buscan este estándar en el Protocolo de Contexto de Modelo.

Existe una especificación abierta llamada Protocolo de Contexto de Modelo (Model Context Protocol), abreviada como MCP. Un protocolo es un método de comunicación acordado mutuamente. MCP es un estándar que conecta la inteligencia artificial con las herramientas de software. Informa a la inteligencia artificial qué herramientas están disponibles mediante un formato estándar. La inteligencia artificial reconoce esas herramientas de forma autónoma y las invoca.

Aquí conviene aclarar un punto que suele prestarse a confusión. MCP es un protocolo de nivel superior que conecta software entre sí. No puede controlar directamente equipos físicos reales como balanzas u hornos eléctricos. La tarea de accionar motores y válvulas corresponde a los controladores específicos de cada equipo. Un controlador es un programa de bajo nivel que maneja directamente las señales de los dispositivos. MCP se monta sobre estos controladores y los vincula con la inteligencia artificial. Por ello, los equipos no se integran de inmediato únicamente con MCP. MCP no es una norma de control industrial en tiempo real ni un estándar de seguridad. Para conectar los equipos, se debe disponer de una capa de controladores independiente por debajo.

Un caso que introdujo este protocolo en los laboratorios autónomos es el controlador NIMO. Es una investigación preliminar publicada en 2026 por el equipo de Naruki Yoshikawa y Ryo Tamura. Este estudio expone diversas funciones del laboratorio autónomo a través de servidores MCP. Aquí, NIMO es el motor de decisión que determina el siguiente experimento. El controlador es la capa de orquestación que conecta ese motor con múltiples herramientas mediante MCP. En este caso, la inteligencia artificial gestionó un laboratorio autónomo para igualar colores. Modificó de forma autónoma los componentes y realizó experimentos hasta alcanzar el color objetivo. Lo que este estudio demostró en la práctica fue esta orquestación y el experimento de ajuste de color.

El propósito de NIMO que destaca este estudio es muy claro. En un laboratorio autónomo existen múltiples métodos de optimización y diversas herramientas. Conectar cada método con cada herramienta de forma individual representa un gran cuello de botella. El controlador NIMO conecta ambos mediante un método común. El ser humano puede diseñar el flujo experimental sin necesidad de escribir código, seleccionando y conectando herramientas directamente en la pantalla. No obstante, dado que este estudio se encuentra en una etapa inicial, aún es prematuro considerarlo un estándar de uso generalizado.

Veamos por qué esta capa de comunicación es importante en un laboratorio de materiales para cohetes. En un laboratorio de materiales coexisten equipos de diferentes fabricantes. Las balanzas, los hornos eléctricos y los equipos de difracción provienen de distintas marcas. Para que una sola inteligencia artificial los controle, se requiere un lenguaje común. Un protocolo de comunicación estándar integra estos diversos dispositivos en un flujo unificado. Permite incorporar equipos nuevos sin necesidad de grandes modificaciones. Así, resulta más sencillo ampliar o transformar el laboratorio.

La comunicación estándar ofrece otra ventaja: permite a las personas diseñar experimentos sin necesidad de saber programar. La inteligencia artificial detecta por sí misma qué herramientas están disponibles y puede desplegarlas en la pantalla. El usuario solo necesita seleccionar las herramientas y conectarlas en orden. Se diseña de forma visual un flujo continuo que incluye pesar con la balanza, mezclar, hornear y medir mediante difracción. Al reducirse la barrera para diseñar experimentos, más investigadores pueden utilizar laboratorios autónomos.

Esta capa también tiene limitaciones. Los protocolos de comunicación son aún recientes y se encuentran en proceso de consolidación. El grado de adopción de las normas varía según el fabricante de los equipos. Los equipos antiguos resultan difíciles de adaptar a estas normas. Si la comunicación se interrumpe o la señal se retrasa, el experimento puede fallar. Por ello, la propia capa de comunicación

requiere mecanismos de seguridad. Estos mecanismos verifican los comandos antes de ejecutarlos directamente, y eso es precisamente la máquina de estados.

3.3.1 La última línea de seguridad protegida por la máquina de estados

Esta subsección aborda la máquina de estados que salvaguarda la seguridad de los equipos. Se examina qué es una máquina de estados y cómo bloquea comandos peligrosos. También se analiza por qué este dispositivo constituye la última línea de seguridad en un laboratorio autónomo. La máquina de estados analizada aquí no es una función preinstalada en un producto específico. El controlador NIMO visto anteriormente es una investigación enfocada en la orquestación para conectar herramientas. La máquina de estados y los enclavamientos son diseños de seguridad comunes que deben implementarse de forma independiente sobre dicha orquestación. Este libro considera ese diseño de seguridad como una capa esencial que todo laboratorio autónomo debe incorporar.

Una máquina de estados es un método que clasifica el estado actual del equipo en casillas predeterminadas. Una máquina de estados finitos (Finite State Machine) define de antemano las casillas posibles. Por ejemplo, existen estados como espera, preparación, pesado, transporte, reacción, análisis, error y parada de emergencia. El equipo siempre se encuentra en uno de estos estados. Las rutas para pasar de un estado al siguiente también se determinan con antelación.

El principio fundamental es que solo se ejecutan las acciones permitidas en un estado determinado. En el estado de reacción, el horno eléctrico está caliente. En ese momento, se bloquea la orden para que el brazo robótico abra la puerta del horno. La puerta solo se abre tras recibir una señal de que la temperatura ha descendido por debajo de un valor seguro. Un sensor confirma esta señal. Antes de que llegue la señal de confirmación, el comando queda bloqueado en la etapa de la máquina de estados. Este tipo de bloqueo se denomina enclavamiento. Un enclavamiento es un bloqueo de seguridad que permanece cerrado si no se cumplen las condiciones.

Veamos por qué este método es seguro con un ejemplo sencillo. Un horno microondas detiene su funcionamiento cuando se abre la puerta. Solo funciona si la puerta está cerrada. Con la puerta abierta, no se encenderá por más que se presione el botón. Esta es una restricción de acción basada en el estado. La máquina de estados de un laboratorio autónomo opera bajo el mismo principio. En un estado peligroso, los comandos peligrosos ni siquiera llegan a ejecutarse.

Las instrucciones en lenguaje natural de la inteligencia artificial deben pasar obligatoriamente por estas reglas. Por mucha confianza con la que la inteligencia artificial emita un comando, las reglas siempre tienen prioridad. Solo los comandos que superan las reglas se envían a los controladores de los equipos. El controlador es el programa que acciona físicamente los motores y las válvulas. Los comandos interceptados por las reglas no se ejecutan y únicamente quedan registrados. Este registro sirve posteriormente como material para investigar las causas.

En los experimentos con materiales para cohetes, esta línea de seguridad es valiosa. Esto se debe a que se manejan hornos eléctricos de temperatura ultraalta y materias primas reactivas. Incluso un pequeño error puede provocar un gran accidente. Es aún más peligroso por la noche, cuando no hay nadie presente. La máquina de estados es la última puerta de seguridad entre la inteligencia artificial y los equipos. Incluso si la inteligencia artificial comete un error, el equipo no realiza operaciones peligrosas. La libertad del laboratorio autónomo solo se permite dentro de esta puerta.

Este dispositivo también tiene limitaciones. Los estados no definidos de antemano son difíciles de manejar. Las averías imprevistas no encajan en las casillas predeterminadas. Si un sensor envía una señal errónea, el juicio también será erróneo. Por lo tanto, una máquina de estados por sí sola no es

suficiente. Debe haber un ojo que observe el equipo en tiempo real y detecte anomalías. Ese ojo es la fusión de sensores que veremos en la siguiente sección.

3.4 Marco de ejecución robótica que resiste entornos experimentales desordenados (PUDA)

Esta sección aborda las dificultades que experimentan los robots en entornos experimentales reales. Un laboratorio es diferente de una fábrica. Los polvos fluyen, se aglomeran y adquieren electricidad estática. Los crisoles se desalinean poco a poco. El ruido se mezcla en los sensores. Esta sección examina un marco de ejecución robótica que resiste estos entornos desordenados. Como caso representativo, aborda la investigación de PUDA.

Los robots de laboratorio se enfrentan a tareas más difíciles que los robots de fábrica. En una fábrica, las mismas piezas llegan a posiciones fijas. En un laboratorio, las propiedades del polvo son diferentes cada vez. Algunos polvos fluyen bien. Algunos polvos se aglomeran y obstruyen la boquilla. Otros adquieren electricidad estática y se adhieren por todas partes. Los crisoles también se desvían ligeramente de su posición cada vez que se colocan.

Este desorden se denomina perturbación. Una perturbación es una interferencia no planificada. Si hay perturbaciones, el trabajo no se puede realizar únicamente con movimientos fijos. El robot debe estar equipado con ojos y sentidos. Los ojos son cámaras y los sentidos son sensores de fuerza. El robot lee esta información en tiempo real y corrige sus movimientos.

PUDA es un caso que aborda este entorno de ejecución robótica. La explicación de su nombre es Arquitectura Física de Dispositivos Unificados (Physical Unified Device Architecture). También se denomina arnés de hardware adaptado a la IA para laboratorios autónomos. Un arnés es un marco que agrupa y gestiona varios equipos en uno solo. Es una investigación inicial publicada en 2026. El equipo de investigación está formado por Zhenkun Ren, Hongzhao Tan, Jiaen Yi y Kedar Hippalgaonkar. Lo que este estudio demostró en la práctica fue un entorno de ejecución de línea de comandos fácil de manejar para la inteligencia artificial y un sistema de registro en su interior. La fusión de sensores y la recuperación de errores que se verán a continuación no son aspectos que este artículo haya demostrado en términos de rendimiento. Son la dirección hacia la que pretende avanzar dicho marco de ejecución.

El método de PUDA es ligeramente diferente de los sistemas anteriores. No prioriza pantallas pensadas para humanos. En su lugar, crea un entorno de línea de comandos fácil de manejar para agentes de inteligencia artificial. Cada equipo se presenta como una herramienta que se puede encontrar en la línea de comandos. Los comandos y las respuestas quedan guardados como registros estándar. Qué comando se envió a qué muestra y qué resultado se obtuvo se conectan como una cadena. Este historial se llama procedencia. La procedencia es el registro histórico de dónde y cómo provienen los datos.

Veamos por qué este registro histórico es valioso en la investigación de materiales para cohetes. Los componentes aeroespaciales exigen una estricta trazabilidad de calidad. Se debe registrar todo: con qué materias primas y bajo qué condiciones se fabricaron. Si surge un problema más adelante, la causa se busca a través de estos registros. Un sistema como PUDA registra automáticamente este historial. Hay menos omisiones que cuando los humanos escriben a mano. Los datos generados por el laboratorio autónomo también se vuelven confiables y utilizables.

Hay una razón por la que PUDA eligió el método de línea de comandos. Las pantallas para humanos son agradables a la vista. Sin embargo, resultan incómodas de manejar para los agentes de inteligencia artificial. Los botones en la pantalla son difíciles de presionar para las máquinas. Las herramientas de línea de comandos son fáciles de invocar directamente por la inteligencia artificial. La inteligencia artificial examina la situación, juzga y emite comandos. El equipo ejecuta esos comandos de manera

precisa y reproducible. Es como si los humanos hubieran separado el cerebro que planifica los experimentos de las manos que mueven los equipos. Al dividirlos de esta manera, incluso si se cambia el cerebro, las manos se siguen utilizando tal como están.

Este método también tiene limitaciones. El entorno de línea de comandos resulta incómodo para que los humanos lo visualicen de inmediato. Se deben crear nuevas herramientas de línea de comandos para cada equipo. Los equipos antiguos son difíciles de adaptar a este marco. Dejar registros estandarizados requiere esfuerzo de almacenamiento y gestión. Esta investigación también se encuentra todavía en una etapa temprana. Se necesitará más tiempo para que se utilice ampliamente en diversos laboratorios.

3.4.1 Combinar múltiples sensores para juzgar y revertir fallos

Esta subsección aborda la fusión de sensores, que combina la información de varios sensores para tomar decisiones. También examina el proceso de recuperación que revierte los fallos por sí mismo. Se analiza por qué estas dos funciones preservan la calidad experimental. Ambas son métodos ampliamente investigados en robótica. Los marcos de ejecución como PUDA de la sección anterior son recipientes destinados a albergar estos métodos. Aquí no examinamos los logros de un sistema específico, sino los principios de los métodos.

La fusión de sensores consiste en combinar información de múltiples sensores para emitir juicios. Las cámaras observan la posición y la forma. Los sensores de fuerza observan la fuerza ejercida sobre la muñeca del robot. Los sensores de temperatura y de presión comprueban si el equipo está seguro. Si solo se observa un tipo de sensor, se pasan por alto muchas cosas. Al observar varios sensores juntos, los juicios se vuelven precisos.

Como ejemplo, veamos la escena en la que un robot recoge un crisol. Si la cámara queda bloqueada, no puede ver la posición. En este momento, el sensor de fuerza detecta la colisión. Si la fuerza es pequeña y no se percibe, la cámara observa si se derrama polvo. Sensores diferentes cubren los puntos ciegos de los demás. Es similar a cómo los humanos usan las manos y los ojos juntos. En una habitación oscura se tantea con las manos, y en un lugar iluminado se mira con los ojos.

Si el robot choca contra algo, el valor del sensor de fuerza se dispara bruscamente. Si supera un valor límite establecido, el robot detiene inmediatamente su movimiento. Después de detenerse, vuelve a examinar la situación. Comprueba con la cámara si se ha derramado polvo o si el crisol se ha resbalado. Este proceso de reevaluación es el comienzo de la recuperación de errores.

La recuperación de errores suele desarrollarse en tres etapas. Primero, se detecta la anomalía para saber qué salió mal. Luego, el robot regresa al estado seguro inmediatamente anterior al fallo. Finalmente, se modifica ligeramente la posición o la fuerza de agarre y se intenta una vez más. Tras pasar por estas tres etapas, los pequeños errores se superan sin intervención humana.

La recuperación de errores también es importante para la calidad de los experimentos. Si se continúa con una muestra mal pesada, se acumulan datos deficientes. Los datos deficientes conducen a los modelos de inteligencia artificial en una dirección equivocada. Por eso, el laboratorio autónomo no oculta los fallos, sino que los registra. Los fallos registrados se convierten en datos para el siguiente aprendizaje. Si se sabe dónde y por qué falló, la próxima vez se evitará ese error.

Esta función también tiene limitaciones. El ruido también se mezcla en los valores de los sensores. Si el ruido se interpreta erróneamente como una anomalía, se detiene un experimento que funciona perfectamente. Si se detiene con demasiada frecuencia, los experimentos se ralentizan. Por el contrario, si los valores límite se establecen de forma demasiado flexible, se pasan por alto accidentes reales. Por tanto, la sensibilidad de este dispositivo debe calibrarse con cuidado. Este ajuste todavía depende en gran medida de la experiencia humana. La recuperación autónoma no es un dispositivo para eliminar a

los humanos. Es un dispositivo para reducir repeticiones peligrosas y llamar a los humanos cuando sea necesario.

3.5 Un laboratorio autónomo en funcionamiento real: el A-Lab de Berkeley

Esta sección examina un caso representativo de un laboratorio autónomo que funcionó en la práctica. Se trata del A-Lab del Laboratorio Nacional Lawrence Berkeley de Estados Unidos. Se analiza qué logró este caso y con qué obstáculos se encontró. También se señala cómo deben interpretarse las cifras de rendimiento.

El A-Lab es una plataforma de síntesis autónoma de materiales inorgánicos creada por el Laboratorio Nacional Lawrence Berkeley. Los materiales inorgánicos son sustancias compuestas por elementos como metales y oxígeno. Los óxidos y los fosfatos se incluyen aquí. El A-Lab seleccionó candidatos viables a partir de datos computacionales. Luego, manipuló los polvos con robots y los sometió a tratamiento térmico. Los resultados se confirmaron con equipos de difracción de rayos X. Este proceso se llevó a cabo casi sin intervención humana.

Es importante interpretar con precisión las cifras de rendimiento. El A-Lab se publicó en un artículo de Nature en 2023. La cifra de éxito publicada inicialmente se revisó posteriormente y se redujo. Otros químicos de materiales volvieron a analizar los patrones de difracción. En 2026, Nature publicó una corrección. La cifra corregida es de treinta y seis de cincuenta y siete objetivos. Este libro toma como referencia la cifra corregida. No utiliza el número más alto inicial.

Este proceso de corrección en sí mismo es una valiosa lección. El núcleo de la corrección consta de dos aspectos. Uno es el error de juicio que sufrió el análisis automático. La determinación del éxito estuvo a cargo del análisis automático de difracción de rayos X. Este análisis interpretó erróneamente las señales de las materias primas residuales no reaccionadas como señales del nuevo material objetivo. En apariencia, parecía que se había creado un nuevo cristal. Este error de juicio solo salió a la luz después de que los humanos examinaran minuciosamente los patrones de difracción. El otro aspecto es el significado del término «nuevo material». Al principio, se interpretó como si se hubieran creado materiales que nunca antes habían existido en el mundo. Al revisarlo, significaba que una parte considerable eran materiales que no figuraban en la base de datos de este laboratorio. Esto es diferente de un material que es el primero para toda la comunidad científica. La corrección señaló que los humanos reconfirmaron los éxitos reportados y dejaron algunos como indeterminados. Por lo tanto, el veredicto de un laboratorio autónomo no es la respuesta definitiva. Solo se convierte en un verdadero éxito después de someterse a análisis de precisión y pruebas de reproducibilidad. Si se omiten las verificaciones solo porque los experimentos autónomos son rápidos, los errores se acumularán.

Veamos el significado que este caso aporta a la investigación de materiales para cohetes. El A-Lab demostró que la síntesis autónoma no es una fantasía fuera del laboratorio. Las series de óxidos y cerámicas están vinculadas a los materiales de protección térmica de los cohetes. Se abrió el camino para explorar estos materiales más rápido que los humanos. Sin embargo, esto debe ir acompañado de rigor en la determinación del éxito. En los componentes aeroespaciales, incluso una pequeña impureza puede convertirse en un gran problema. Por lo tanto, la verificación es tan importante como la velocidad.

El A-Lab reveló tres obstáculos claros. En la siguiente subsección, examinaremos cada obstáculo en detalle. Estos obstáculos son problemas que otros laboratorios autónomos también experimentan. Conocer los obstáculos con precisión sirve de base para construir los próximos sistemas.

3.5.1 Configuración del sistema y rendimiento operativo

Esta subsección examina de qué componentes estaba compuesto el A-Lab. También analiza cómo encajaban y funcionaban esos componentes. Dibuja de manera concreta la imagen real de un laboratorio autónomo.

La estructura del A-Lab se puede dividir en tres partes. La capa de cálculo selecciona los materiales candidatos y los métodos de síntesis. En la capa de ejecución, los robots y los equipos fabrican realmente las muestras. En la capa de medición, los equipos de difracción de rayos X comprueban los resultados. Estas tres capas se suceden del cálculo a la ejecución, y de la ejecución a la medición, completando un ciclo.

La capa de cálculo seleccionó candidatos de una gran base de datos. Utilizó datos públicos de Materials Project. También utilizó datos computacionales de Google. Estos datos predicen sustancias que probablemente sean termodinámicamente estables. Estable significa que la sustancia no se descompone fácilmente y se mantiene. El A-Lab tomó como objetivo los materiales predichos de esta manera.

En la capa de ejecución había varios brazos robóticos. Los robots trasladaban las muestras entre las estaciones de pesaje, mezclado, prensado y cocción del polvo. Pesaban el polvo con precisión y lo mezclaban con un molino de bolas. Un molino de bolas es un dispositivo que muele y mezcla polvos en un recipiente que contiene bolas pequeñas. El polvo mezclado se prensaba para formar pequeños gránulos. Esos gránulos se cocían en un horno eléctrico a alta temperatura.

El equipo central de la capa de medición es el difractómetro de rayos X. La difracción de rayos X (X-ray Diffraction) es un método para observar la estructura cristalina irradiando rayos X. Cada cristal produce un patrón diferente de rayos X reflejados. Este patrón es como la huella dactilar de una sustancia. La inteligencia artificial lee este patrón para determinar si se ha creado el material objetivo. Si el patrón coincide con el objetivo, se cuenta como un éxito.

El A-Lab ejecutó este proceso continuamente durante varios días. Los experimentos continuaron incluso mientras los humanos dormían por la noche. Esto demuestra que el laboratorio autónomo pasó de ser una idea en un artículo científico a un equipo real. Se produjeron múltiples materiales objetivo sin intervención humana. Este logro permanece como un caso que abrió las puertas a la síntesis autónoma de materiales.

Sin embargo, no se trató de una operación completamente desatendida. Los investigadores humanos prepararon y supervisaron el sistema. Los experimentos fallidos se examinaron junto con humanos. La determinación del éxito también fue reconfirmada posteriormente por humanos. El laboratorio autónomo reduce el trabajo humano, pero no elimina a los humanos. Los humanos se encargan de juicios y verificaciones de mayor nivel. La máquina se encarga de las repeticiones detalladas. Esta división del trabajo es la realidad actual de los laboratorios autónomos.

3.5.2 Tres obstáculos que reveló el A-Lab

Esta subsección examina uno por uno los tres obstáculos que reveló el A-Lab. Estos obstáculos son desafíos comunes que los laboratorios autónomos deben superar. Conocer los problemas con precisión es fundamental para construir bien los próximos sistemas.

El primer obstáculo es la dificultad para juzgar los resultados de la síntesis. Algunos de los materiales contados como éxitos no eran monofásicos. Monofásico es un estado en el que solo existe puramente el cristal deseado. Algunas muestras dejaron materias primas sin reaccionar. Otras muestras contenían fases secundarias. Una fase secundaria es un cristal distinto del objetivo. Incluso si el patrón de

difracción es similar al del objetivo, pueden mezclarse impurezas. Por lo tanto, se debe verificar estrictamente si se trata de una sola fase.

El segundo obstáculo es el procesamiento de polvos. Las propiedades de los polvos varían según el tipo. Algunos polvos acumulan electricidad estática. Este es el caso del dióxido de titanio o de los nanopolvos. Estos polvos se adhieren al interior de la boquilla. Como resultado, la boquilla de dosificación se obstruye. Las materias primas que absorben fácilmente la humedad también son difíciles de manejar. Los dosificadores automáticos se detienen con frecuencia ante polvos tan diversos. Lo que un ser humano resolvería con unos ligeros golpes con los dedos detiene a una máquina.

El tercer obstáculo es el control ambiental. A-Lab operaba en un entorno expuesto a la atmósfera. En el aire hay oxígeno y humedad. Los materiales vulnerables al oxígeno y a la humedad no se pueden manipular en este tipo de entorno, ya que sus propiedades cambian o se descomponen nada más ser creados. Los materiales de sulfuro y haluro para baterías de estado sólido se enfrentan a este problema. Estos materiales quedan inservibles si entran en contacto con el aire. Por ello, el A-Lab expuesto a la atmósfera no pudo explorar estos materiales.

Estos obstáculos muestran con claridad por qué la automatización es difícil. Los seres humanos superan estos problemas gracias a la experiencia de sus ojos y manos. Si el polvo se apelmaza, golpean suavemente el recipiente. Si una muestra presenta anomalías, lo notan por el olor y el color. A una máquina hay que enseñarle estas experiencias una por una. Lo que una persona hace sin pensarlo resulta difícil para una máquina. El desarrollo de los laboratorios autónomos es el proceso de transferir estas pequeñas sensaciones a las máquinas.

Estos tres obstáculos están interconectados. Si el procesamiento del polvo es deficiente, la dosificación es inexacta. Si la dosificación falla, la evaluación de los resultados también se vuelve inestable. Si no hay control ambiental, existen materiales que ni siquiera se pueden sintetizar. Para expandir los laboratorios autónomos, es necesario superar estas barreras de manera conjunta. Corregir solo una de ellas difícilmente generará un gran avance.

En la investigación de materiales para cohetes, estos obstáculos se sienten aún mayores. Los nuevos materiales para baterías o compuestos especiales destinados a cohetes suelen ser altamente reactivos. También hay muchos materiales vulnerables al aire. Para manipular dichos materiales, el control ambiental debe ser prioritario. Con un laboratorio expuesto a la atmósfera es difícil abordar esta familia de materiales. Por lo tanto, el tercer obstáculo, el control ambiental, representa un reto especialmente grande en la investigación de materiales para cohetes.

Los esfuerzos para superar este obstáculo son el tema de la siguiente sección. Para manipular materiales sensibles al aire, se requiere un entorno sellado. Ese entorno es la caja de guantes. La siguiente sección examina los laboratorios autónomos que operan en entornos sellados. Se trata de un sistema dirigido directamente a superar el tercer obstáculo de A-Lab.

3.6 Búsqueda de materiales sensibles al aire en entornos sellados (A-Lab GPSS)

Esta sección aborda los laboratorios autónomos que operan en entornos sellados. Se trata de un sistema concebido para superar el tercer obstáculo visto en la sección anterior. Se examinan los métodos para manipular materiales sensibles al aire.

Una caja de guantes es una caja sellada equipada con guantes. En las paredes de la caja hay fijados unos gruesos guantes de goma. Las personas introducen las manos en los guantes desde el exterior para manipular el interior. Por lo general, la caja se llena con un gas inerte como el argón. Un gas inerte es un

gas que no reacciona fácilmente con otras sustancias. Por ello, en el interior de la caja apenas hay oxígeno ni humedad.

El nombre de este sistema es A-Lab GPSS. GPSS significa sistema de síntesis protegido por caja de guantes. Es la siguiente generación creada para superar las limitaciones del A-Lab expuesto a la atmósfera. Los equipos se colocaron dentro de una gran caja de guantes llena de argón de ultraalta pureza. Mantiene los niveles de humedad y oxígeno por debajo de una parte por millón. En su interior se realizan el almacenamiento de materias primas, la dosificación, el moldeo, el sellado y el tratamiento térmico. La clave reside en eliminar cualquier momento de contacto entre la muestra y el aire.

Este caso se constata en un preprint de 2026. El equipo de investigación está integrado por Pei Yu y otros. Es un estudio inicial que abordó conductores de espinela de haluro de litio sensibles al aire. El artículo indica que se fabricaron 352 muestras. Había 171 combinaciones posibles de pares metálicos. De estas, el 72 % se abordó mediante experimentos. Aquí, el 72 % no es la tasa de éxito; es la proporción de combinaciones posibles cubiertas experimentalmente. Este número no debe interpretarse erróneamente como una tasa de éxito.

El artículo señala que los resultados mejoraron a medida que avanzaban los experimentos. En las primeras 75 muestras, la proporción de buenos resultados fue baja. En las últimas 75 muestras, esa proporción aumentó. La inteligencia artificial aprendió de los experimentos anteriores y mejoró los posteriores. Esto demuestra que el aprendizaje en bucle cerrado funciona en la práctica. No obstante, dado que esta investigación se encuentra en una etapa temprana, no deben extraerse conclusiones generalizadas.

Veamos por qué este laboratorio autónomo sellado es valioso en la industria de los cohetes. Las naves espaciales necesitan baterías de alto rendimiento. Las baterías de estado sólido son candidatas seguras y de alta densidad energética. Su elemento central es el electrolito sólido. Sin embargo, los buenos electrolitos sólidos suelen ser sensibles al aire. Un laboratorio autónomo sellado explora estos materiales de forma segura. El mismo entorno se utiliza en la investigación de materiales para cohetes con alta reactividad.

3.6.1 Electrolitos sólidos de haluro de litio sensibles al aire

Esta subsección examina en detalle el material tratado aquí. Explica en términos sencillos qué es un electrolito sólido de haluro de litio. Se analiza por qué este material es vulnerable al aire y por qué es valioso.

Primero, aclaremos los términos relacionados con las baterías. El ion de litio es una pequeña partícula que transporta carga eléctrica dentro de la batería. Cuando la batería se carga y se descarga, estos iones se desplazan de un lado a otro. El electrolito es la vía por la que pasan estos iones. Las baterías comunes utilizan electrolitos líquidos. El líquido puede filtrarse o incendiarse.

Un electrolito sólido transforma esta vía en un sólido. El electrolito sólido debe permitir el paso fluido de los iones de litio, pero debe bloquear eficazmente los electrones. Un sólido que deja pasar los iones y bloquea los electrones es un buen electrolito. Los sólidos no presentan fugas y son resistentes al fuego. Por esta razón, las baterías de estado sólido se consideran baterías más seguras.

El electrolito de haluro de litio es un electrolito sólido que contiene elementos halógenos. Los halógenos son elementos como el cloro, el bromo y el yodo. En su fórmula química, el litio, los metales y los halógenos se combinan en proporciones fijas. Este material tiene el potencial de ser muy compatible con electrodos de alto voltaje. Si el voltaje es alto, se puede almacenar más energía en el mismo tamaño. Por ello, llama la atención como material para baterías de próxima generación.

El problema es que este material es muy vulnerable al agua. Esta propiedad se denomina delicuescencia. La delicuescencia es la propiedad de absorber la humedad del aire y disolverse por sí misma. Incluso al contacto con trazas de humedad, su fase se descompone. Durante la descomposición, pueden liberarse gases nocivos como el cloruro de hidrógeno. Por lo tanto, el almacenamiento de materias primas, la síntesis y las mediciones deben realizarse íntegramente en un entorno sellado.

Esta condición complica la labor del laboratorio autónomo. El interior de la caja sellada es estrecho e incómodo de manejar. El robot debe moverse con precisión en un espacio reducido. No se puede permitir la exposición al aire ni por un solo instante. La inteligencia artificial propone composiciones candidatas y condiciones de síntesis. El robot repite los mismos procedimientos sin exposición al aire. Los equipos analíticos verifican rápidamente la conductividad iónica y la fase cristalina. Todo este proceso se desarrolla de manera coordinada dentro del recinto sellado.

Medir la conductividad iónica también es crucial. La conductividad iónica es un valor que indica con qué facilidad se desplazan los iones de litio. Cuanto mayor sea este valor, mejor rinde la batería. El laboratorio autónomo mide de inmediato la conductividad de la muestra fabricada. Y establece ese valor como objetivo para la inteligencia artificial. Selecciona el siguiente experimento orientándose hacia composiciones con mayor conductividad. Al examinar conjuntamente la fase cristalina y la conductividad, se va acotando la búsqueda de un buen electrolito.

La exploración de este material aún presenta limitaciones. En un entorno sellado es difícil introducir y extraer equipos. Debido al espacio reducido, los equipos que se pueden utilizar son limitados. Los materiales delicuescentes quedan inutilizados ante el más mínimo error. Por eso, cada experimento exige extrema precaución. Aunque la automatización acelere el proceso, esta necesidad de cautela no disminuye. Los laboratorios autónomos sellados aún manejan una gama restringida de materiales.

3.6.2 Doble bucle que alterna entre inducción y abducción

Esta subsección aborda dos modos de pensamiento que utiliza el laboratorio autónomo: la inducción y la abducción. Se analiza por qué alternar entre ambos acelera la exploración. Este método constituye la estructura de razonamiento revelada por la investigación de A-Lab GPSS en esta sección.

La inducción es el pensamiento que busca reglas comunes a partir de múltiples resultados experimentales. Recopila patrones de difracción y resultados de medición de conductividad. Examina qué composiciones se sintetizaron con éxito. Compara varias muestras en paralelo y lee las tendencias. Analiza cómo cambia la conductividad a medida que se expande la red cristalina. Las reglas acumuladas de este modo sirven de base para los siguientes experimentos.

La abducción es el pensamiento que busca la causa cuando se encuentra un resultado inesperado. La abducción es la inferencia que selecciona la causa que mejor explica un resultado anómalo. Supongamos que en una muestra aparece una fase de impureza imprevista. Se evalúa si la materia prima no reaccionó por completo o si la temperatura fue demasiado baja. Se selecciona la causa más plausible entre ellas y se modifica el siguiente experimento. La abducción busca inmediatamente el siguiente paso a partir de un único resultado sorprendente. Su dirección es diferente de la inducción, que extrae reglas a partir de múltiples resultados.

El laboratorio autónomo alterna estos dos modos de pensamiento. Los cálculos de la teoría del funcional de la densidad (Density Functional Theory) reducen de antemano los candidatos viables. Este cálculo determina la energía de los materiales a nivel atómico. El experimento pone a prueba los candidatos seleccionados aquí. A medida que se acumulan los resultados experimentales, la inducción encuentra reglas. Si surge un resultado inesperado, la abducción señala la causa y cambia el rumbo. De este modo,

el cálculo y el experimento se refinan mutuamente. El pensamiento no fluye en una sola dirección, sino que gira como un bucle.

Veamos este método con un ejemplo sencillo. Imaginemos la escena de buscar el camino en una ciudad desconocida. Caminar varias veces y descubrir las reglas de los atajos es la inducción. Cuando el camino se bloquea repentinamente, intuir la causa y tomar un desvío es la abducción. Si solo se confía en las reglas, uno se detiene ante situaciones imprevistas. Si solo se siguen conjeturas sobre las causas, se pierde la visión general. Al utilizar ambas, se llega al destino con rapidez. El laboratorio autónomo también utiliza ambos modos de pensamiento para alcanzar rápidamente buenos materiales.

Veamos por qué este doble bucle es valioso en la investigación de materiales para cohetes. Los nuevos materiales no se pueden comprender en su totalidad solo mediante cálculos. El cálculo asume condiciones ideales. En la síntesis real existen impurezas y ruido. Tampoco se puede saber todo solo con experimentos. Los experimentos requieren mucho tiempo y costes elevados. Al combinar ambos métodos, el cálculo reduce el alcance y el experimento lo verifica. En la costosa búsqueda de materiales para cohetes, esta combinación demuestra un gran poder.

Este método también presenta limitaciones. El lenguaje del cálculo y el del experimento son diferentes entre sí. Alinear los resultados computacionales con los resultados experimentales resulta complejo. Si el cálculo es erróneo, se desperdician experimentos con candidatos equivocados. Si el ruido experimental es elevado, el cálculo se corrige de forma incorrecta. Por lo tanto, es necesario construir adecuadamente el puente que une ambos bucles. Tender este puente aún depende en gran medida del criterio humano.

3.6.3 Cooperación mediante la división de roles entre múltiples inteligencias artificiales

Esta subsección aborda el modo en que múltiples inteligencias artificiales cooperan dividiéndose las funciones. Esto se denomina multiagente. Se analiza por qué varios agentes son mejores que una sola inteligencia artificial y qué riesgos existen.

El sistema multiagente es una estructura en la que múltiples inteligencias artificiales se reparten los roles. Un agente es una unidad de inteligencia artificial que juzga y actúa de forma autónoma. Es equivalente a que un equipo se divida el trabajo en lugar de que una sola persona lo haga todo. En un laboratorio autónomo, esta división de roles es natural, ya que la experimentación requiere conocimientos de química, planificación, seguridad y análisis.

Al dividir los roles, cada agente se concentra en una sola tarea. El agente de química verifica si la ecuación de reacción es correcta. Comprueba si la masa se conserva y si los estados de oxidación coinciden. También examina si las materias primas reaccionan negativamente con el crisol. El agente de planificación establece la secuencia de los equipos. Divide de antemano tareas que suponen cuellos de botella, como tratamientos térmicos prolongados. Asimismo, determina qué tareas pueden ejecutarse en paralelo.

El agente de seguridad es el ojo que protege el laboratorio. Supervisa las concentraciones de oxígeno y humedad dentro de la caja de guantes. También observa las variaciones de presión y el flujo de agua de refrigeración del horno eléctrico. Verifica estos valores repetidamente en intervalos muy cortos. Si detecta un valor anómalo, emite una alerta y registra el evento. Sin embargo, este agente no ejecuta directamente la parada de emergencia real. Esto se debe a que un agente generativo puede demorarse en responder o equivocarse en su juicio. Si el oxígeno o la presión alcanzan valores peligrosos, sensores independientes y controladores de seguridad verificados detienen el equipo de inmediato. Este enclavamiento por hardware no espera la decisión del modelo de lenguaje. El agente se encarga de la monitorización y el registro, mientras que la detención queda en manos de dispositivos de seguridad

verificados. El agente de análisis interpreta los resultados de las mediciones. Coteja los patrones de difracción con la base de datos. Cuantifica numéricamente la cantidad de impurezas. E introduce ese valor numérico como valor objetivo en el modelo de aprendizaje.

En 2026, las investigaciones sobre laboratorios con sistemas multiagente aumentaron notablemente. Un estudio coordinó múltiples agentes de inteligencia artificial bajo un orquestador. Cada agente compartió y utilizó diferentes herramientas. Se dividieron en un grupo de búsqueda de conocimiento, un grupo para convertir el diseño en procedimientos de ejecución, un grupo para operar los robots y un grupo para interpretar los datos. Al agruparse de este modo, lograron una exploración de materiales en bucle cerrado. Otro estudio, ChatMat, conectó automáticamente el flujo de predicción de propiedades físicas mediante múltiples agentes. También se publicaron artículos que abordaban el significado del enfoque multiagente en la gestión de laboratorios autónomos.

Las ventajas de dividir las funciones son evidentes. Permite ver con claridad las tareas asignadas a cada agente. Si surge un problema, es fácil identificar en qué etapa se produjo. Modificar un agente no requiere alterar el sistema completo. También resulta sencillo incorporar nuevas herramientas. Esto equivale a las ventajas que ofrece la colaboración en un equipo humano.

Los sistemas multiagente también transforman la manera misma de hacer ciencia. Antes, los investigadores humanos registraban cada experimento uno por uno. Ahora, los agentes documentan el proceso de decisión como datos. Queda registrado por qué se eligieron estas condiciones y por qué se consideró un éxito este resultado. Este registro se utiliza posteriormente para revisar la investigación. Sin embargo, el juicio de la inteligencia artificial no siempre es transparente. Resulta difícil para los humanos comprender plenamente por qué se tomó tal decisión. Esta falta de transparencia sigue siendo una tarea pendiente por resolver en el futuro.

Sin embargo, los riesgos también aumentan. Los modelos de lenguaje grandes pueden inventar hechos inexistentes. El error de un agente puede propagarse a otros agentes. Si múltiples agentes confían entre sí y se transmiten los datos, los errores no se detectarán. Por esta razón, la autoridad final debe recaer en las reglas de seguridad y en los supervisores humanos. Aunque la inteligencia artificial se exprese con confianza, los equipos pueden ser peligrosos. Cuantos más agentes haya, más importantes se vuelven la verificación cruzada humana y los dispositivos de seguridad de hardware.

Resumen

En este capítulo se ha examinado cómo funcionan los laboratorios autónomos. El núcleo es una estructura de circuito cerrado. La inteligencia artificial elige el siguiente experimento. Los robots lo fabrican. Los instrumentos lo miden. Los resultados regresan a la inteligencia artificial. Este ciclo funciona día y noche, reduciendo rápidamente el amplio espacio de materiales.

El método por el cual la inteligencia artificial selecciona el siguiente experimento es el aprendizaje activo bayesiano. Este método evalúa conjuntamente el rendimiento previsto y la incertidumbre. Elige de manera equilibrada tanto las áreas prometedoras como las desconocidas. Por ello, alcanza materiales óptimos con un menor número de experimentos. Esto representa una gran ventaja en la costosa búsqueda de materiales para cohetes.

Los métodos de síntesis descritos en los artículos científicos son extraídos por modelos de lenguaje grandes. Investigaciones como MSP-LLM realizan esta tarea. Los métodos extraídos se convierten de manera segura a lenguajes de planificación como PDDL. El lenguaje de planificación sirve como una barrera que previene de antemano los accidentes del robot. La inteligencia artificial y los equipos se interconectan mediante estándares como el protocolo de contexto de modelos. El controlador NIMO es

un ejemplo de ello. La máquina de estados actúa como la última compuerta que bloquea las órdenes peligrosas.

Los robots deben tolerar entornos experimentales desordenados. La fusión de sensores, que lee múltiples sensores simultáneamente, contribuye a ello. La recuperación de errores, que subsana fallos de forma autónoma, también opera al mismo tiempo. Sistemas como PUDA registran las órdenes y los resultados. Este historial resulta muy valioso para la trazabilidad de la calidad de los componentes aeroespaciales.

Como casos prácticos, se examinaron A-Lab y A-Lab GPSS. A-Lab sintetizó treinta y seis de cincuenta y siete objetivos durante diecisiete días. Esta cifra corresponde a un valor corregido en 2026. Este caso demuestra que las determinaciones de los laboratorios autónomos también requieren la revisión humana. A-Lab GPSS manipuló materiales vulnerables al aire en un entorno sellado. Con trescientas cincuenta y dos muestras, cubrió el setenta y dos por ciento de las combinaciones posibles de pares metálicos. Este número no representa una tasa de éxito, sino la cobertura combinatoria.

Al llegar 2026, este campo se ha expandido rápidamente. Han aumentado los laboratorios multiagente en los que varias inteligencias artificiales se dividen las funciones. También se han sucedido intentos de conectar equipos mediante comunicaciones estandarizadas. Asimismo, fue intensa la investigación para adaptar la optimización bayesiana a los flujos de trabajo experimentales. Se publicaron artículos de revisión y manuales que analizan exhaustivamente los laboratorios autónomos. La tendencia general es clara: liberar la mano del ser humano de las repeticiones detalladas. Sin embargo, el lugar del juicio y la validación sigue perteneciendo a las personas.

La mayoría de los sistemas presentados en este capítulo son investigaciones iniciales. Abundan los artículos preliminares y los estándares de campo aún son escasos. Por ello, en lugar de poner por delante las cifras de rendimiento, se examinaron conjuntamente los principios y las limitaciones. El laboratorio autónomo no es una herramienta para sustituir al ser humano. Las máquinas asumen las repeticiones peligrosas, mientras que las personas se encargan del juicio y la verificación. El verdadero sentido del laboratorio autónomo no radica en la rapidez del experimento, sino en su fiabilidad. Solo sobre este principio puede desarrollarse con seguridad la automatización en la investigación de materiales para cohetes.

Fuentes y referencias relacionadas

King, R. D., Rowland, J., Oliver, S. G., Young, M., Aubrey, W., Byrne, E., Liakata, M., Markham, M., Pir, P., Soldatova, L. N., Sparkes, A., Whelan, K. E., & Clare, A. (2009). The automation of science. Science, 324(5923), 85-89. https://doi.org/10.1126/science.1165620

MacLeod, B. P., Parlane, F. G. L., Morrissey, T. D., Häse, F., Roch, L. M., Dettelbach, K. E., Moreira, R., Yunker, L. P. E., Rooney, M. B., Deeth, J. R., Lai, V., Ng, G. J., Situ, H., Zhang, R. H., Elliott, M. S., Haley, T. H., Dvorak, D. J., Aspuru-Guzik, A., Hein, J. E., & Berlinguette, C. P. (2020). Self-driving laboratory for accelerated discovery of thin-film materials. Science Advances, 6(20), eaaz8867. https://doi.org/10.1126/sciadv.aaz8867

Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., Milsted, D., McDermott, M. J., Gallant, M., Cubuk, E. D., Merchant, A., Kim, H., Jain, A., Bartel, C. J., Persson, K. A., Zeng, Y., & Ceder, G. (2023). An autonomous laboratory for the accelerated synthesis of inorganic materials. Nature, 624(7990), 86-91. https://doi.org/10.1038/s41586-023-06734-w

Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., Milsted, D., McDermott, M. J., Gallant, M., Cubuk, E. D., Merchant, A., Kim, H., Jain, A., Bartel, C. J., Persson, K. A., Zeng, Y., & Ceder, G. (2026). Author correction: An autonomous laboratory for the accelerated synthesis of inorganic materials. Nature, 650, E1. https://doi.org/10.1038/s41586-025-09992-y

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. Nature, 624(7990), 80-85. https://doi.org/10.1038/s41586-023-06735-9

McDermott, M. J., Dwaraknath, S. S., & Persson, K. A. (2021). A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. Nature Communications, 12, 3097. https://doi.org/10.1038/s41467-021-23339-x

Tom, G., Schmid, S. P., Baird, S. G., Cao, Y., Darvish, K., Hao, H., Lo, S., Pablo-García, S., Rajaonson, E. M., Skreta, M., Yoshikawa, N., Corapi, S., Akkoc, G. D., Strieth-Kalthoff, F., Seifrid, M., & Aspuru-Guzik, A. (2024). Self-driving laboratories for chemistry and materials science. Chemical Reviews, 124(16), 9633-9732. https://doi.org/10.1021/acs.chemrev.4c00055

Boiko, D. A., MacKnight, R., Kline, B., & Gomes, G. (2023). Autonomous chemical research with large language models. Nature, 624, 570-578. https://doi.org/10.1038/s41586-023-06792-0

Noh, H., Na, G. S., Lee, N., & Park, C. (2026). MSP-LLM: A unified large language model framework for complete material synthesis planning. arXiv preprint, arXiv:2602.07543. https://arxiv.org/abs/2602.07543

Yoshikawa, N., & Tamura, R. (2026). NIMO Controller: a self-driving laboratory orchestrator based on the Model Context Protocol. arXiv preprint, arXiv:2605.15227. https://arxiv.org/abs/2605.15227

Tamura, R., & Yoshikawa, N. (2026). NIMO: A software platform for closed-loop materials exploration with diverse AI algorithms. arXiv preprint, arXiv:2606.15522. https://arxiv.org/abs/2606.15522

Ren, Z., Tan, H., Yee, J., & Hippalgaonkar, K. (2026). PUDA: An AI-native hardware harness for self-driving laboratories. arXiv preprint, arXiv:2607.26464. https://arxiv.org/abs/2607.26464

Fei, Y., Rendy, B., Yang, X., Woo, J., Huang, X., Li, C., Wang, S., Milsted, D., Zeng, Y., & Ceder, G. (2026). Agentic LLM reasoning in a self-driving laboratory for air-sensitive lithium halide spinel conductors. arXiv preprint, arXiv:2604.11957. https://arxiv.org/abs/2604.11957

Kusne, A. G., & McDannald, A. (2026). Managing autonomous materials labs with multi-agent AI and its implications for the science of science. Communications Materials, 7, 173. https://doi.org/10.1038/s43246-026-01219-5

Lv, S., Peng, L., Jiao, S., Yao, Y., Wu, W., & Hu, W. (2026). ChatMat: a multi-agent chemist for autonomous material prediction and exploration. Digital Discovery, 5(7), 2886-2898. https://doi.org/10.1039/d5dd00582e

Lee, H., Yoo, H. J., Jang, H. S., Park, B., Park, Y. J., & Han, S. S. (2026). Toward self-driving laboratory 2.0 for chemistry and materials discovery. Materials Horizons, 13(10), 4712-4739. https://doi.org/10.1039/d5mh01984b

Zhu, J., Zhang, L., Zhu, Y., Lin, X., Wu, Y., Di, S., Liu, B., Luo, Y., & Zhang, T. (2026). A survey of agentic materials science and engineering: where are we and where are we going? Journal of Materials Informatics, 6, 7. https://doi.org/10.20517/jmi.2026.07

Tobias, A. V., & Wahab, A. (2025). Autonomous 'self-driving' laboratories: a review of technology and policy implications. Royal Society Open Science, 12(7), 250646. https://doi.org/10.1098/rsos.250646

Wakabayashi, Y. K., & Otsuka, T. (2026). Bayesian optimization for self-driving materials laboratories: From algorithms to physics-informed workflows. arXiv preprint, arXiv:2608.26016. https://arxiv.org/abs/2608.26016

Ntagiantas, A., Tsilimidos, P., Giannakopoulos, G., Rekatsinas, C., & Krokidas, P. (2026). Active learning guided design space refinement for scalable multi-objective Bayesian optimization in materials discovery. arXiv preprint, arXiv:2608.04651. https://arxiv.org/abs/2608.04651

He, C., Singull, M., & Jacobsson, T. J. (2026). Bayesian optimisation for the experimental sciences: A practical guide to data-efficient optimisation of laboratory workflows. Advanced Intelligent Systems, 8(5), e202501149. https://doi.org/10.1002/aisy.202501149

Liang, H., Sun, Y., Paxson, R., Lee, C.-Y., Hall, A. T., Warecki, Z., Cumings, J., Koinuma, H., Kusne, A. G., Lippmaa, M., & Takeuchi, I. (2026). Autonomous epitaxial atomic-layer synthesis via real-time computer vision of electron diffraction. arXiv preprint, arXiv:2602.20432. https://arxiv.org/abs/2602.20432

Anthropic. (2025). Model Context Protocol specification (version 2025-06-18). https://modelcontextprotocol.io/specification/2025-06-18

Anthropic. (2024). Introducing the Model Context Protocol. https://www.anthropic.com/news/model-context-protocol

Capítulo 4. Predicción de coherencia física y optimización de la composición de aleaciones de alta entropía (HEA) y superaleaciones

Introducción

Cuando se usa una olla en la cocina durante mucho tiempo, el fondo se ennegrece. La parte que está expuesta con mayor frecuencia al fuego es la primera en deteriorarse. Cuanto más caliente está un metal, más blando y débil se vuelve. Esta propiedad de debilitarse ante el calor la comparten todos los metales por igual.

Los motores cohete sufren este problema con mayor severidad que cualquier otra máquina. En el interior de la cámara de combustión, el combustible arde y genera gases a miles de grados. Ese gas se expulsa rápidamente a través de un estrecho conducto llamado tobera. La pared de la tobera debe soportar este gas caliente justo a su lado. Es un entorno hostil que difícilmente puede compararse con el fondo de una olla.

Los vehículos que reingresan a la atmósfera enfrentan el mismo desafío. Al penetrar el aire a gran velocidad, este se comprime y se calienta. La parte delantera del vehículo se pone al rojo vivo. Si no puede soportar este calor, la estructura se derrite o se fractura.

Este calentamiento del aire se puede percibir en la palma de la mano. Al andar rápido en bicicleta, el viento golpea con fuerza la cara. Al sacar la mano por la ventana, el viento la empuja. Un vehículo aéreo recibe ese viento miles de veces más rápido. El aire desplazado se comprime y se calienta, y ese calor se transfiere a la superficie frontal. La temperatura del punto de estancamiento en la superficie de reentrada supera los 1.600 grados Celsius.

Los vehículos hipersónicos se encuentran con el mismo calor. La velocidad hipersónica es más de cinco veces más rápida que el sonido. Al volar a esta velocidad, los bordes de ataque de las alas están más calientes que otras partes. Esto se debe a que el calor se concentra en las zonas estrechas y afiladas. Por eso, en estas secciones se emplean materiales con una excelente capacidad para soportar altas temperaturas.

Por ello, se buscan metales que no pierdan su resistencia incluso en entornos extremadamente calientes. El metal que ha ocupado ese lugar durante mucho tiempo es la superaleación de base níquel (Nickel-based superalloy). Las superaleaciones son aleaciones diseñadas específicamente para resistir bien a altas temperaturas. Sin embargo, este metal también tiene un límite de temperatura.

Este capítulo aborda nuevos metales que buscan superar ese límite. El protagonista es la aleación de alta entropía (High-Entropy Alloy, HEA). Es una aleación en la que se mezclan varios metales en proporciones similares. Entre ellas, si se seleccionan y mezclan únicamente metales de alto punto de fusión, se obtienen nuevos candidatos para materiales de cohetes. A este tipo de aleaciones se les denomina aleaciones refractarias de alta entropía (Refractory High-Entropy Alloy, RHEA).

Este capítulo explica un nuevo método para encontrar tales metales. En lugar de que los humanos experimenten uno por uno, la inteligencia artificial (AI) selecciona rápidamente a los candidatos. Se desglosa paso a paso cómo la disposición de los átomos, las leyes de la termodinámica y los cálculos atómicos a gran escala se encuentran con la inteligencia artificial. Aunque trata sobre metales, se explica mediante ejemplos cercanos a nosotros, como las estaciones del año y la escuela.

Veamos de antemano por qué es necesaria la inteligencia artificial. Para crear una buena aleación, es necesario determinar los tipos de elementos y sus proporciones. Incluso con solo cinco elementos, el número de combinaciones de proporciones resulta incontable. Si se fabrica y prueba cada una por separado, una vida entera no bastaría para verlas todas. La inteligencia artificial reduce previamente este amplio abanico de candidatos.

El papel de la inteligencia artificial se asemeja al de un guía. Un guía que conoce el camino en la montaña enseña los atajos. Sin embargo, alcanzar la cima es, en última instancia, tarea del ser humano. Cuando la inteligencia artificial reduce los candidatos, el ser humano los confirma mediante experimentos. Este capítulo muestra a través de ejemplos cómo se lleva a cabo esta colaboración.

4.1 Requisitos para entornos aeroespaciales extremos y límites estructurales de las aleaciones de alta temperatura

Lo que se aborda en esta sección

Esta sección describe primero cuán severo es el entorno térmico al que se enfrentan los cohetes y los vehículos de reentrada. Luego, examina dónde encuentra sus límites la superaleación de base níquel utilizada durante tanto tiempo. Por último, explica por qué las aleaciones refractarias de alta entropía han surgido como una nueva alternativa.

En pocas palabras, esta sección es el espacio donde se define el problema. Solo conociendo con precisión el problema se puede apreciar la utilidad de las herramientas de inteligencia artificial que se presentarán más adelante. Por ello, se da prioridad a comprender la situación antes que a las cifras.

La cámara de combustión de un cohete es un recipiente sometido a una presión muy elevada. En su interior, el combustible y el oxidante arden juntos, generando una enorme presión. Esta presión alcanza decenas de megapascales. La presión del neumático de un automóvil es de aproximadamente 0,2 megapascales. La cámara de combustión de un cohete soporta cientos de veces esa presión.

El gas generado en la cámara de combustión se expulsa por la tobera. La temperatura de este gas puede ascender a miles de grados según los materiales. Sin embargo, la temperatura del metal de la pared no sube tanto como la del gas. La refrigeración y la capa límite bloquean el calor y mantienen baja la temperatura del metal. Aun así, la cantidad de calor que recibe la pared es inmensa. Si no soporta este calor, la pared se funde, se perfora y el motor queda inservible.

El calor no actúa una sola vez. El motor se enciende y se apaga repetidamente. Al encenderse se calienta súbitamente y al apagarse se enfría de golpe. A estas fluctuaciones bruscas de temperatura se les llama choque térmico (Thermal shock). Es el mismo principio por el cual un vaso de vidrio se agrieta al verter agua caliente de forma repentina.

A esto se suma la fuerza mecánica. Cuando un metal se somete a una fuerza durante mucho tiempo a altas temperaturas, se estira lentamente. Este fenómeno se denomina fluencia (Creep). Es similar a cómo un caramelo blando sostenido durante mucho tiempo termina cediendo por sí solo. A altas temperaturas, esta fluencia y las cargas cíclicas actúan conjuntamente para degradar el material.

La superaleación de base níquel es un metal que ha resistido en tales entornos durante mucho tiempo. La resistencia de este metal proviene de su microestructura. Dentro de la fase matriz gamma (γ), pequeños granos llamados gamma prima (γ′) están incrustados densamente. Estos granos actúan como obstáculos que impiden la deformación del metal. Es como pequeños frutos secos incrustados en una masa que evitan que esta se estire con facilidad.

El tipo de estos obstáculos difiere ligeramente según la superaleación. El Inconel 718 (Inconel 718), con alto contenido de niobio (Nb), es un ejemplo de ello. En esta aleación, una fase llamada gamma doble prima (γ″, Ni3Nb) aporta mayor resistencia que la gamma prima. Aunque el nombre del obstáculo sea diferente, el principio es el mismo: los pequeños granos impiden la deformación y endurecen el metal.

Los cohetes utilizan diferentes metales según la ubicación. En la cámara de combustión, donde el combustible refrigera la pared, se usan con frecuencia aleaciones de cobre, debido a su excelente capacidad para disipar el calor rápidamente. Las superaleaciones de base níquel se emplean en componentes sometidos a altas temperaturas durante períodos prolongados, como las turbinas. Para la garganta de la tobera o las superficies de reentrada, que son aún más calientes, los candidatos son metales refractarios o cerámicas. Este capítulo aborda los nuevos metales destinados a los lugares con las temperaturas más extremas.

El problema surge cuando la temperatura sube aún más. Por encima de aproximadamente 1.150 grados, estos pequeños granos crecen o se desmoronan. Esto se conoce como engrosamiento o coalescencia. Al desaparecer los obstáculos, el metal se deforma con facilidad y su resistencia cae abruptamente. Dado que el punto de fusión de las superaleaciones de base níquel ronda los 1.350 grados, el límite operativo efectivo resulta estar muy por debajo de su punto de fusión.

Además, pueden formarse fases perjudiciales. Al permanecer a altas temperaturas durante mucho tiempo, los átomos se agrupan siguiendo patrones específicos y forman cúmulos duros y quebradizos. Estas fases se denominan fases perjudiciales. Cuando se forman fases perjudiciales, el metal se fractura fácilmente. Cuanto más avanzada es una superaleación, más se diseña para suprimir la formación de tales fases.

Debido a estos problemas, las superaleaciones de base níquel requieren refrigeración. En los motores reales, se crean pequeños canales en el interior de los componentes. Al hacer fluir combustible frío por esos canales, se enfrían las paredes. Gracias a esta refrigeración, los componentes pueden utilizarse a temperaturas inferiores a su punto de fusión. Sin embargo, añadir canales de refrigeración complejiza la estructura de los componentes y aumenta su peso.

El objetivo de las aleaciones refractarias de alta entropía es reducir esta refrigeración. Si el propio material puede soportar temperaturas más elevadas, se necesita menos refrigeración. Al reducir la refrigeración, la estructura se vuelve más ligera y sencilla. Los investigadores se fijan como meta en el desarrollo de nuevas aleaciones materiales capaces de resistir por encima de 1.400 grados sin refrigeración.

Por esta razón, se buscó un camino completamente diferente. La idea consiste en construir la estructura base con metales que posean por sí mismos un alto punto de fusión. Metales como el niobio (Nb), el molibdeno (Mo), el tántalo (Ta) y el tungsteno (W) pertenecen a este grupo. Sus puntos de fusión superan con creces los 2.000 grados. A estos metales se les conoce como metales refractarios (Refractory metal).

Las aleaciones refractarias de alta entropía son aleaciones en las que se mezclan varios de estos metales refractarios en cantidades similares. Investigaciones recientes informan que estas aleaciones muestran una resistencia superior a la de las superaleaciones de base níquel incluso en ensayos de compresión a 1.600 grados. Por ello, se consideran nuevas candidatas para la propulsión de cohetes, los bordes de ataque de vehículos hipersónicos y los escudos térmicos de reentrada.

Sin embargo, no todo son ventajas. Estas aleaciones tienden a ser quebradizas a temperatura ambiente. A esto se le llama baja ductilidad. Además, cuando entran en contacto con el oxígeno a altas temperaturas, experimentan una oxidación en la que la superficie se desmenuza y se desmorona. Estos

dos aspectos son tareas pendientes que deben resolverse antes de utilizarlas en componentes reales. La inteligencia artificial se emplea precisamente para resolver estas tareas con rapidez.

Examinemos el problema de la oxidación un poco más a fondo. Las aleaciones ricas en tungsteno o molibdeno reaccionan violentamente con el oxígeno a altas temperaturas. La capa de óxido superficial no se adhiere firmemente y se desprende como polvo. A esta oxidación que se desmorona se le denomina oxidación por plaga (Pest oxidation). El nombre proviene de la idea de que carcome la superficie como las plagas que devoran los granos.

La forma de prevenir este problema es una capa de óxido protectora. Si se añade una cantidad adecuada de aluminio o cromo, se forma una capa de óxido densa en la superficie. Esta capa impide que el oxígeno penetre hacia el interior. Es similar a cómo una capa adherida al fondo de una olla protege el metal interior. Cuánto y cómo añadir estos elementos para que se forme una buena capa constituye el núcleo del diseño.

La compatibilidad con el recubrimiento también es importante. Sobre la aleación se suele aplicar un recubrimiento resistente a la oxidación. El problema radica en que el metal y el recubrimiento se expanden de manera diferente según la temperatura. A este grado de expansión se le llama coeficiente de dilatación térmica. Si los coeficientes de dilatación térmica de ambos materiales difieren significativamente, la interfaz se fractura cuando la temperatura aumenta. Por lo tanto, la aleación no solo debe poseer resistencia mecánica, sino también ser compatible con el recubrimiento.

De este modo, las aleaciones para el revestimiento de toberas deben satisfacer múltiples condiciones a la vez. Se requieren un alto punto de fusión, baja densidad, resistencia a altas temperaturas, resistencia a la oxidación y compatibilidad con recubrimientos. En muchos casos, estas condiciones entran en conflicto entre sí. Resolver estos problemas entrelazados únicamente mediante experimentos humanos resulta abrumador. La inteligencia artificial interviene como una herramienta para ordenar con rapidez estos complejos problemas.

4.2 Modelado físico de los cuatro efectos centrales y termodinámica del equilibrio de fases

Lo que se aborda en esta sección

Esta sección trata el marco explicativo inicial de por qué las aleaciones de alta entropía son especiales. Cuando este campo se abrió por primera vez, los investigadores propusieron cuatro efectos. A estos cuatro se les conoce como los cuatro efectos centrales: entropía configuracional, distorsión de la red, difusión lenta y efecto cóctel.

Estos cuatro efectos no son leyes físicas consolidadas, sino más bien hipótesis iniciales. Como se verá más adelante, la difusión lenta y la estabilización de una sola fase debido a la alta entropía no han sido demostradas como leyes generales. Aun así, es necesario conocer este marco para entender qué intenta calcular la inteligencia artificial. Desglosaremos cada uno mediante ejemplos de la vida cotidiana.

Primero, veamos qué es una aleación de alta entropía. Las aleaciones convencionales añaden pequeñas cantidades de otros metales a un metal principal. En el acero, el hierro es el protagonista y se añade una pequeña cantidad de carbono. Las aleaciones de alta entropía difieren de este enfoque. En ellas se mezclan cinco o más metales en proporciones similares. Es como tener múltiples protagonistas.

Este concepto fue presentado de forma independiente en 2004 por dos equipos de investigación: el equipo de Yeh en Taiwán y el equipo de Cantor en el Reino Unido. Demostraron que se pueden crear metales estables sin depender de un único elemento principal. Este descubrimiento abrió las puertas al diseño de nuevas aleaciones.

El primer principio es la entropía configuracional (Configurational entropy). La entropía es un valor relacionado con el número de posibles formas de mezcla. Cuantas más maneras haya de guardar objetos en un casillero, mayor será el desorden. Al mezclar varios metales en cantidades similares, las formas de disponer los átomos se vuelven extraordinariamente numerosas. Este gran desorden contribuye a mantener el metal en un estado homogéneamente mezclado.

A este estado mezclado se le denomina solución sólida (Solid solution). Una solución sólida es un estado en el que múltiples elementos se disuelven uniformemente dentro de una misma red cristalina. Se asemeja a cómo el azúcar se disuelve en agua para formar un único líquido. Una alta entropía impide que los elementos se agrupen por separado formando compuestos duros e intermetálicos. Por ello, las aleaciones de alta entropía tienden a formar soluciones sólidas en lugar de compuestos.

El segundo principio es la distorsión de la red (Lattice distortion). Los átomos metálicos tienen diferentes tamaños. Cuando átomos grandes y pequeños se mezclan en una misma red, esta se deforma y se vuelve irregular. Es como un muro construido con piedras de diferentes tamaños en lugar de ladrillos lisos y uniformes. Esta red distorsionada actúa como una barrera que frena la deformación.

La deformación ocurre cuando pequeños defectos llamados dislocaciones (Dislocation) se deslizan a través de la red. Si la red es irregular, a las dislocaciones les resulta difícil deslizarse. Es como la dificultad de empujar una carretilla por un terreno de grava. Por lo tanto, las aleaciones con gran distorsión de red no se deforman fácilmente y resultan duras.

El tercer principio es la difusión lenta (Sluggish diffusion). La difusión es el movimiento por el cual los átomos cambian de posición. En una aleación de alta entropía, el entorno de átomos vecinos varía para cada átomo. Algunas posiciones son fáciles de ocupar y otras son difíciles. Se considera que esta irregularidad ralentiza, en promedio, el desplazamiento de los átomos.

Cuando los átomos se mueven lentamente, la microestructura no cambia fácilmente incluso a altas temperaturas. Por ello, los granos cristalinos crecen despacio y pueden ofrecer resistencia a la fluencia. Sin embargo, este efecto no siempre se manifiesta con fuerza en todas las aleaciones de alta entropía. La velocidad de difusión real varía según la combinación de elementos, la temperatura y la estructura cristalina. Por esta razón, las investigaciones recientes tratan este efecto no como una ley general, sino como una propiedad que debe verificarse según cada caso.

El cuarto principio es el efecto cóctel (Cocktail effect). Un cóctel es una bebida que mezcla varios ingredientes para crear un nuevo sabor. El sabor tras la mezcla no es igual a la simple suma de los sabores de cada ingrediente. Lo mismo ocurre con los metales. Cuando múltiples elementos coexisten, surgen propiedades difíciles de explicar mediante la suma individual de las propiedades de cada elemento.

Esta propiedad no lineal se manifiesta en la resistencia mecánica, la resistencia a la corrosión y la tenacidad. El problema radica en que resulta difícil para el ser humano prever de antemano este efecto de combinación. Esto se debe a que la cantidad de combinaciones de elementos es inmensa. La inteligencia artificial se utiliza como una herramienta para descubrir estos efectos de combinación ocultos a partir de los datos.

Estos cuatro principios se vinculan directamente con los materiales para cohetes. La entropía configuracional ayuda a evitar que la microestructura se fracture a altas temperaturas. La distorsión de la red aumenta la resistencia. La difusión lenta puede resultar ventajosa frente a la fluencia. El efecto cóctel aporta beneficios inesperados como la resistencia a la oxidación. Estos cuatro principios deben actuar en conjunto para soportar entornos extremos.

Sin embargo, estos principios no son una llave maestra. En algunas combinaciones se manifiestan con fuerza y en otras con debilidad. Por ello, aun conociendo los principios, es necesario verificar cada composición real una por una. Los principios indican la dirección, mientras que los cálculos y los experimentos se encargan de la confirmación. La inteligencia artificial acelera notablemente esta verificación.

Estos cuatro principios convergen en última instancia en una sola pregunta: qué composición permanecerá estable en una sola fase. La disciplina que responde a esta pregunta es la termodinámica del equilibrio de fases. En las dos subsecciones siguientes se examinan dos indicadores que ayudan a realizar este juicio.

4.2.1 El parámetro omega para medir la probabilidad de permanecer en una sola fase

La termodinámica es la ciencia que estudia la energía y la estabilidad. La estabilidad de un estado se determina mediante la energía libre (Free energy). Un estado con menor energía libre es más estable. Así como el agua fluye de lugares altos a lugares bajos, la materia tiende a desplazarse hacia un estado de menor energía libre.

La energía libre queda determinada por dos valores de manera conjunta. Uno es la entalpía (Enthalpy). La entalpía es un valor relacionado con la afinidad o repulsión mutua entre los átomos. Los átomos con buena afinidad mutua tienden a mezclarse, mientras que los que se repelen tienden a separarse. El otro es la entropía vista anteriormente. La entropía representa el desorden de la mezcla.

La pugna entre ambos factores varía con la temperatura. Cuanto más alta es la temperatura, mayor es la fuerza de la entropía. Es similar a cómo las personas se dispersan con mayor libertad en una habitación calurosa. Por ello, a altas temperaturas, una solución sólida desordenada tiende a estabilizarse con facilidad. Esta es la razón por la que las aleaciones de alta entropía se conservan bien en una sola fase a altas temperaturas.

La entropía configuracional aumenta cuando los elementos se mezclan en proporciones equimolares. La proporción equimolar consiste en incorporar cada elemento en la misma cantidad. Al mezclar cinco elementos en cantidades iguales, la entropía configuracional alcanza su valor máximo. Cuanto mayor es el número de elementos, mayor resulta este valor. De esta propiedad de poseer una elevada entropía configuracional proviene el nombre de alta entropía.

El parámetro omega es un valor que permite estimar de forma sencilla la probabilidad de formar una solución sólida monofásica. Este valor considera tres factores a la vez: el punto de fusión promedio, la entropía configuracional y la magnitud de la entalpía de mezcla. Cuando la influencia favorable de la entropía supera el impacto negativo de la entalpía, el valor de omega aumenta. Si omega es grande, se considera fácil que permanezca en una sola fase.

A esto se añade la comprobación de la diferencia de tamaño atómico. Si los tamaños atómicos difieren demasiado, la red no lo soporta y se fractura. Por lo tanto, la diferencia de tamaño debe ser inferior a un valor determinado para que se forme adecuadamente una solución sólida monofásica. Los investigadores emplean la regla empírica de que se forma una solución sólida monofásica cuando omega es grande y la diferencia de tamaño es pequeña.

Esta regla es práctica, pero no es una herramienta definitiva. Las aleaciones reales se ven afectadas por el ordenamiento local, la velocidad de enfriamiento, las impurezas y el tratamiento térmico. Incluso con la misma composición, la microestructura varía según el método de fabricación. Por esta razón, la regla de omega se utiliza únicamente para un cribado inicial.

El dictamen preciso queda a cargo de métodos que requieren un mayor volumen de cálculo. El método representativo es CALPHAD. CALPHAD es un método que calcula el equilibrio de fases acumulando datos de la energía libre de múltiples fases. A esto se suman los cálculos de la teoría del funcional de la densidad (Density Functional Theory, DFT) y los experimentos. La teoría del funcional de la densidad es un método de cálculo cuántico que calcula la energía a partir del comportamiento de los electrones.

Traslademos la pugna entre entalpía y entropía a un ejemplo cotidiano. Supongamos que en un aula se sientan mezclados chicos y chicas. La fuerza de los grupos de amigos por juntarse entre sí se parece a la entalpía. La fuerza que busca mezclarse libremente de asiento se parece a la entropía. En los momentos tranquilos, los grupos se unen. En los momentos bulliciosos, se mezclan libremente. Esto equivale a cómo, a mayor temperatura, más fuerte es la fuerza que impulsa a mezclarse.

La inteligencia artificial acelera este proceso. Indicadores simples como omega se utilizan como características de entrada para la inteligencia artificial cuando se cuenta con grandes cantidades de datos. La inteligencia artificial examina estas características y selecciona rápidamente qué composiciones permanecerán en una sola fase. Únicamente los candidatos así preseleccionados se comprueban mediante CALPHAD y experimentos. Esta colaboración es la clave para abordar un amplio espacio de composiciones.

Investigaciones recientes han elevado notablemente la precisión de esta predicción. Al introducir múltiples características físicas de manera conjunta, la precisión en la predicción de fases sube de aproximadamente el 70 por ciento a superar el 90 por ciento. Con este nivel, es posible cribar con rapidez más de un millón de composiciones. En la práctica, la inteligencia artificial reduce enormemente el esfuerzo humano. El tiempo y los recursos se concentran en los pocos candidatos así seleccionados.

4.2.2 Estimación de la estructura cristalina mediante la concentración de electrones de valencia

Esta subsección aborda el indicador para estimar qué estructura cristalina adoptará una aleación. Dicho indicador es la concentración de electrones de valencia (Valence Electron Concentration, VEC), que significa el número promedio de electrones externos. Primero, definamos qué es una estructura cristalina. Los átomos metálicos no se disponen al azar, sino que se apilan en una red regular. Es como apilar ladrillos siguiendo una regla. Las propiedades varían según el modo de apilamiento.

Existen dos estructuras importantes en los materiales para cohetes. Una es la cúbica centrada en el cuerpo (Body-Centered Cubic, BCC), una disposición con átomos en los ocho vértices y en el centro de un cubo. La otra es la cúbica centrada en las caras (Face-Centered Cubic, FCC), una disposición con átomos en los ocho vértices y en el centro de las seis caras.

Estas dos estructuras tienen propiedades distintas. La cúbica centrada en el cuerpo suele ser ventajosa para la resistencia a altas temperaturas. En cambio, tiende a presentar fragilidad a temperatura ambiente, fracturándose con facilidad. La cúbica centrada en las caras suele mostrar una buena ductilidad para estirarse bien. Si se conoce de antemano qué estructura se formará, el diseño resulta más sencillo.

La concentración de electrones de valencia es el indicador para estimar esta estructura. Los electrones de valencia son los electrones de la capa externa del átomo que participan en el enlace. El valor promedio del número de electrones externos aportados por cada elemento según la proporción de la composición es la concentración de electrones de valencia. Este valor señala la orientación de la estructura cristalina.

La regla empírica es la siguiente: si este valor es inferior a aproximadamente 6,87, la estructura cúbica centrada en el cuerpo es estable. Si es superior a 8, la estructura cúbica centrada en las caras es estable. En el intervalo intermedio, ambas estructuras se mezclan. Las aleaciones refractarias de alta entropía tienen un valor bajo, por lo que predomina la cúbica centrada en el cuerpo. La conocida aleación de Cantor presenta un valor alto, por lo que es cúbica centrada en las caras.

La utilidad de esta regla se percibe también en investigaciones recientes de inteligencia artificial. Existe un preimpreso publicado a principios de 2026. Un preimpreso es un estudio preliminar que no ha pasado por revisión por pares. Dicho estudio predijo el punto de fusión de aleaciones refractarias de alta entropía. No introdujo directamente la estructura cristalina, sino únicamente características obtenidas de la composición de los elementos. Características con claro sentido físico, como la concentración de electrones de valencia, contribuyeron en gran medida a la predicción. De este modo, el conocimiento físico acumulado por el ser humano y las características elegidas por la inteligencia artificial se dan la mano.

Este caso demuestra dos cosas: primero, que los indicadores simples siguen siendo útiles en la actualidad; segundo, que la inteligencia artificial puede aprender sin entrar en contradicción con la física. Estas dos ideas atraviesan la totalidad de este capítulo.

Profundicemos aquí un poco más en por qué la estructura condiciona las propiedades. La deformación se produce cuando las dislocaciones se deslizan dentro de la red. La estructura cúbica centrada en las caras posee planos donde los átomos están estrechamente empaquetados, lo que permite que las dislocaciones se desplacen con suavidad. Por ello, se deforma bien sin romperse y no se fractura con facilidad. La estructura cúbica centrada en el cuerpo carece de tales planos compactos. Esto no ocurre porque haya pocas trayectorias de deslizamiento, sino porque la barrera de energía que debe superar la dislocación para avanzar es grande. Esta barrera es muy sensible a la temperatura. Por esta razón, la cúbica centrada en el cuerpo es frágil a temperatura ambiente, pero al subir la temperatura las dislocaciones se mueven con facilidad manteniendo la resistencia.

Esta diferencia se traslada directamente a la elección de materiales para cohetes. En el revestimiento de la tobera, la resistencia a altas temperaturas es lo más importante. Por ello, las aleaciones refractarias de alta entropía de la familia cúbica centrada en el cuerpo son candidatas idóneas. Sin embargo, acarrean la debilidad de la fragilidad a temperatura ambiente. Para reducir esta flaqueza, es necesario ajustar minuciosamente la composición. La concentración de electrones de valencia sirve como la brújula inicial de dicho ajuste.

No obstante, la regla de la concentración de electrones de valencia también tiene límites. Los umbrales de 6,87 y 8 son valores empíricos procedentes de un grupo concreto de aleaciones de metales de transición. No se aplican directamente a otros grupos de aleaciones. Al ser un promedio, este valor tampoco refleja el entorno local. Incluso con el mismo valor medio, la estructura real puede variar según la combinación de elementos. Además, esta regla no puede indicar cuándo se formarán compuestos intermetálicos duros y quebradizos. Las fases de Laves o la fase sigma son ejemplos de ello. Si estas fases se mezclan, la aleación se debilita. Por tanto, este indicador también se utiliza para el cribado inicial y se confirma con cálculos precisos. La inteligencia artificial permite realizar este cribado de manera mucho más rápida y amplia.

4.3 Extracción de características de materiales a partir de textos científicos

La inteligencia artificial no puede ver los metales con los ojos. Solo conoce los materiales a través de características convertidas en números. Estas características se denominan descriptores (Descriptor). Esta sección examina cómo extraer estas características a partir de textos de artículos científicos. Primero, destaquemos cuán amplio es el espacio de composiciones. Si se mezclan cinco o más elementos y se varían sus proporciones, el número de combinaciones posibles se dispara exponencialmente. Si hay dos ingredientes para el café, las combinaciones son apenas unas pocas. Si hay diez ingredientes, resulta difícil contar el número de combinaciones. Las aleaciones multicomponente se hallan exactamente en esa situación.

Es imposible explorar todo este vasto espacio mediante experimentos. Fabricar y probar una sola aleación cuesta tiempo y dinero. Los cálculos CALPHAD también requieren tiempo para cada composición. Por eso se necesita un guía que reduzca rápidamente el amplio espacio. La inteligencia artificial es ese guía.

Para emplear la inteligencia artificial, es necesario transformar los materiales en números. Los descriptores son esos números. El radio atómico, la electronegatividad, el punto de fusión y el número de electrones externos son descriptores representativos. La electronegatividad es la fuerza con la que un átomo atrae electrones. Estos valores se calculan según la composición y se introducen en la inteligencia artificial.

El problema reside en que estos números por sí solos no bastan. Incluso con la misma composición, las propiedades varían drásticamente según cómo se haya fabricado el material. La temperatura de tratamiento térmico, el tiempo de mantenimiento y el método de enfriamiento modifican la microestructura. Esta información del proceso es difícil de plasmar en un solo número. Sin embargo, en los artículos científicos, dicha información está descrita en forma de texto.

Los descriptores derivados del lenguaje natural (Natural language-derived descriptor) salvan esta brecha. Consisten en características que la inteligencia artificial extrae tras leer las oraciones de los artículos científicos y convertirlas en valores numéricos. Es el equivalente a trasladar el texto que leen los humanos a números que interpretan las máquinas. Investigaciones recientes han incorporado incluso las descripciones de los procesos a las características mediante este método.

Un preimpreso publicado en 2026 es un buen ejemplo. Se trata de un resultado del equipo de investigación de la Universidad Estatal de Pensilvania en Estados Unidos. Estos investigadores introdujeron en un modelo de lenguaje grande oraciones que describían procesos de tratamiento térmico. A partir de las representaciones numéricas de dichas oraciones, pudieron recuperar casi con total exactitud los valores del proceso original. Y, utilizando esta representación como característica, mejoraron en torno al 20 por ciento la predicción de la dureza de la aleación. Este estudio es un trabajo inicial que aún no ha pasado por revisión por pares.

La minería bibliográfica es la labor de recopilar estas características a gran escala. Extrae automáticamente de innumerables artículos sobre metales valores como temperaturas de transformación de fase, límites de solubilidad sólida, resistencia y ductilidad. Antiguamente, las personas confeccionaban las tablas a mano. Hoy en día, los modelos de lenguaje grande leen las oraciones y rellenan las tablas. Los datos acumulados de este modo constituyen la base para el entrenamiento de la inteligencia artificial.

También existen estudios que han utilizado modelos de lenguaje directamente para la predicción de propiedades. Es el resultado publicado en una revista científica en 2025. Dicho estudio empleó la estructura de modelo de lenguaje llamada Transformer. El Transformer es una inteligencia artificial

que interpreta las relaciones entre las palabras de una oración. Al procesar la composición y las condiciones de la aleación como si fueran texto, predijo sus propiedades. Es un ejemplo de transferencia de una tecnología diseñada para el lenguaje al ámbito de la investigación de materiales.

Recientemente, también han aparecido estudios que construyeron bases de datos de aleaciones de alta entropía mediante modelos de lenguaje grande. Es un resultado publicado en una revista científica en 2026. Esta base de datos organizó de manera automática composiciones y propiedades extraídas de artículos científicos. Aceleró en gran medida un trabajo que a mano habría tomado mucho tiempo. Los datos reunidos de esta forma sirven de base para los modelos predictivos que se examinarán en la siguiente sección.

Este método también tiene sus límites. Las condiciones de medición y las formas de notación varían de un artículo a otro. Si la máquina malinterpreta una frase, se mezclan valores erróneos en los datos. Por eso, los humanos deben revisar nuevamente los datos recopilados. La colaboración en la que la inteligencia artificial recopila y los humanos verifican es el estándar actual.

La calidad de los datos determina la calidad de las predicciones. Si se aprende con datos erróneos, se obtienen respuestas incorrectas. Si el plano está mal, incluso un carpintero hábil no puede construir una casa sólida. Por eso, la etapa de recolección de datos es tan importante como la predicción. Los buenos datos son la base de una buena inteligencia artificial.

Recientemente, la inteligencia artificial también amplía las reglas por sí misma. Un estudio publicado en una revista científica en 2026 presentó un marco de inteligencia artificial enriquecido con conocimiento. Este estudio superó la restricción de tener que mezclar elementos en cantidades iguales. Diseñó aleaciones de alta entropía monofásicas cambiando libremente las proporciones de los elementos según el objetivo. Es el resultado de sumar a la inteligencia artificial el conocimiento metalúrgico acumulado por los humanos.

Esta tendencia muestra la colaboración entre la inteligencia artificial y el conocimiento humano. Si solo se utiliza la inteligencia artificial, esta flaquea fuera de los datos. Si solo se emplean reglas humanas, resulta difícil explorar un espacio amplio. Al combinar ambas, se compensan las debilidades mutuas. El diseño se vuelve sólido cuando la minería de literatura y el enriquecimiento de conocimiento avanzan de la mano.

La integración con CALPHAD también es importante. La inteligencia artificial puede emular rápidamente los cálculos de CALPHAD o seleccionar candidatos para el siguiente cálculo. Al entrelazar los datos recopilados de la literatura, los cálculos de CALPHAD, la teoría del funcional de la densidad (DFT) y los experimentos, el diseño de candidatos se acelera. Este proceso integrado es el método práctico para abordar un amplio espacio de composición.

4.4 Redes neuronales de grafos para representar aleaciones desordenadas

Veamos cómo la inteligencia artificial representa el desorden de las aleaciones de alta entropía. La herramienta clave es la red neuronal de grafos (Graph Neural Network, GNN). Es un método que aprende transformando los materiales en diagramas de nodos y aristas. Primero, repasemos qué es un grafo. Un grafo es un diagrama que representa relaciones mediante nodos y aristas. El mapa del metro es un grafo: las estaciones son nodos y las vías son aristas. Los cristales también se pueden dibujar como grafos: los átomos son nodos y los enlaces entre átomos son aristas.

La red neuronal de grafos lee este diagrama para predecir las propiedades del material. Cada nodo contiene información del átomo. Cada arista contiene la distancia y la dirección entre los átomos. La red neuronal comprende el material intercambiando información entre nodos vecinos. Este intercambio de

información se denomina paso de mensajes (Message passing). Es como intercambiar notas con el compañero de al lado para captar el ambiente de toda la clase.

Este método se adapta bien a monocristales ideales. Un monocristal es un cristal en el que cada sitio está ocupado exactamente por un solo elemento. Las reglas son claras, como en un cristal de sal. Las redes neuronales de grafos cristalinos convencionales se construyeron bajo esta premisa. Colocar un elemento en cada nodo simplifica el cálculo.

El problema radica en que las aleaciones de alta entropía no son así. En estas aleaciones, varios elementos se mezclan probabilísticamente en un mismo sitio. Un sitio determinado puede ser niobio, molibdeno o tantalio. A esto se le llama ocupación fraccionaria (Fractional occupancy). Es un estado en el que un solo sitio alberga múltiples posibilidades al mismo tiempo.

A esto se suma el orden de corto alcance (Short-Range Order, SRO) local. Los átomos no se mezclan de forma completamente caótica. Ciertos elementos prefieren ser vecinos entre sí, mientras que otros se evitan. Esta preferencia a corta distancia influye notablemente en las propiedades. Los modelos convencionales no logran capturar este orden sutil.

Las redes neuronales de grafos de cristales fraccionarios son un intento de resolver este problema. La idea detrás de este enfoque es la siguiente: en lugar de un solo elemento, se asigna un vector de composición a cada nodo. Un vector de composición es una lista numérica que indica la probabilidad de que cada elemento se encuentre en esa posición. Por ejemplo, una posición se representa como 25 por ciento de niobio, 25 por ciento de molibdeno y 50 por ciento de tantalio. De este modo, se incorpora la ocupación fraccionaria de forma natural.

Lo que realmente demostró el estudio preliminar bajo este nombre es limitado y concreto. Dicho estudio predijo la energía del cristal integrando un entorno local de unos 16 átomos junto con la fracción de elementos de toda la aleación. Fue el resultado del entrenamiento con unas mil estructuras. No es un estudio que haya predicho a la vez el orden de corto alcance, el punto de fusión, la entalpía de mezcla y la resistencia a altas temperaturas. Predecir estas propiedades conjuntamente es una dirección por resolver en el futuro, no un logro ya alcanzado. El método de añadir información sobre pares de elementos a las aristas para capturar el orden de corto alcance se encuentra en la misma etapa de propuesta.

Veamos cómo el paso de mensajes conduce al aprendizaje. Al principio, cada nodo solo conoce su propia información. Al intercambiar información una vez, conoce a sus vecinos. Si se repite varias veces, llega a conocer áreas más distantes. Con esta información ampliada, predice las propiedades de todo el material. Si la predicción difiere del valor real, se ajustan los valores de la red neuronal. Al repetir estos ajustes innumerables veces, el modelo se vuelve más inteligente.

Predecir múltiples propiedades de manera conjunta con un solo modelo también es una ventaja. A esto se le llama predicción multiobjetivo. El punto de fusión y la resistencia están relacionados entre sí. La información adquirida al aprender una propiedad ayuda a predecir la otra. Es como estudiar varias materias juntas, donde el aprendizaje de una profundiza la comprensión de las demás.

Sin embargo, esta línea de investigación aún se encuentra en una fase inicial. Las redes neuronales de grafos de cristales fraccionarios se han propuesto recientemente como estudio preliminar. Las celdas grandes, el orden complejo de corto alcance y el historial térmico real siguen siendo problemas difíciles. Por lo tanto, las predicciones de este modelo deben verificarse mediante experimentos y simulaciones atómicas.

Por otra parte, otros estudios de inteligencia artificial que abordan el orden de corto alcance están creciendo con rapidez. Un estudio publicado en una revista científica en 2025 comparó métodos para

reproducir el orden de corto alcance mediante potenciales de aprendizaje automático. Los potenciales de aprendizaje automático son herramientas que calculan las fuerzas interatómicas mediante inteligencia artificial. Este estudio demostró que, aun con una alta precisión energética, es posible no predecir adecuadamente las propiedades. Esto significa que la selección de los datos de entrenamiento determina la predicción de las propiedades.

Ese mismo año, otro estudio aceleró los cálculos de Montecarlo mediante inteligencia artificial. Montecarlo es un método computacional que encuentra estados de equilibrio mediante ensayos aleatorios. Dicho estudio utilizó un modelo subrogado de inteligencia artificial para revelar nanoestructuras a una escala de miles de millones de átomos. En cierto sentido, abrió el camino para observar el orden microscópico dentro de las aleaciones de alta entropía a gran escala. Las redes neuronales de grafos de cristales fraccionarios también evolucionan sobre esta tendencia.

Veamos por qué esta representación del desorden es importante para los materiales de cohetes. La resistencia del revestimiento de la tobera depende de la eficacia con la que se bloquean las dislocaciones. El orden de corto alcance altera la dificultad del camino por el que se desplazan las dislocaciones. Cuando ciertos elementos se agrupan, bloquean las dislocaciones con mayor fuerza. Por lo tanto, incluso con la misma composición, la resistencia varía según el orden de corto alcance.

Este orden microscópico cambia según el método de fabricación. Si se enfría rápidamente, el desorden queda congelado. Si se enfría lentamente, los átomos se acomodan y surge el orden. La temperatura y el tiempo de tratamiento térmico también modifican el orden. Por ello, es necesario considerar no solo la composición, sino también el proceso de fabricación para predecir las propiedades con precisión.

La inteligencia artificial se utiliza para aprender estas relaciones complejas. Conecta a la vez la composición, el proceso, el orden y las propiedades. Para los humanos resulta difícil manejar tantas variables simultáneamente. La inteligencia artificial descubre esas conexiones si dispone de suficientes datos. Las redes neuronales de grafos de cristales fraccionarios representan un intento de capturar estas conexiones a nivel de disposición atómica.

Esta dirección se encuentra todavía en una fase temprana. En la actualidad, tanto la precisión de las predicciones como el volumen de datos son insuficientes. A medida que se acumulen datos experimentales y los cálculos se aceleren, las predicciones mejorarán. Hasta entonces, el procedimiento de verificar las predicciones de la inteligencia artificial mediante experimentos es imprescindible.

4.5 Dinámica molecular a gran escala que observa átomo por átomo

El método para simular el movimiento de cada átomo en una computadora se denomina dinámica molecular (Molecular Dynamics, MD). Es una herramienta para observar el proceso de fractura de los metales a escala atómica. La clave para escalar este cálculo a dimensiones masivas es una inteligencia artificial llamada potencial de neuroevolución (Neuroevolution Potential, NEP). Primero, examinemos por qué se necesita una gran escala. Las grietas en los metales comienzan a partir de unos pocos átomos. Esa pequeña fisura crece, se propaga hacia los límites de grano y finalmente fractura la estructura. Este proceso solo se comprende observando simultáneamente lo diminuto y lo macroscópico. A estos fenómenos se les llama fenómenos multiescala (Multi-scale).

La teoría del funcional de la densidad convencional difícilmente puede realizar esta labor. Este cálculo es preciso, pero computacionalmente muy costoso. Rara vez supera unos cientos de átomos. En cambio, la microestructura real por la que se propaga una grieta contiene más de miles de millones de átomos. Por lo tanto, no es posible ver el panorama completo únicamente con cálculos cuánticos precisos.

Los potenciales empíricos tienen el problema opuesto. Un potencial es una regla para calcular las fuerzas que actúan entre átomos. Los potenciales empíricos son ligeros y rápidos. Sin embargo, su precisión disminuye en aleaciones complejas con más de cinco elementos. El entorno vecino difiere según el elemento y no logran capturar esa sutileza.

Por ello, se necesita una herramienta que combine solo las ventajas de ambos métodos: una herramienta que reúna la precisión de los cálculos cuánticos y la rapidez de las reglas empíricas. Ese espacio lo ocupa el potencial interatómico de aprendizaje automático. La inteligencia artificial aprende a partir de datos de cálculos cuánticos para predecir las fuerzas. El potencial de neuroevolución es una herramienta destacada por su velocidad dentro de esta categoría.

Repasemos también qué es la dinámica molecular. Este cálculo aplica las leyes del movimiento de Newton a todos los átomos. Calcula la fuerza que actúa sobre cada átomo y lo desplaza durante un intervalo de tiempo sumamente corto. Este proceso se repite innumerables veces. Así, es posible observar cómo se mueven los átomos a lo largo del tiempo, tal como en una película.

4.5.1 Potencial de neuroevolución: aunando precisión y velocidad

El potencial de neuroevolución es una herramienta que combina dos ideas. Una es la red neuronal. Una red neuronal es una inteligencia artificial que aprende reglas a partir de datos. La otra es la estrategia evolutiva. La estrategia evolutiva es un método que genera múltiples candidatos, selecciona los mejores y los perfecciona. Es un método de aprendizaje que imita la selección natural.

Este potencial convierte primero el entorno que rodea al átomo en números. Convierte en características la distancia a la que se encuentra un átomo de sus vecinos y el ángulo en el que se ubican. Estas características se introducen en la red neuronal para predecir la energía de un solo átomo. Al sumar la energía de todos los átomos, se obtiene la energía total. Y a partir de la variación de energía, se obtienen las fuerzas que actúan sobre los átomos.

Veamos por qué se analizan conjuntamente la distancia y el ángulo. Si solo se observa la distancia, únicamente se sabe cuán cerca están los átomos. Si también se considera el ángulo, se conoce la configuración geométrica en la que están dispuestos los átomos. Aun a la misma distancia, una disposición triangular y una disposición lineal tienen propiedades diferentes. Por eso, es indispensable incluir tanto la distancia como el ángulo para representar el entorno con fidelidad.

Este enfoque conecta con la esencia del enlace químico. Los átomos se enlazan de manera diferente según su entorno circundante. Si los vecinos están densamente empaquetados, el enlace es fuerte; si están dispersos, es débil. El potencial de neuroevolución aprende esta relación a partir de los datos. Con la relación aprendida, estima rápidamente la energía de una nueva configuración.

La ventaja de este método radica en lograr simultáneamente precisión y velocidad. La red neuronal aprende una precisión cercana a la de los cálculos cuánticos. El cálculo en sí es tan ligero como el de las reglas empíricas. Así, se obtienen fuerzas precisas de manera rápida. Esta es la clave que abre la puerta a los cálculos a gran escala.

El secreto de la velocidad reside en la unidad de procesamiento gráfico (Graphics Processing Unit, GPU). Este dispositivo se creó originalmente para renderizar gráficos de videojuegos. Sobresale en pintar innumerables puntos de la pantalla de forma simultánea. Esta capacidad de procesamiento concurrente se adapta a la perfección al cálculo atómico, ya que permite calcular las fuerzas de innumerables átomos a la vez.

El potencial de neuroevolución está diseñado para ejecutarse eficientemente en estos dispositivos. Se utiliza junto con un programa para unidades de procesamiento gráfico llamado GPUMD. Esta

combinación se ha empleado ampliamente en investigaciones sobre conducción térmica, defectos y deformación a altas temperaturas. Recientemente, también han surgido estudios posteriores que aceleran aún más el método de aprendizaje. Estas mejoras hacen posibles cálculos de escala todavía mayor.

Repasemos también el proceso de construcción de este potencial. Primero, se calculan las energías y fuerzas de diversas configuraciones atómicas mediante la teoría del funcional de la densidad (DFT). Estos valores se convierten en el libro de texto de la inteligencia artificial. La inteligencia artificial consulta este libro de texto para perfeccionar sus predicciones. La estrategia evolutiva selecciona valores en la dirección que mejor se ajusta a los datos. Al concluir este aprendizaje, se completa una calculadora de fuerzas rápida y precisa.

En el entrenamiento se utiliza de forma conjunta el aprendizaje activo. La inteligencia artificial detecta por sí misma aquellas configuraciones en las que tiene poca certeza. Selecciona únicamente esas configuraciones y las calcula de nuevo mediante la teoría del funcional de la densidad. Los datos obtenidos de este modo se reincorporan al entrenamiento. Es un método inteligente para cubrir un amplio rango con un costo computacional reducido.

Sin embargo, una respuesta rápida no siempre es una respuesta correcta. Los potenciales de inteligencia artificial son precisos dentro del rango en el que han sido entrenados. Pueden fallar en composiciones o temperaturas que no estaban en los datos de entrenamiento. En las aleaciones de alta entropía, el entorno cambia drásticamente con solo variar ligeramente la combinación de elementos. Por ello, la representatividad de los datos de entrenamiento es más importante que cualquier otra cosa. Se requiere un proceso separado para verificar la incertidumbre de las predicciones.

4.5.2 Deformación y fractura observadas mediante cálculos a gran escala

Un billón es un número difícil de imaginar. Al alinear un billón de átomos, se alcanza un tamaño cercano al de una microestructura real. Esto significa que es posible representar una probeta a escala micrométrica hasta el nivel de cada átomo individual. En cálculos pequeños, solo se observan una sola grieta y unos pocos límites de grano. En cálculos a gran escala, se puede apreciar cómo interactúan entre sí numerosos límites de grano y defectos.

Un estudio preliminar publicado en 2026 superó este umbral. Esta investigación calculó 1,6 billones de átomos mediante potenciales neuroevolucionarios. Se utilizaron conjuntamente cerca de 45 000 dispositivos de cálculo en una nueva supercomputadora de China. Lo que demostró este estudio fue la escala y la velocidad de cómputo. Informa haber ejecutado el mismo cálculo con mucha mayor rapidez que el récord anterior. No se trata de un estudio que haya producido resultados que analicen de forma efectiva las dislocaciones y grietas de aleaciones de alta entropía a esta escala. Una demostración que abre la escala y lo que se revela con esa escala son cuestiones distintas. Este estudio es aún una investigación preliminar que no ha pasado por la revisión por pares.

Comparemos brevemente para entender el significado de esta cifra. En el pasado, calcular un millón de átomos ya era una tarea colosal. Hace unos años se superaron los 100 millones y pronto se alcanzaron los 10 000 millones. Ahora se ha sobrepasado el billón. El tamaño que se puede calcular está creciendo con rapidez. Sin embargo, el cálculo de un billón de átomos es una demostración simbólica que exprimió al límite una supercomputadora colosal. En el diseño real de aleaciones, los cálculos que se utilizan a diario tienen una escala de entre miles y cientos de miles de átomos. Los potenciales de inteligencia artificial y las computadoras veloces impulsan este crecimiento.

La ventaja que ofrece una gran escala reside en la estadística. Cuantos más átomos haya, mejor se capturan los sucesos poco frecuentes. Las semillas de las grietas se forman en raras ocasiones. En

cálculos pequeños, es posible no ver esas semillas. En cálculos grandes, se observan múltiples semillas a la vez. Por tanto, las predicciones se acercan mucho más a la realidad.

Veamos primero por qué este cálculo resulta útil para los materiales de cohetes. La estructura real del revestimiento de la tobera contiene una cantidad incontable de átomos. Es necesario saber qué ocurre en esta gran estructura para conocer su vida útil. Los cálculos a gran escala proporcionan una imagen cercana a esta estructura real. Por ello, pueden utilizarse para estimar qué composición resistirá durante más tiempo.

En principio, hay tres fenómenos que pueden observarse mediante esta dinámica molecular a gran escala. Lo siguiente describe los objetivos que persigue este método y no resultados que la demostración anterior ya haya esclarecido. El primero es la nucleación de dislocaciones. Una dislocación es un pequeño defecto donde se inicia la deformación. Se rastrea en tiempo real cómo surgen y cómo se deslizan las dislocaciones en una red cúbica centrada en el cuerpo. En las aleaciones de alta entropía, las fluctuaciones de la composición local atrapan las dislocaciones. Este atrapamiento constituye el principio que incrementa la resistencia.

El segundo es el deslizamiento de límites de grano. Los límites de grano son las fronteras entre los granos cristalinos. Cuando se someten a fuerzas a altas temperaturas, estos límites se deslizan. Los cálculos a gran escala miden a nivel atómico este deslizamiento y la aglomeración de huecos. Este comportamiento determina la vida útil frente a la fluencia.

El tercero es la propagación de grietas. La punta de la grieta es el frente extremo de la fisura. En este punto, la disposición atómica cambia y se produce la deformación. Los cálculos a gran escala observan cómo crecen y cómo se detienen las grietas. Esto aporta pistas para estimar qué composiciones resisten mejor el agrietamiento.

Estos tres fenómenos están interconectados. Las dislocaciones se acumulan y convergen en los límites de grano. Los límites de grano se deslizan y generan huecos. Esos huecos se unen y se transforman en grietas. En cálculos pequeños resulta difícil observar este encadenamiento. En cálculos a gran escala es posible abarcarlo de principio a fin de un solo vistazo.

Veamos por qué resulta útil observar este encadenamiento. La vida útil de un material se decide en su eslabón más débil. Una cadena se rompe por el eslabón más frágil. Si se conoce qué etapa se desmorona primero, se puede reforzar ese punto. Los cálculos a gran escala ayudan a localizar de antemano dónde se encuentra este eslabón débil.

Esta información es útil para la investigación de revestimientos de toberas de cohetes. El revestimiento de la tobera recibe tanto gases a alta temperatura como choque térmico. Estimar qué composición resistirá durante más tiempo antes de hacer experimentos ahorra tiempo. Los cálculos a gran escala actúan como una balanza que permite reducir los candidatos.

Sin embargo, este cálculo también está sujeto a condiciones. La precisión del resultado depende de la calidad del potencial de inteligencia artificial. En situaciones que el potencial no ha aprendido, el cálculo también arroja errores. Por ello, los resultados de los cálculos a gran escala deben verificarse mediante experimentos y cálculos precisos a pequeña escala. Que la escala sea enorme y que el resultado sea correcto son cuestiones distintas.

4.6 Inteligencia artificial para el diseño de metales coherente con las leyes de la física

El aprendizaje basado únicamente en datos puede generar predicciones peligrosas. El método que previene ese peligro mediante las leyes de la física se denomina red neuronal informada por la física (Physics-Informed Neural Network, PINN). Veamos cómo se aplica este método a la metalurgia, es decir, la ciencia que estudia los metales. Primero examinemos la debilidad del aprendizaje exclusivo a partir de datos. La inteligencia artificial responde bien dentro de los datos aprendidos. Sin embargo, en zonas donde no hay datos produce respuestas absurdas. A estas áreas se las denomina fuera del rango de entrenamiento (Out-of-Distribution). Se parece a un estudiante que se enfrenta a una pregunta fuera del temario del examen.

En el diseño de metales, estas respuestas absurdas violan las leyes de la física. Por ejemplo, la temperatura sube pero el valor de estabilidad se mueve en sentido contrario. También se llega a predecir que la energía de red disminuye a medida que el volumen aumenta. Tales respuestas no pueden darse en la naturaleza. Si se confía ciegamente en estas predicciones, el diseño falla.

Las redes neuronales informadas por la física previenen este problema en la etapa de entrenamiento. El método consiste en incorporar las leyes de la física a la función de pérdida. La función de pérdida es el patrón que mide cuánto se equivoca la inteligencia artificial. Por lo general, solo mide la diferencia entre la predicción y los datos reales. A esto se le suma el grado de infracción de las leyes de la física.

Al hacer esto, la inteligencia artificial intenta cumplir dos cosas a la vez. Una es ajustarse a los datos. La otra es respetar las leyes de la física. Incluso en regiones donde no hay datos, las leyes orientan la dirección de la respuesta. Es como un estudiante que intenta resolver un problema fuera del temario aplicando principios fundamentales.

¿Qué leyes se incorporan en el diseño de metales? Un ejemplo es cómo debe variar la energía libre según la temperatura. También se introduce la relación en la que la resistencia se define como la suma de diversas contribuciones de refuerzo. Se añaden además las condiciones que debe cumplir el equilibrio de fases. Estas leyes reducen la libertad de la inteligencia artificial. Esa libertad reducida produce, por el contrario, respuestas más fiables.

Expliquemos con más detalle el modo de incluir leyes en la función de pérdida. Las leyes de la física suelen formularse como ecuaciones. Si se introduce el valor predicho en esa ecuación, ambos lados deben coincidir. Si no coinciden, queda una diferencia. A esta diferencia se le llama residuo. Es un método que penaliza si el residuo es grande.

De esta manera, la inteligencia artificial se esfuerza por evitar las penalizaciones. Intenta ajustarse a los datos al tiempo que cumple las ecuaciones. Se parece a un estudiante que, al entregar las respuestas del examen, cuadra incluso las fórmulas de cálculo y las unidades. Es más fuerte ante problemas aplicados que un estudiante que solo memorizó la respuesta. He aquí por qué las redes neuronales informadas por la física son sólidas fuera del rango de datos.

El valor de este método aumenta cuando los datos son escasos. En los materiales para condiciones extremas de cohetes, los datos experimentales son valiosos y raros. Cada prueba cuesta mucho dinero y tiempo. Si los datos son pocos, el aprendizaje puro se desestabiliza fácilmente. Las leyes de la física suplen los datos insuficientes.

Es preciso distinguir estos dos métodos. El estudio de predicción del punto de fusión visto en la sección anterior no es una red neuronal informada por la física. Dicha investigación es un modelo de regresión que seleccionó e incorporó descriptores físicamente significativos. No introdujo de forma directa ecuaciones de la física en la función de pérdida. Una regresión que utiliza buenos descriptores y un

aprendizaje que inserta leyes físicas en la función de pérdida son métodos distintos. Sin embargo, ambos persiguen el mismo objetivo: lograr que las predicciones de la inteligencia artificial no contradigan la metalurgia. Cuando los buenos descriptores y las restricciones físicas van de la mano, nos acercamos más a esa meta.

4.6.1 Integración de herramientas para el diseño de revestimientos de toberas para temperaturas ultraaltas

Esta subsección examina cómo se utilizan conjuntamente las herramientas anteriores en el diseño real. El objetivo es una aleación refractaria de alta entropía para revestimientos de toberas de temperaturas ultraaltas. Se pone mayor peso en los procedimientos de diseño y verificación que en un único resultado terminado. También se analiza hasta dónde han llegado las investigaciones recientes y qué desafíos quedan pendientes. Una buena aleación para el revestimiento de toberas debe satisfacer múltiples propiedades a la vez. Debe tener un punto de fusión elevado. Si la densidad es excesiva, el cohete se vuelve pesado. Debe mantener su resistencia a altas temperaturas. Su superficie debe resistir bien al entrar en contacto con el oxígeno. El revestimiento y la dilatación térmica deben acoplarse adecuadamente. Si solo una propiedad es buena, no puede utilizarse como componente.

Alcanzar todos estos objetivos simultáneamente resulta abrumador para un ser humano. Si se añade mucho wolframio, el punto de fusión sube. Sin embargo, la densidad aumenta y puede volverse quebradizo. Si se añade aluminio o cromo, se vuelve resistente a la oxidación. Pero otras propiedades pueden verse comprometidas. A estos objetivos entrelazados se les denomina optimización multiobjetivo.

En los cohetes, la densidad es un asunto sumamente delicado. Si un componente es pesado, todo el cohete se vuelve pesado. Un cohete pesado viaja menos distancia con la misma cantidad de combustible. Por tanto, aunque su resistencia sea excelente, si es demasiado pesado no se puede utilizar. Hay que sopesar al mismo tiempo el punto de fusión, la resistencia y el peso.

En este balanceo no existe una única respuesta correcta. Si se cede un poco de resistencia, se puede reducir el peso. Si se cede un poco de peso, se puede aumentar la resistencia. Así surge un conjunto de candidatos con compensaciones mutuas. Frente a este conjunto, el ser humano elige la opción que mejor se ajuste a su propósito. La inteligencia artificial traza este conjunto con gran rapidez.

La inteligencia artificial aborda conjuntamente estos objetivos entrelazados. Convierte las composiciones en números mediante los descriptores vistos anteriormente. Filtra con rapidez el punto de fusión y la resistencia mediante modelos de predicción. Descarta candidatos inverosímiles aplicando restricciones físicas. Examina la deformación y el agrietamiento mediante cálculos atómicos a gran escala. Este flujo de trabajo reduce drásticamente los candidatos.

Estudios recientes demuestran la utilidad de este flujo. Una investigación publicada en una revista académica en 2024 diseñó aleaciones refractarias de alta entropía para temperaturas ultraaltas mediante aprendizaje automático. Se redujo el amplio espacio de composición con inteligencia artificial y se seleccionaron candidatos prometedores. Fue un intento de mejorar simultáneamente la resistencia y la ductilidad. Tales estudios demuestran que el diseño mediante inteligencia artificial funciona en la etapa de laboratorio.

La inteligencia artificial también interviene en el problema de la oxidación. Existe un estudio preliminar publicado a finales de 2025. Dicho estudio exploró aleaciones concentradas complejas refractarias resistentes a la oxidación a alta temperatura mediante aprendizaje activo. El aprendizaje activo es un método en el que la inteligencia artificial elige por sí misma la siguiente composición que se

va a probar. Es una vía para encontrar buenos candidatos con pocos experimentos. Esta investigación también es un estudio preliminar que aún no ha pasado por la revisión por pares.

Sin embargo, es prematuro hablar de ello como un logro definitivo. Las cifras que afirman que cierta composición alcanza con exactitud determinado nivel de resistencia a una temperatura específica requieren verificación en el texto original. En este libro no se incluyen valores de rendimiento no verificados. En su lugar, esta sección hace hincapié en el procedimiento de validación por el que debe pasar el diseño.

La verificación consta de tres etapas. Primero, se comprueba si la composición candidata es estable en los cálculos. Se confirma el equilibrio de fases mediante CALPHAD y la teoría del funcional de la densidad (DFT). A continuación, se comprueba si la aleación puede fabricarse en la realidad. De nada sirve que sea excelente en el cálculo si no se puede fabricar. Por último, se miden la resistencia y la resistencia a la oxidación a altas temperaturas. Las pruebas en condiciones cercanas al entorno real representan la barrera final.

No se debe olvidar que, al atravesar estas etapas, los errores se propagan. Un pequeño error en los cálculos atómicos se traslada a la predicción de la microestructura y luego se amplifica en la predicción de la resistencia del componente. Si la superficie se desgasta por la oxidación y los límites se desplazan, la predicción se vuelve aún más inestable. A esto se suman las desviaciones que surgen en cada proceso de fabricación. Por eso, un buen diseño ofrece un rango de vida útil y no un único valor puntual. Al final, debe someterse a pruebas de certificación en condiciones idénticas a las del componente real.

Estas tres etapas no transcurren únicamente de forma secuencial. Si en las pruebas se detecta una debilidad, se vuelve a modificar la composición. La composición corregida se calcula de nuevo y se vuelve a fabricar. Este proceso cíclico se repite continuamente. A este ciclo se le llama diseño en bucle cerrado. El laboratorio autónomo es un intento de ejecutar este ciclo de manera automatizada.

La inteligencia artificial interviene en cada etapa de este ciclo. Elige la siguiente composición para probar, aprende de los resultados y vuelve a seleccionar. El ciclo sigue girando sin necesidad de que una persona haga guardia toda la noche. Con ello se acorta el camino para alcanzar una buena composición. Esta tendencia está transformando el futuro del desarrollo de aleaciones para altas temperaturas.

En el campo de las superaleaciones de base níquel, el diseño mediante inteligencia artificial también es muy activo. Diversos estudios de 2025 predijeron la vida útil frente a la fluencia de superaleaciones monocristalinas mediante aprendizaje automático. Introdujeron información sobre la composición y la microestructura para seleccionar composiciones más duraderas. La fracción volumétrica y la morfología de los precipitados gamma prima (γ') se utilizaron como descriptores clave. Tales investigaciones demuestran que la inteligencia artificial aporta una nueva utilidad incluso a metales tradicionales.

También existen investigaciones que combinan cálculos de campo de fases con la inteligencia artificial. El campo de fases es un método computacional que describe la evolución de la microestructura como un campo continuo. Un estudio de 2025 enlazó estos cálculos con el aprendizaje automático. Predijo de forma conjunta el impacto que la composición y la microestructura ejercen sobre la deformación por fluencia. Este enfoque que conecta la computación y la inteligencia artificial es cada vez más frecuente.

Estos casos ofrecen una lección común. La inteligencia artificial despliega su fuerza cuando cuenta con buenos descriptores y datos suficientes. Si los descriptores están vinculados a la física, las predicciones son robustas. Si los datos están sesgados hacia una región específica, el modelo se desestabiliza fuera de ella. Por ello, recopilar datos con amplitud y seleccionar adecuadamente los descriptores es lo prioritario.

En este procedimiento, el lugar de la inteligencia artificial es claro. La inteligencia artificial es un guía que reduce el espacio de composición. Si la inteligencia artificial reduce 100 candidatos a 3, los humanos prueban en profundidad solo esos 3. Al dividir los roles de este modo, el desarrollo se acelera y los fracasos disminuyen. Esta es la verdadera imagen de una inteligencia artificial para el diseño de metales coherente con las leyes de la física.

Resumen

Este capítulo ha tratado sobre los metales destinados a proteger los puntos donde la temperatura alcanza su máximo en un cohete. Esos puntos son la tobera de la cámara de combustión y las superficies de reentrada. En estos lugares, el calor, la presión y las fuerzas mecánicas castigan conjuntamente al material.

Las superaleaciones a base de níquel utilizadas durante mucho tiempo alcanzan su límite alrededor de los 1.150 grados. El nuevo candidato que apunta a temperaturas superiores es la aleación refractaria de alta entropía. Esta aleación se obtiene mezclando varios metales de alto punto de fusión en cantidades similares.

Las investigaciones iniciales explicaron las propiedades de esta aleación mediante cuatro efectos: la entropía configuracional, la distorsión de la red, la difusión lenta y el efecto cóctel. Algunos de ellos fueron refutados posteriormente, por lo que resulta difícil considerarlos leyes generales. Aun así, este marco sirve como punto de partida para evaluar qué composiciones permanecerán estables en una sola fase. El parámetro omega y la concentración de electrones de valencia son indicadores sencillos que ayudan en esa evaluación.

Para abordar un amplio espacio de composiciones se necesita inteligencia artificial. La inteligencia artificial interpreta los materiales mediante números llamados descriptores. Los modelos de lenguaje grandes extraen estas características de los textos de los artículos. Las redes neuronales de grafos aprenden los materiales como puntos y líneas. Las redes neuronales de grafos de fracciones cristalinas son un nuevo intento de capturar el desorden de las aleaciones de alta entropía.

El comportamiento a nivel atómico se observa mediante dinámica molecular. La inteligencia artificial llamada potencial neuroevolutivo acelera estos cálculos. Investigaciones recientes han alcanzado una escala de más de 1 billón de átomos. Estos cálculos masivos permiten observar dislocaciones, deslizamiento de límites de grano y grietas a nivel atómico.

La idea que une todas estas herramientas es la consistencia física. El aprendizaje basado puramente en datos puede generar respuestas que violan la física. Las redes neuronales informadas por la física previenen esto incorporando leyes a la función de pérdida. La inteligencia artificial es una guía que reduce los candidatos, y el juicio final lo determinan los experimentos.

La mayoría de los casos de este capítulo son aún investigaciones preliminares. Las cifras de los estudios previos no son valores definitivos. Los casos de empresas son solo ejemplos industriales y no demostraciones académicas. Por eso, un buen diseño siempre pasa por la verificación conjunta del cálculo y el experimento. La colaboración donde la inteligencia artificial acota el camino y las personas lo deciden representa el presente de este campo.

Si se resume este capítulo en una sola frase, es la siguiente. La inteligencia artificial cambia la velocidad del diseño de metales para altas temperaturas. En el pasado, se tardaba mucho tiempo en probar un solo candidato. Ahora, la inteligencia artificial filtra de antemano numerosos candidatos. Solo los pocos restantes se comprueban mediante cálculos precisos y ensayos.

Los desafíos futuros también son claros. Es necesario recopilar datos experimentales de manera más amplia y uniforme. Se debe informar conjuntamente sobre la incertidumbre en las predicciones de la inteligencia artificial. Hay que incorporar las leyes físicas de manera más exhaustiva en el aprendizaje. Cuanto mejor se disponga de estos tres elementos, más rápido será el diseño de materiales para cohetes.

Fuentes y referencias relacionadas

Yeh, J. W., Chen, S. K., Lin, S. J., Gan, J. Y., Chin, T. S., Shun, T. T., Tsau, C. H., & Chang, S. Y. (2004). Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. Advanced Engineering Materials, 6(5), 299-303. https://doi.org/10.1002/adem.200300567

Cantor, B., Chang, I. T. H., Knight, P., & Vincent, A. J. B. (2004). Microstructural development in equiatomic multicomponent alloys. Materials Science and Engineering A, 375-377, 213-218. https://doi.org/10.1016/j.msea.2003.10.257

Senkov, O. N., Wilks, G. B., Miracle, D. B., Chuang, C. P., & Liaw, P. K. (2010). Refractory high-entropy alloys. Intermetallics, 18(9), 1758-1765. https://doi.org/10.1016/j.intermet.2010.05.014

Senkov, O. N., Wilks, G. B., Scott, J. M., & Miracle, D. B. (2011). Mechanical properties of Nb25Mo25Ta25W25 and V20Nb20Mo20Ta20W20 refractory high entropy alloys. Intermetallics, 19(5), 698-706. https://doi.org/10.1016/j.intermet.2011.01.004

Zhang, Y., Zuo, T. T., Tang, Z., Gao, M. C., Dahmen, K. A., Liaw, P. K., & Lu, Z. P. (2014). Microstructures and properties of high-entropy alloys. Progress in Materials Science, 61, 1-93. https://doi.org/10.1016/j.pmatsci.2013.10.001

Miracle, D. B., & Senkov, O. N. (2017). A critical review of high entropy alloys and related concepts. Acta Materialia, 122, 448-511. https://doi.org/10.1016/j.actamat.2016.08.081

George, E. P., Raabe, D., & Ritchie, R. O. (2019). High-entropy alloys. Nature Reviews Materials, 4(8), 515-534. https://doi.org/10.1038/s41578-019-0121-4

Rao, Z., Tung, P. Y., Xie, R., Wei, Y., Zhang, H., Ferrari, A., Klaver, T. P. C., Kormann, F., Sukumar, P. T., da Silva, A. K., Chen, Y., Li, Z., Ponge, D., Neugebauer, J., Gutfleisch, O., Bauer, S., & Raabe, D. (2022). Machine learning-enabled high-entropy alloy discovery. Science, 378(6615), 78-85. https://doi.org/10.1126/science.abo4940

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Fan, Z., Zeng, Z., Zhang, C., Wang, Y., Dong, H., Chen, Y., & Ala-Nissila, T. (2021). Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport. Physical Review B, 104, 104309. https://doi.org/10.1103/PhysRevB.104.104309

Fan, Z., Wang, Y., Ying, P., Song, K., Wang, J., Wang, Y., Zeng, Z., Xu, K., Lindgren, E., Rahm, J. M., Gabourie, A., Liu, J., Dong, H., Wu, J., Chen, Y., Zhong, Z., Sun, J., Erhart, P., Su, Y., & Ala-Nissila, T. (2022). GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations. Journal of Chemical Physics, 157, 114801. https://doi.org/10.1063/5.0106617

He, Q., Wu, D., Wang, H., et al. (2024). Machine learning assisted design of refractory high-entropy alloy for ultrahigh temperature applications. Materials & Design, 246, 113326. https://doi.org/10.1016/j.matdes.2024.113326

Divilov, S., et al. (2024). A comparative study of predicting high entropy alloy phase fractions with traditional machine learning and deep neural networks. npj Computational Materials, 10, 68. https://www.nature.com/articles/s41524-024-01335-1

Hasnain, M. (2026). Physically Consistent Machine Learning for Melting Temperature Prediction of Refractory High-Entropy Alloys. arXiv Preimpresión (sin revisión por pares), arXiv:2601.03801. https://arxiv.org/abs/2601.03801

Suo, P., Cao, W., Wu, X., Zhang, W., Fan, Z., Xian, S., et al. (2026). Trillion-atom molecular dynamics simulations with ab initio accuracy. arXiv Preimpresión (sin revisión por pares), arXiv:2604.24816. https://arxiv.org/abs/2604.24816

Hsiao, L. C., Liu, Z. K., & Reinhart, W. (2026). Modeling High Entropy Alloys' Mechanical Property through Natural Language-Derived Descriptors. arXiv Preimpresión (sin revisión por pares), arXiv:2604.21807. https://arxiv.org/abs/2604.21807

Kotama, T., & Huang, Y. (2026). Crystal Fractional Graph Neural Network for Energy Prediction of High-Entropy Alloys. arXiv Preimpresión (sin revisión por pares), arXiv:2605.08103. https://arxiv.org/abs/2605.08103

Bejjipurapu, A., Karumuri, S., Flanagan, J. C., Tucker, V., Bilionis, I., Strachan, A., Sandhage, K. H., & Titus, M. S. (2025). Active Learning Discovery of High Temperature Oxidation Resistant Refractory Complex Concentrated Alloys. arXiv Preimpresión (sin revisión por pares), arXiv:2512.15958. https://arxiv.org/abs/2512.15958

Cao, Y., Sheriff, K., & Freitas, R. (2025). Capturing short-range order in high-entropy alloys with machine learning potentials. npj Computational Materials, 11, 268. https://doi.org/10.1038/s41524-025-01722-2

Liu, X., Yang, K., Liu, Y., Zhou, F., Fan, D., Pei, Z., Xu, P., & Tian, Y. (2025). Revealing nanostructures in high-entropy alloys via machine-learning accelerated scalable Monte Carlo simulation. npj Computational Materials, 11, 267. https://doi.org/10.1038/s41524-025-01762-8

Huang, H., Yu, Y., Deng, W., Qian, Q., & Zheng, H. (2026). Breaking equiatomic constraints: knowledge-enhanced AI framework for function-oriented single-phase high-entropy alloy design. npj Computational Materials, 12, 209. https://doi.org/10.1038/s41524-026-02071-4

Kamnis, S., & Delibasis, K. (2025). High entropy alloy property predictions using a transformer-based language model. Scientific Reports, 15, 11861. https://doi.org/10.1038/s41598-025-95170-z

Chizhevskiy, V., Marković, G., Benrazzouq, S., Cvelbar, U., Nominé, A., & Zavašnik, J. (2026). High Entropy Alloys Database generated with Large Language Model. Scientific Data, 13, 612. https://doi.org/10.1038/s41597-026-06930-z

Yao, J., Yu, W., Wang, J., Zhang, L., Liu, F., Li, W., Tan, L., Huang, L., & Liu, Y. (2025). Integrating Machine Learning and Thermodynamic Descriptors for Enhanced Ni-Based Single Crystal Superalloys Creep Life Prediction and Alloy Design. Metals and Materials International, 31, 2730-2748. https://doi.org/10.1007/s12540-025-01906-x

Song, J., Shan, Y., Zhang, Z., Wang, S., Zhang, H., Muhammad, S., Huang, H., & Li, Y. (2025). Phase-field-informed machine learning on creep behavior of Ni-based single-crystal superalloys. Journal of Materials Informatics, 5, 2. https://doi.org/10.20517/jmi.2025.02

Capítulo 5. Control de la microestructura y evaluación de la vida útil térmica de materiales compuestos (CMC) y cerámicas de ultraalta temperatura (UHTC) para aplicaciones aeroespaciales

Introducción

Una vasija de cerámica no se derrite aunque se introduzca en un horno caliente. Sin embargo, si un recipiente caliente se sumerge directamente en agua fría, se agrieta. Esto se debe a que, aunque es resistente al calor, resulta vulnerable a los cambios bruscos de temperatura y a los impactos. Los materiales que protegen la superficie exterior de cohetes y vehículos hipersónicos también buscan una respuesta entre estas dos propiedades. No obstante, las temperaturas y las fuerzas en esos entornos no se pueden comparar con las de una cocina.

El interior de la cámara de combustión de un motor cohete se llena de gases calientes a miles de grados. La parte delantera de un vehículo hipersónico se calienta por sí misma al chocar contra el aire. Las superficies en estos lugares soportan calor, oxidación, abrasión y grandes fuerzas de manera simultánea. La oxidación es un fenómeno en el que el oxígeno entra en contacto con el material y lo va degradando.

Los metales son resistentes, pero se ablandan a temperaturas muy altas. Es similar a cómo un caramelo blando se dobla en la mano. Por esta razón, los investigadores dirigen su atención a las cerámicas. Al igual que la porcelana, las cerámicas son resistentes al calor y soportan bien las altas temperaturas.

Los dos materiales que se abordan en este capítulo son los materiales compuestos de matriz cerámica (CMC) y las cerámicas de ultraalta temperatura (UHTC). Los materiales compuestos de matriz cerámica se fabrican introduciendo fibras finas dentro de una matriz cerámica. Las cerámicas de ultraalta temperatura son cerámicas con un punto de fusión sumamente alto. Ambos tienen como objetivo mantener su forma en entornos de calor extremo.

El problema radica en que resulta difícil prever con antelación cuándo se dañarán estos materiales. Es complicado recrear con exactitud el entorno real de vuelo en un laboratorio. El coste de cada ensayo también es elevado. Por ello, la inteligencia artificial (AI) surge como una herramienta para aliviar esta dificultad.

La inteligencia artificial se utiliza para interpretar las microestructuras del interior del material y predecir su vida útil. Por microestructura se entiende la forma de los granos, las fibras, los poros y los límites que se observan al microscopio. Este capítulo explica paso a paso cómo convertir estas pequeñas estructuras en números y cómo anticipar el futuro a partir de ellas. Se expone con palabras sencillas, sin emplear fórmulas matemáticas.

En este capítulo existe un hilo conductor principal: convertir el mundo microscópico de los materiales en números. Las fotografías de microscopio, la orientación cristalina y las vibraciones atómicas son objeto de este proceso. Esta información se transforma en un formato que la inteligencia artificial puede interpretar. Así, la inteligencia artificial puede anticipar el futuro y diseñar la estructura.

El orden de este capítulo es el siguiente. En primer lugar, se examina cómo se degradan los materiales en entornos extremos. A continuación, se analiza el método para predecir la vida útil convirtiendo los granos cristalinos en grafos. Seguidamente, se revisa la forma de diseñar inversamente el trenzado de

fibras. Por último, se describe cómo conectar el flujo de calor en la escala microscópica con el flujo de calor en la pieza completa.

Conviene dejar claro un principio de antemano. La inteligencia artificial en este capítulo no sustituye a los ensayos físicos. Los valores de rendimiento calculados por la inteligencia artificial deben verificarse con piezas reales. Los datos no verificados quedan solo como referencia. Este libro incluye en su texto únicamente información comprobada. Este principio constituye la base para garantizar la seguridad en el desarrollo de materiales.

5.1 Entornos extremos de vuelo hipersónico y propulsión de cohetes de nueva generación y mecanismos de degradación de estructuras térmicas

Contenido de esta sección

Esta sección describe primero cuán severo es el entorno que los materiales deben soportar. Se toman como ejemplos los vehículos que vuelan a velocidades hipersónicas y el interior de los motores cohete de alta presión. Se explica por qué los metales convencionales no son suficientes y por qué se necesitan cerámicas.

A continuación, se examina cómo se deterioran en la práctica los materiales cerámicos. Se analizan sucesivamente la oxidación de los materiales compuestos de fibra de carbono, y la ablación y los cambios de fase en las cerámicas de ultraalta temperatura. La ablación es el fenómeno por el cual la superficie se funde o se erosiona debido a los gases calientes. Es necesario comprender estos procesos de daño para entender qué aprende la inteligencia artificial en la sección siguiente al realizar sus predicciones.

5.1.1 Entornos de calentamiento aerodinámico fuera del equilibrio y gases de combustión de alta entalpía

El calentamiento aerodinámico es el fenómeno por el cual el aire calienta la superficie durante el vuelo a alta velocidad. La fricción con la superficie también aporta calor. Sin embargo, a velocidades hipersónicas, la contribución del calentamiento por una fuerte compresión del aire es aún mayor. Cuanto mayor es la velocidad, más rápido aumenta este calor, calentando intensamente la superficie.

La velocidad hipersónica supera cinco veces la velocidad del sonido. Mach 7 equivale a siete veces la velocidad del sonido. A esta velocidad, el aire no fluye de manera ordinaria. Se genera una onda de choque delante del vehículo. La onda de choque es una delgada pared de presión que se forma al comprimirse repentinamente el aire.

El aire que atraviesa la onda de choque se desacelera y se calienta en un instante. Cuando la temperatura se eleva enormemente, las moléculas de aire se disocian. Las moléculas de oxígeno y nitrógeno se separan en átomos individuales. Este oxígeno atómico disociado presenta una reactividad mayor que el oxígeno ordinario y ataca con más fuerza a los materiales.

Las partes de un vehículo que reciben más calor que cualquier otra zona son el borde de ataque y el morro. El borde de ataque es el extremo del ala que entra primero en contacto con el aire. El morro es la punta delantera del vehículo. La superficie en estos puntos alcanza miles de grados. Además, sobre ellos actúa de forma superpuesta una fuerza cortante que empuja la superficie lateralmente.

El interior de un motor cohete resulta igualmente severo. Al quemar metano y oxígeno líquido se generan gases calientes. Estos gases contienen una mezcla de agua, dióxido de carbono y monóxido de carbono. También queda oxígeno residual. La presión en la cámara de combustión alcanza varias decenas de atmósferas. Este gas a alta presión golpea sin cesar las paredes de la cámara de combustión.

Este entorno se denomina entorno de alta entalpía. La entalpía es la cantidad de energía térmica que contiene un gas. Alta entalpía significa que el gas posee una enorme cantidad de calor. Este gas transfiere calor masivamente a las paredes. No obstante, la temperatura de la pared es menor que la del gas, debido a que el calor se disipa hacia el interior del material y también se lleva a cabo una refrigeración.

A esto se añade el factor del tiempo. El cohete realiza la combustión durante cientos de segundos desde el lanzamiento hasta alcanzar la órbita. Los vehículos hipersónicos planean por la atmósfera durante periodos prolongados. Esto significa que reciben calor no solo durante un breve instante, sino a lo largo de un tiempo extenso. Por tanto, no basta con que el material resista solo por un momento.

Hagámonos una idea de cuán caliente es. El punto de estancamiento en el borde de ataque hipersónico roza los límites del material. El punto de estancamiento es el lugar donde el flujo se detiene y el calor se concentra al máximo. La carga térmica en este punto no se puede comparar con la llama de gas de una cocina. La fuerza cortante que empuja la superficie hacia un lado también actúa en este lugar.

Un único material debe soportar todas estas condiciones. El calor, la oxidación, la fuerza cortante y los largos tiempos de exposición se superponen. Una revisión de 2024 que recopila materiales para vuelo hipersónico señala este entorno complejo como el principal desafío en la selección de materiales. Este desafío constituye el punto de partida de la investigación en materiales cerámicos.

La tarea de la inteligencia artificial en este campo es clara. Los ensayos reales son costosos y escasos. Por tanto, a partir de los datos de unos pocos ensayos, es necesario prever una amplia gama de condiciones. La inteligencia artificial es hábil extrayendo patrones a partir de datos reducidos. Sin embargo, sin la ayuda de las leyes físicas, produce respuestas erróneas. Por esta razón, diversos métodos descritos en este capítulo combinan la física con la inteligencia artificial.

5.1.2 Proceso de degradación por oxígeno en compuestos de fibra de carbono (C/SiC)

El carburo de silicio reforzado con fibra de carbono (C/SiC) es un material termoestructural representativo. La fibra de carbono es fina y ligera, y no se rompe con facilidad. El carburo de silicio (SiC) resiste relativamente bien las altas temperaturas y la oxidación. Al combinar ambos, se obtiene un material compuesto ligero y resistente. En adelante, este material se denominará C/SiC.

Existen principalmente dos métodos para fabricar este material. Uno es la infiltración química de vapor. Se hace penetrar un gas para rellenar los espacios entre las fibras con cerámica. El otro es la infiltración y pirólisis de polímeros. Se introduce resina y se hornea para transformarla en cerámica. Ambos métodos rellenan los espacios entre los haces de fibras con cerámica.

El problema comienza ya en el proceso de fabricación. La fibra de carbono y la matriz de carburo de silicio se expanden en distinta medida al recibir calor. Esto se denomina diferencia en el coeficiente de dilatación térmica. Al enfriarse desde un estado caliente hasta la temperatura ambiente, se contraen en proporciones diferentes. Como resultado, se forma una red de microgrietas dentro de la matriz. Estas microgrietas se generan de forma inevitable durante el proceso de fabricación.

Las grietas se convierten en vías de entrada para el oxígeno. Aunque por fuera parezca intacto, el oxígeno penetra hacia el interior. Es similar a cómo el agua se filtra por una pequeña rendija en el papel de pared. El oxígeno que ingresa ataca las fibras de carbono y la interfase. La interfase es el delgado límite entre la fibra y la matriz.

La matriz de carburo de silicio reacciona de dos maneras al entrar en contacto con el oxígeno. Cuando la temperatura no es excesivamente alta y hay suficiente oxígeno, se forma una película vítrea sobre la

superficie. Esta película es una delgada capa protectora de dióxido de silicio. A esto se le llama oxidación pasiva. Dicha película vítrea bloquea la penetración del oxígeno y protege el interior del material.

Por el contrario, cuando la temperatura es sumamente alta o la presión parcial de oxígeno es baja, la situación cambia. No se puede formar la película vítrea protectora. En su lugar, el carburo de silicio se volatiliza directamente en forma gaseosa. Esto se conoce como oxidación activa. En la oxidación activa, la superficie continúa erosionándose sin ninguna capa protectora. Este riesgo aumenta en el vuelo hipersónico a gran altitud.

Cuando la interfase desaparece, las ventajas del material compuesto se derrumban. Originalmente, las fibras sostienen la grieta como un puente a medida que esta se propaga. Esto se denomina puenteado de fibras. Si la interfase se destruye por oxidación, esta fuerza de sujeción desaparece. Entonces, el material se vuelve propenso a fracturarse de manera repentina. También se produce un daño por delaminación, en el que las capas se separan unas de otras.

Este daño no se manifiesta con facilidad en el exterior. Aunque la superficie parezca intacta, la interfase interna se degrada primero. Por ello, una inspección visual no resulta suficiente. Es necesario evaluar la penetración de oxígeno y el estado de la interfase dentro del material. Los modelos de vida útil basados en inteligencia artificial detectan este daño oculto en el interior y no solo en la superficie.

Un estudio publicado en 2026 ilustra este daño mediante cifras. El equipo de Wu predijo el daño por oxidación de un C/SiC trenzado en 2.5D mediante un modelo físico. En este estudio, la resistencia a la tracción a temperatura ambiente fue de unos 264 megapascales. Al subir a 1100 grados, la resistencia cayó a unos 184 megapascales. El módulo de elasticidad también se redujo alrededor de un 23 por ciento en el mismo intervalo. Este estudio emplea un modelo constitutivo basado en la física y no un modelo de inteligencia artificial.

Existe otra vía para bloquear el oxígeno: aplicar un recubrimiento protector sobre la superficie. Esto se denomina recubrimiento de barrera ambiental. Este recubrimiento impide que el oxígeno y el vapor de agua penetren al interior. En la investigación hipersónica, el objetivo son recubrimientos capaces de soportar más de 2000 grados Celsius. La inteligencia artificial se utiliza para determinar el espesor y la composición de estos recubrimientos.

Este resultado sugiere algo importante. Si solo se observan las propiedades promedio del material, se pasa por alto el daño en la interfase. Es necesario analizar la ubicación de por dónde y en qué medida ha penetrado el oxígeno. También se deben examinar conjuntamente el historial de temperatura y el historial de cargas. Los modelos de vida útil con inteligencia artificial que se presentarán más adelante se encargan precisamente de esta labor.

Conviene señalar también las limitaciones en este punto. El modelo del equipo de Wu es un cálculo basado en la física. Predice el daño incluso sin inteligencia artificial. La inteligencia artificial se utiliza para acelerar los cálculos aprendiendo a partir de los datos de estos modelos físicos. Los modelos físicos y la inteligencia artificial no son rivales, sino compañeros que se complementan mutuamente.

5.1.3 Proceso de fusión y desprendimiento en cerámicas de ultraalta temperatura (UHTC)

Las cerámicas de ultraalta temperatura son cerámicas con un punto de fusión sumamente elevado. Los materiales representativos son el diboruro de circonio y el diboruro de hafnio. El diboruro de circonio es un compuesto de circonio y boro. El punto de fusión de estos materiales supera los 3000 grados Celsius. Por esta razón, se investigan como candidatos para bordes de ataque hipersónicos y componentes de cohetes.

Un alto punto de fusión no garantiza la tranquilidad. El verdadero problema es la oxidación. Cuando el diboruro de circonio entra en contacto con el oxígeno, se forman dos tipos de óxidos. Uno es un óxido duro llamado circonia. El otro es un óxido de boro denominado trióxido de diboro.

El trióxido de diboro muestra una propiedad interesante a bajas temperaturas. Se funde y se convierte en líquido a unos 450 grados Celsius. Este líquido sella las pequeñas grietas de la superficie. Funciona de manera semejante a una pomada que cubre y rellena una herida. A esto se le denomina película autorreparable. En la etapa inicial de oxidación, esta capa protege el material.

Sin embargo, a medida que la temperatura aumenta, esta defensa se desmorona. Al superar aproximadamente los 1100 grados Celsius, el trióxido de diboro se evapora gradualmente. Hacia los 1500 grados Celsius, esta capa prácticamente desaparece. La temperatura límite exacta varía según la presión parcial de oxígeno, la velocidad del gas y la presión. Por ello, resulta difícil definirla con un único valor fijo.

Al añadir carburo de silicio, la situación mejora. El vidrio líquido procedente de la oxidación del carburo de silicio se mezcla con el trióxido de diboro. Ambos componentes se combinan para formar una fase vítrea llamada borosilicato. Este vidrio cubre la superficie hasta temperaturas más elevadas.

Cuando la temperatura sube aún más, ni siquiera este vidrio logra resistir. Al superar aproximadamente los 1700 grados Celsius, la viscosidad del vidrio disminuye. El vidrio se evapora y es arrastrado por el rápido flujo de gas. Al desaparecer la capa protectora, la superficie se desgasta con rapidez. En última instancia, sobre la superficie solo queda un esqueleto poroso de circonia. Este esqueleto poroso permite que el oxígeno penetre hacia el interior con mayor facilidad.

La circonia presenta otra trampa adicional. Su estructura cristalina cambia según la temperatura. Durante el proceso de enfriamiento, pasa de una estructura tetragonal a una monoclínica. Estos dos nombres se refieren a diferencias en el empaquetamiento atómico. Cuando ocurre esta transformación, el volumen aumenta aproximadamente un 4.5 por ciento.

Al expandirse repentinamente el volumen, se genera una gran tensión dentro de la capa de óxido. Es similar a cómo el agua, al congelarse en invierno, rompe una botella. Esta fuerza produce grietas en la capa de óxido. Cuando las grietas se conectan, la capa de óxido se desprende en fragmentos. Esto se conoce como desprendimiento. En el lugar desprendido queda expuesta una nueva superficie y la oxidación comienza nuevamente.

Los boruros de alta entropía representan un nuevo enfoque para este problema. Las cerámicas de alta entropía son cerámicas en las que se mezclan varios elementos metálicos en cantidades similares. En ocasiones se introducen cinco metales en un solo cristal. La amplitud de sus composiciones las hace idóneas para la búsqueda mediante inteligencia artificial. Es posible filtrar primero por ordenador qué combinaciones son resistentes a la oxidación.

Un artículo de revisión de 2026 sobre cerámicas de ultraalta temperatura ilustra bien esta tendencia. Esta revisión aborda de forma conjunta la experimentación, la simulación multiescala y el diseño basado en datos. El enfoque multiescala consiste en conectar múltiples niveles, desde la escala atómica hasta el tamaño de los componentes. El diseño basado en datos es un método en el que la inteligencia artificial selecciona la composición y la microestructura.

La inteligencia artificial también se utiliza para predecir el espesor de la capa de óxido. Un equipo de investigación predijo el daño por oxidación de los boruros de ultraalta temperatura mediante aprendizaje automático. Utilizaron un método llamado Random Forest para estimar el espesor de la capa de óxido. Random Forest es un método que reúne múltiples reglas de decisión para ofrecer una respuesta. No obstante, el comportamiento real de la capa de óxido debe confirmarse mediante ensayos

en túnel de viento y arco de plasma (arcjet). El arcjet es un dispositivo que expulsa gas a temperaturas ultraaltas para simular el entorno de vuelo.

Los potenciales interatómicos de aprendizaje automático (MLIP) también se emplean para predecir el punto de fusión. Este es un método en el que la inteligencia artificial calcula rápidamente las fuerzas entre los átomos. En 2025, un equipo de investigación utilizó MLIP para estimar el punto de fusión de cerámicas de alta entropía. Filtraron primero por ordenador las combinaciones con puntos de fusión elevados. Este método reduce los candidatos antes de realizar los experimentos. De este modo, disminuye el número de costosas síntesis y pruebas. Cuando el cálculo filtra previamente de esta forma, la búsqueda general de materiales se acelera.

Aun así, queda una pregunta: ¿son realmente resistentes a la oxidación las combinaciones seleccionadas por el ordenador? El cálculo asume condiciones ideales. En la superficie real existen impurezas y defectos. El flujo y la presión del gas también difieren de los cálculos. Por ello, el veredicto final lo determinan las pruebas reales. La inteligencia artificial se limita a seleccionar de manera inteligente los candidatos que se van a probar.

5.2 Cómo convertir los granos cristalinos en grafos para predecir la vida útil

Lo que se aborda en esta sección

Esta sección aborda cómo transformar el interior de los materiales en un formato legible para la inteligencia artificial. La herramienta empleada es la red neuronal de grafos espaciotemporales. Un grafo es una representación que dibuja relaciones mediante puntos y líneas. El mapa del metro es un buen ejemplo: las estaciones son puntos y los tramos son líneas.

Los granos cristalinos del interior del material se convierten en puntos, y los límites de grano en líneas. A esto se le añade el paso del tiempo, ya que el daño se acumula al recibir calor y fuerzas de manera repetida. Esta sección explica cómo se construye dicho grafo, cómo se propaga el daño y cómo se integran las leyes físicas en el aprendizaje. Se denomina modelo espaciotemporal porque considera el tiempo y el espacio de forma conjunta.

5.2.1 Conversión de la microestructura en grafos

Los granos cristalinos son pequeños granos de cristal dentro del material. Los metales o las cerámicas no son un solo cristal grande, sino que están formados por innumerables granos pequeños agrupados. En cada grano, la orientación hacia la que apuntan los átomos varía ligeramente. A esta dirección se la denomina orientación cristalográfica.

El equipo representativo para leer la orientación cristalográfica es la difracción de retrodispersión de electrones (EBSD). Dentro de un microscopio electrónico, se disparan electrones para leer la orientación de cada grano. Este instrumento escanea la superficie y registra la orientación de cada uno de los granos. El resultado se presenta como un colorido mapa de orientaciones.

El límite de grano es el borde donde se encuentran dos granos. Este límite tiende a ser el eslabón débil del material. El oxígeno se infiltra a lo largo de este camino y las grietas también crecen siguiendo estos bordes. Por tanto, conocer las propiedades de los límites es fundamental para predecir la vida útil.

El método para convertirlo en un grafo parte de una idea sencilla. Cada grano se define como un punto, es decir, un nodo. Si dos granos están en contacto, se unen mediante una línea, es decir, una arista (edge). De este modo, una fotografía de microscopio compleja se convierte en un diagrama de relaciones ordenado. La inteligencia artificial interpreta este diagrama de relaciones mejor que una fotografía.

En los nodos se almacena la información de los granos. Se incluyen el tamaño, la forma y la orientación del grano. También se incorporan la temperatura y el esfuerzo en esa ubicación. El esfuerzo es la intensidad de la fuerza que comprime o estira un objeto. De esta manera, un solo punto resume el estado de dicho grano.

En las aristas se almacena la información de los límites. Se incluye la diferencia de orientación entre dos granos. También se incorporan el área y la orientación del límite, así como el grado de daño causado por la oxidación. Cuanto mayor es la diferencia de orientación en el límite, más tiende a penetrar el oxígeno.

Este enfoque ya ha sido probado en otros materiales. En 2023, un equipo de investigación aplicó este método de grafos al acero. Clasificaron los puntos donde el daño por fatiga era severo en cada grano del acero. La fatiga es el daño que se produce al recibir fuerzas de forma repetida. Convirtieron los mapas EBSD en grafos y los entrenaron. En ese estudio, el método de grafos demostró su utilidad para identificar las zonas de daño grano a grano. Sin embargo, lo que abordó ese trabajo fue la clasificación de la fatiga en el acero; no fue un estudio que predijera la vida útil de las cerámicas, la rotura por fluencia lenta (creep) ni los cambios a lo largo del tiempo.

Trasladar este método de grafos a la predicción de la vida útil de las cerámicas es la dirección que propone este libro. Aún no se ha llevado a cabo una demostración empírica en cerámicas. La siguiente explicación no describe hechos comprobados, sino el rumbo hacia el que debe avanzar la investigación.

También surgen nuevos métodos para tratar la información de los límites. En 2026, un equipo de investigación rellenó las áreas vacías de mapas EBSD mediante aprendizaje basado en grafos. Se trata de una técnica para completar las regiones omitidas en la medición de acuerdo con las reglas cristalográficas. Este método introduce las reglas cristalinas en la inteligencia artificial para evitar que se rellenen orientaciones erróneas. Dado que los datos de materiales no siempre son completos, estas tecnologías complementarias resultan muy útiles.

El método de grafos tiene ventajas frente a las fotografías. Una fotografía procesa píxel por píxel, por lo que no sabe por sí misma dónde empieza y termina un grano. El grafo, en cambio, agrupa el grano como una unidad. Esto se parece más a la forma en que los seres humanos observan los materiales, lo que facilita el aprendizaje de las relaciones entre los granos.

Los grafos son fáciles de manejar incluso cuando el tamaño varía. Las fotografías deben tener un tamaño fijo. Los grafos se procesan de la misma manera aunque el número de granos aumente o disminuya. Esta propiedad resulta útil porque el número de granos difiere de un material a otro. Las reglas aprendidas en probetas pequeñas pueden trasladarse y probarse en componentes grandes. Sin embargo, al aumentar el tamaño, la distribución de defectos y los gradientes térmicos también cambian. Por ello, se requiere una etapa de reverificación en componentes grandes.

Los grafos construidos de este modo también pueden transferirse y aplicarse a las cerámicas, ya que estas también están compuestas por granos cristalinos y límites de grano. La oxidación y la fluencia lenta (creep) de las cerámicas de ultraalta temperatura también comienzan en los límites. El creep es el fenómeno por el cual un material se deforma lentamente al soportar fuerzas durante mucho tiempo. Por eso, este método de grafos sirve como base para predecir la vida útil de las cerámicas.

5.2.2 Aprendizaje de grafos que considera conjuntamente los vecinos y el tiempo

Ahora se añade el paso del tiempo al grafo. El daño no ocurre en un solo instante; se acumula poco a poco a medida que el calor y las fuerzas se repiten. Por ello, incluso en el mismo material, el grafo cambia con el tiempo. El seguimiento de esta evolución es la tarea del modelo espaciotemporal.

Primero se examina el aspecto espacial. El estado de un grano se ve afectado por sus vecinos. Si un grano adyacente recibe mucha fuerza, esta se transmite también al grano propio. El cálculo que reúne esta información de los vecinos es la convolución de grafos. Cada nodo recibe la información de sus nodos vecinos y actualiza su propio estado.

No todos los vecinos se tratan por igual. Algunos límites contribuyen en gran medida al daño, mientras que otros lo hacen en menor grado. Por lo tanto, el modelo asigna diferentes pesos a cada vecino. El método para determinar estos pesos es el mecanismo de atención, una técnica que presta mayor atención a los vecinos importantes.

La clave radica en incorporar el conocimiento físico en el mecanismo de atención. Si la diferencia de orientación en el límite es grande, la energía del límite es alta. Si la energía del límite es alta, el oxígeno penetra con mayor facilidad. Al reflejar estas reglas físicas en el cálculo de los pesos, el modelo evita conclusiones erróneas incluso con pocos datos.

El aspecto temporal se aborda de forma similar. El estado anterior da lugar al estado siguiente. El daño de hoy se suma al daño acumulado hasta ayer. El modelo aprende esta cadena temporal, para lo cual utiliza una arquitectura de red neuronal que procesa secuencias temporales.

Los daños importantes tratados en esta sección tienen una dirección. En los compuestos de matriz cerámica, primero se generan finas microgrietas en la matriz. A lo largo de estas grietas, la interfaz entre las fibras y la matriz se desprende. Por ese lugar penetra el oxígeno y crece una capa de óxido. Por lo general, el daño se propaga siguiendo la dirección de las fuerzas y del gas.

Este daño direccional influye notablemente en la vida útil del material. Si las grietas y la delaminación interfacial se conectan en una dirección, esa dirección se debilita. Si los poros se interconectan, los canales para el oxígeno se ensanchan. Las fases vítreas intercaladas en los límites también pueden desplazarse y transmitir el daño. Estos cambios alteran la velocidad de fluencia lenta y aceleran la rotura. La rotura se refiere al momento en que el material no resiste más y se fractura.

El deslizamiento de los límites de grano es también un daño importante en esta sección. A altas temperaturas, los granos se deslizan entre sí a lo largo de los límites, algo parecido a cómo resbalan los ladrillos mojados unos sobre otros. Este deslizamiento incrementa la deformación por fluencia lenta. La orientación y las propiedades de los límites determinan la magnitud del deslizamiento. Por ello, las aristas que contienen la información de los límites son fundamentales.

El gas caliente generalmente fluye en una dirección y empuja la superficie. Por ello, el daño en las cerámicas también se desarrolla con una direccionalidad. El modelo de grafos espaciotemporales busca incorporar esta direccionalidad en el aprendizaje, lo que permite prever hacia qué dirección crecerá el daño. Este tipo de predicción direccional todavía carece de suficientes datos de validación.

Existen dos métodos para abordar el tiempo. Uno consiste en leer el orden cronológico de forma secuencial, similar a leer una oración desde el principio. El otro consiste en tratar directamente la velocidad de cambio, aprendiendo con qué rapidez crece el daño. Ambos métodos reflejan el proceso mediante el cual se acumula el daño.

El mecanismo de atención es hoy en día una herramienta central de la inteligencia artificial. La inteligencia artificial que traduce textos también utiliza este método, cuyo principio es concentrarse más en las palabras importantes de una oración. En los grafos de materiales, se concentra en los vecinos más relevantes para el daño. El mismo principio de atención se aplica tanto a la traducción de idiomas como al análisis de materiales.

Este enfoque de grafos espaciotemporales es una tecnología que aún se encuentra en desarrollo. No abundan las validaciones a gran escala en compuestos de matriz cerámica. Para realizar buenas predicciones, se requieren abundantes datos de calidad. La tarea misma de recopilar datos de ensayos a temperaturas ultraaltas resulta difícil. Por ello, es necesaria la ayuda de las leyes de la física, que se tratarán en el siguiente subapartado.

5.2.3 Rotura por fluencia lenta direccional y aprendizaje con reglas físicas

La fluencia lenta (creep) es el fenómeno por el cual un material se estira lentamente al soportar fuerzas durante mucho tiempo. Es similar a cómo una cuerda de tender la ropa colgada durante mucho tiempo se comba poco a poco. A altas temperaturas, este fenómeno se acelera. Los componentes de cohetes soportan fuerzas durante periodos prolongados a altas temperaturas. Por ello, el creep se convierte en una variable decisiva en la vida útil de las piezas de cohetes.

El daño por fluencia lenta se acumula de forma diferente según la dirección. A esto se le llama anisotropía. La anisotropía es la propiedad por la cual las características varían según la dirección, de manera similar a cómo la madera se divide fácilmente a lo largo de sus vetas. Los materiales reciben fuerzas simultáneamente desde múltiples direcciones, lo que se conoce como carga multiaxial.

Los investigadores resumen este daño en una única variable, denominada variable de daño. Si la variable de daño es 0, el material está intacto; si es 1, significa que esa zona se ha fracturado. A medida que el daño aumenta, la sección restante soporta una fuerza mayor. Entonces, el daño crece con mayor rapidez. Debido a esta retroalimentación, el daño aumenta de forma incontrolable y precipita la rotura.

Al aumentar la temperatura, la fluencia lenta se acelera debido a que los átomos se mueven con mayor actividad. Esta dependencia de la temperatura se modela mediante una función exponencial. Incluso un leve incremento de temperatura dispara la tasa de daño. Por eso, reducir la temperatura máxima superficial resulta crucial. Aquí convergen los recubrimientos de la sección anterior y el flujo térmico de esta sección.

La oxidación acelera aún más este daño. A medida que la capa de óxido se vuelve más gruesa, los límites se debilitan. Por ello, el espesor de la capa de óxido se incorpora en la ley de daño. Esto significa que la fluencia lenta y la oxidación no se analizan por separado, sino conjuntamente, ya que en el entorno de un cohete ambos fenómenos ocurren siempre a la vez.

Existe aquí una regla física fundamental: el daño solo crece en una dirección. Una grieta que ya se ha formado no cicatriza por sí sola. Con el paso del tiempo, la variable de daño no disminuye. A esta regla se la denomina irreversibilidad del daño, la cual se sitúa en el mismo contexto que la segunda ley de la termodinámica.

Si la inteligencia artificial infringe esta regla, las predicciones resultan erróneas. Ofrecer una respuesta en la que el daño disminuye contradice la realidad. Por lo tanto, se aplica una penalización durante el entrenamiento. Si se produce una respuesta que viola las reglas, se impone una pérdida elevada. Incorporar esta penalización a la función de pérdida es el método del aprendizaje informado por la física.

En el aprendizaje informado por la física también se integran otras reglas. El equilibrio de fuerzas es una de ellas: la suma de las fuerzas que empujan y tiran en un mismo punto debe cuadrar. También se introduce la regla de que la masa no desaparece ni se crea. Al añadir estas reglas a la función de pérdida, el modelo evita respuestas contrarias a la física.

El criterio de rotura también evalúa la dirección. Los materiales son vulnerables a las fuerzas de tracción. Por ello, se analiza específicamente la dirección en la que la fuerza de tracción es máxima.

También se considera la presión interna que tiende a expandirlo. Las fuerzas procedentes de múltiples direcciones se consolidan en un único nivel de riesgo. Este nivel de riesgo determina la velocidad a la que progresa el daño.

La raíz de este enfoque es la red neuronal informada por la física (PINN). En 2019, el equipo de Raissi formalizó este concepto. Ellos incorporaron directamente las ecuaciones de la física en el entrenamiento de la red neuronal. La idea consiste en que, incluso con pocos datos, las reglas físicas cubren los vacíos. Esta noción se aplica directamente a la predicción de la vida útil de las cerámicas.

El verdadero valor de las reglas físicas se revela en la extrapolación de las predicciones. Los datos de ensayo suelen concentrarse a bajas temperaturas. Lo que realmente se desea conocer es el comportamiento cerca de los 2000 °C. Apenas existen datos a estas temperaturas tan altas. Al contar con reglas físicas, es posible prever de forma razonable incluso aquellas regiones donde no hay datos.

No obstante, este método también tiene sus límites. Si se introducen las reglas físicas de manera errónea, la predicción puede empeorar. Determinar el peso relativo entre las reglas también resulta complejo. No se puede omitir el paso de validación con datos reales de fractura. Por ello, las predicciones de la inteligencia artificial no reemplazan los ensayos, sino que se utilizan como una herramienta para reducirlos.

5.3 Modelos generativos para el diseño del tejido de fibras y recubrimientos

La sección anterior abordó la predicción de la vida útil de materiales ya fabricados. Esta sección cambia de dirección. Primero se define el rendimiento deseado y luego se busca a la inversa la estructura que lo produce. Esto se denomina diseño inverso. En términos culinarios, equivale a definir el sabor deseado y buscar la combinación de ingredientes adecuada.

El objeto de esta sección son las estructuras de fibras tejidas en tres dimensiones y los recubrimientos interfaciales. Se trata de determinar en qué ángulo tejer las fibras y qué espesor de recubrimiento aplicar. Para ello se utiliza una inteligencia artificial generativa llamada modelo de difusión. Un modelo de difusión es un método que genera una estructura nítida a partir de ruido difuso. A esto se le añade la propiedad de que los resultados se mantengan invariantes ante cambios de orientación.

5.3.1 Estructuras de fibras tejidas tridimensionalmente y aprendizaje equivariante respecto a la orientación

Los materiales compuestos trenzados en 3D son materiales elaborados tejiendo haces de fibras de forma tridimensional. Es similar a una cesta tejida densamente en tres direcciones. Los materiales compuestos convencionales se fabrican apilando capas delgadas. Como resultado, la unión entre capas es débil y se separan con facilidad. Esto se denomina delaminación.

El trenzado en 3D elimina esta debilidad. Al entrelazarse las fibras en múltiples direcciones, no existen límites de capa. Por lo tanto, rara vez se producen daños por separación de capas. Esto resulta ventajoso en zonas sometidas a un fuerte choque térmico, como los bordes de ataque hipersónicos. El choque térmico es el daño que se produce cuando un material se calienta o se enfría repentinamente.

Existen varios tipos de métodos de trenzado. Las fibras pueden tejerse en cuatro direcciones o incluso en muchas más. Cuantas más direcciones haya, más uniformemente resistirá diversas cargas. A cambio, el proceso de tejido se vuelve más complejo. Por lo tanto, se debe seleccionar el patrón de tejido adecuado para el uso de cada componente. Esta elección del tejido también es objeto del diseño inverso.

El problema radica en que el tejido nunca es perfecto. En la fabricación real, las fibras se tuercen. También pueden aplastarse y distorsionar su sección transversal. El ángulo de tejido fluctúa

ligeramente. El espesor del recubrimiento tampoco es uniforme en todas partes. Estas fluctuaciones hacen que la fuerza se concentre en puntos específicos.

Estas fluctuaciones se tratan de manera probabilística, ya que no todos los productos son idénticos y difieren ligeramente entre sí. Por ello, en lugar de una única estructura correcta, se trabaja con una distribución de estructuras. Los modelos de difusión son idóneos para manejar esta distribución. El modelo de difusión puede generar múltiples estructuras verosímiles.

Aquí cobra gran importancia la propiedad llamada equivariancia. La equivariancia es la propiedad por la cual, si se rota un objeto, el resultado rota exactamente de la misma manera. Si una pieza se gira hacia la izquierda, la fuerza predicha también debe girar hacia la izquierda. Sin esta propiedad, una misma estructura se juzgaría de forma diferente según su orientación espacial, lo cual contradice las leyes de la física.

En el espacio tridimensional, los objetos pueden rotar o trasladarse. Existe una rama matemática que aborda estos dos movimientos en conjunto: el grupo euclidiano especial. Aunque el nombre suene difícil, su significado es simple. Es un conjunto de reglas que agrupa las rotaciones y las traslaciones. La inteligencia artificial se diseña para que siempre respete estas reglas de rotación y traslación.

Estos modelos de difusión equivariantes se desarrollaron primero en el diseño molecular. En 2025, el equipo de Chen presentó un modelo de difusión equivariante que combina grafos a múltiples escalas. Con este modelo diseñaron moléculas en 3D. Las moléculas también son objetos que rotan y se trasladan en un espacio tridimensional. Por lo tanto, las mismas matemáticas pueden transferirse a las estructuras de fibras. Sin embargo, dicha investigación abordó moléculas y no el tejido de fibras cerámicas. Los haces de fibras, las restricciones de tejido, la porosidad y la viabilidad de fabricación aún no han sido validados con este modelo. Su adaptación al trenzado cerámico es la dirección que propone este libro.

Veamos un ejemplo de por qué la equivariancia es importante. Imaginemos que colocamos un mismo tornillo en distintas posiciones sobre un escritorio. La resistencia del tornillo es independiente de su orientación. La inteligencia artificial también debe respetar este hecho. Sin equivariancia, juzgaría el mismo tornillo de manera diferente según su dirección. En ese caso, habría que recopilar datos por separado para cada orientación, lo que desperdiciaría enormemente el aprendizaje.

Mantener la equivariancia permite ahorrar datos. Con los datos de una sola orientación se pueden cubrir todas las direcciones. En situaciones donde los datos de materiales son escasos, este ahorro resulta significativo. Por eso es adecuado para campos con pocos datos, como las cerámicas de ultraalta temperatura. Puede considerarse una forma de inscribir las leyes de la física en la propia estructura.

En resumen, convergen tres elementos: la estructura tridimensional, las fluctuaciones probabilísticas y la propiedad de invariancia ante la orientación. Estos tres aspectos se integran en un solo modelo. Solo así se puede reflejar fielmente la variabilidad de la fabricación real. En la siguiente sección veremos cómo este modelo va dando forma a las estructuras.

5.3.2 Modelos de difusión que reconstruyen estructuras a partir del ruido

La idea básica del modelo de difusión se divide en dos etapas. La primera es un proceso que difumina la estructura convirtiéndola en ruido. La segunda es el proceso de reconstruir la estructura a partir del ruido. La inteligencia artificial aprende este proceso de reconstrucción. Es equivalente a reproducir en reversa la difusión de una gota de tinta en el agua.

Veamos primero la etapa inicial. Se añade ruido gradualmente a una estructura de fibras nítida. A medida que avanza el tiempo, la estructura se desdibuja. Al final, solo queda ruido puro. Este proceso se

lleva a cabo lentamente siguiendo una regla predeterminada, la cual se denomina programa de ruido (noise schedule).

La etapa posterior recorre este proceso a la inversa. Comenzando desde el ruido, se reconstruye la estructura paso a paso. En cada iteración se elimina un poco de ruido. En este punto, existe un valor que indica en qué dirección eliminarlo: la puntuación (score). La puntuación señala la dirección en la que la estructura se vuelve más verosímil.

La inteligencia artificial aprende precisamente esta puntuación. Al observar una estructura difusa, predice la dirección que conduce a una forma nítida. Una vez finalizado el entrenamiento, es capaz de forjar estructuras incluso a partir de ruido puro. Este método se denomina difusión basada en puntuación (score-based diffusion). La inteligencia artificial generadora de imágenes actual también dibuja siguiendo este mismo principio.

Existe una razón por la cual este método es adecuado para el diseño de materiales. En las estructuras de los materiales no hay una sola respuesta correcta. Existen múltiples estructuras que ofrecen el mismo rendimiento. El modelo de difusión genera estos múltiples candidatos, permitiendo al ser humano elegir entre ellos el más fácil de fabricar. Su gran utilidad reside en que ofrece diversas alternativas en lugar de una única solución.

Incorporar condiciones a este proceso es el núcleo del diseño inverso. Generar una estructura cualquiera al azar no tiene utilidad; se debe crear una estructura que posea el rendimiento deseado. Por ello, el rendimiento objetivo se introduce como condición. Entre los objetivos se incluyen la conductividad térmica, la resistencia a la tracción, la tenacidad a la fractura y la velocidad de ablación.

La tenacidad a la fractura es la resistencia del material a la propagación de grietas. La velocidad de ablación es la rapidez con la que se erosiona la superficie. Estos valores se fijan en los niveles deseados. Entonces, el modelo selecciona y genera una estructura acorde a dichas condiciones. También es posible regular con qué intensidad se reflejan los objetivos.

Esta tecnología de condicionamiento ya se aplica en diversos materiales. En 2025, un equipo de investigación presentó un modelo de difusión en 3D que incorporaba restricciones físicas. Diseñaron compuestos poliméricos reforzados con fibras adaptados a rendimientos específicos. Definieron la curva de tensión-deformación deseada y buscaron la estructura correspondiente a la inversa. Este método hizo que se respetaran directamente las reglas físicas durante el proceso generativo.

Si se imponen condiciones demasiado estrictas, surge otro problema. Al perseguir únicamente el rendimiento objetivo, se pueden generar estructuras imposibles de fabricar: patrones de tejido que lucen impecables sobre el papel pero que en la práctica no se pueden tejer. Por ello, se guía el modelo para que busque soluciones dentro del rango factible de fabricación, introduciendo las reglas de manufactura como condiciones adicionales. Lograr este equilibrio entre el rendimiento y la viabilidad de fabricación es la parte más compleja del diseño inverso.

Para aplicar este enfoque a los materiales compuestos cerámicos, aún queda un obstáculo por superar: es necesario reunir abundantes datos sobre la estructura tridimensional de las cerámicas. Estos datos se obtienen mediante tomografía computarizada por rayos X, que funciona bajo el mismo principio que las tomografías (TC) de los hospitales. Permite visualizar el interior de una pieza en 3D sin necesidad de cortarla. Las estructuras obtenidas de este modo sirven como base fundamental para el modelo.

Recopilar estos datos requiere tiempo. Tomografiar una sola probeta toma un periodo prolongado, y segmentar los datos capturados en granos y fibras también exige un gran esfuerzo manual. Por eso los datos son siempre escasos. La equivariancia analizada anteriormente ayuda a mitigar esta carencia, ya que permite aprender múltiples configuraciones a partir de una cantidad reducida de datos.

5.3.3 Determinación simultánea del ángulo de fibra y del espesor del recubrimiento

El modelo de diseño inverso toma dos decisiones principales: una es el ángulo de tejido de las fibras y la otra es el espesor del recubrimiento interfacial. Ambas variables determinan en gran medida la resistencia mecánica y la tolerancia térmica del material.

Analicemos primero el ángulo de las fibras. Las fuerzas suelen actuar de manera predominante en una dirección específica. Las toberas de cohete reciben una fuerza de empuje desde el interior, mientras que los bordes de ataque hipersónicos soportan esfuerzos cortantes en la dirección del flujo. Cuando las fibras se orientan alineadas con la dirección de las fuerzas, el material las soporta con mayor eficacia. Por ello, el modelo determina el ángulo teniendo en cuenta dichas fuerzas.

El espesor del recubrimiento es igualmente crucial. Aquí es necesario distinguir entre dos tipos de capas. Una es la delgada capa interfacial que se coloca entre las fibras y la matriz. Esta capa interfacial se fabrica con nitruro de boro o carbono pirolítico. Esta película evita que las grietas fracturen directamente las fibras, haciendo que la grieta se deslice a lo largo de la capa y disipe la energía. Si la capa interfacial es demasiado delgada, este efecto se debilita; si es excesivamente gruesa, surgen otros problemas.

El otro tipo es el recubrimiento de barrera ambiental aplicado en el exterior del material. Este recubrimiento impide que el oxígeno y el vapor de agua penetren en el interior. Los disilicatos de tierras raras son candidatos destacados para este recubrimiento externo, ya que bloquean eficazmente el oxígeno a altas temperaturas. La capa interfacial y el recubrimiento exterior difieren tanto en ubicación como en función. En ocasiones, ambas películas se apilan en múltiples capas, asignando una función distinta a cada una.

El modelo de diseño inverso aborda estas dos decisiones de manera simultánea. Busca aumentar la resistencia mecánica al tiempo que reduce el esfuerzo térmico. Estos dos objetivos a menudo entran en conflicto: ganar en un aspecto suele implicar perder un poco en el otro. Este tipo de problema se conoce como optimización multiobjetivo. El modelo encuentra un punto de equilibrio óptimo entre la resistencia y el esfuerzo térmico.

La concentración de fuerza en un punto específico se denomina concentración de tensiones. Es similar a una prenda que comienza a rasgarse por una costura suelta. Las grietas se originan primero en los puntos con mayor concentración de tensiones. Por esta razón, el diseño busca distribuir las fuerzas de manera uniforme. Este es el motivo por el cual se optimizan el ángulo de las fibras y el espesor del recubrimiento.

Al buscar este equilibrio, también se evalúa la continuidad geométrica. Si una fibra se dobla de forma abrupta, las fuerzas se concentran en ese punto; lo mismo ocurre si el espesor del recubrimiento cambia bruscamente. Por ello, el modelo promueve que los ángulos y los espesores varíen de manera suave y continua, aplicando penalizaciones a los diseños que presenten variaciones drásticas.

Este tipo de diseño inverso se encuentra todavía en una fase cercana a la prueba de concepto. Cualquier afirmación sobre la magnitud de mejora en un valor de rendimiento específico debe tomarse con cautela. Es imprescindible verificar si la estructura generada por el modelo realmente ofrece dicho rendimiento. Esta validación se realiza, nuevamente, mediante fabricación y ensayos experimentales. La inteligencia artificial asume el papel de reducir y filtrar el abanico de candidatos.

En 2026 surgieron investigaciones sobre generación condicionada a partir de texto. Un equipo de investigación generó microestructuras utilizando frases que describían las propiedades deseadas. Aunque se trata de un estudio preliminar que aún no ha pasado por revisión por pares, la dirección es

clara: una tendencia en la que el usuario describe el rendimiento deseado y la inteligencia artificial moldea la estructura. Esta tendencia puede extenderse al diseño de compuestos cerámicos.

5.4 Conexión del flujo térmico entre el mundo microscópico y el componente completo

Esta sección aborda el flujo de calor dentro del material. El calor es la clave que determina la vida útil del material. Es necesario saber hacia dónde y con qué rapidez fluye el calor para poder predecir la temperatura. Y solo conociendo la temperatura se pueden estimar las velocidades posteriores de oxidación y fluencia.

El problema radica en que el flujo de calor se define a escala atómica. Las vibraciones atómicas transportan el calor. Sin embargo, lo que deseamos conocer es la temperatura de todo el componente. Es necesario conectar el mundo microscópico con el macroscópico. Esta conexión se denomina acoplamiento multiescala. Esta sección explica cómo lograr dicha unión mediante redes neuronales de operadores profundos.

5.4.1 Cálculo de la conductividad térmica a partir de vibraciones atómicas (dinámica molecular de Green-Kubo)

En las cerámicas, el calor se transfiere mediante las vibraciones de los átomos. Las partículas cuantizadas de estas vibraciones se denominan fonones. Resulta sencillo concebir los fonones como cuantos de sonido. Las vibraciones se propagan a través de la red cristalina transportando calor. La facilidad con la que se propagan estas vibraciones determina la conductividad térmica.

A altas temperaturas, estas vibraciones colisionan entre sí. Cuando los fonones chocan entre ellos, el flujo de calor se interrumpe. Los límites de grano, los defectos y los átomos dopantes o mezclados también dispersan las vibraciones. Cuanto mayor sea esta dispersión, más lentamente fluirá el calor. Para comprender la conducción térmica de las cerámicas de ultraalta temperatura, es fundamental entender este fenómeno de dispersión.

Las cerámicas de alta entropía son singulares en este aspecto. Al combinar múltiples elementos, la red cristalina se vuelve irregular. Esta irregularidad dispersa intensamente las vibraciones, lo que impide que el calor se propague con facilidad. Si el objetivo es el aislamiento térmico, esta propiedad resulta muy ventajosa, ya que reduce la transferencia de calor desde la superficie hacia el interior.

La dinámica molecular es el método para calcular este mundo diminuto. En la computadora se mueve cada uno de los átomos. Se sigue el rastro de cómo vibran y chocan los átomos. Este cálculo aborda un tiempo breve en unidades de nanosegundos. Un nanosegundo es la milmillonésima parte de un segundo. Es un mundo así de pequeño y rápido.

La fórmula de Green-Kubo es la fórmula para extraer la conductividad térmica. Esta fórmula es un método que observa durante largo tiempo las fluctuaciones del flujo de calor. Es similar a averiguar las propiedades del agua observando las suaves ondas en agua tranquila. Reúne las fluctuaciones del flujo de calor que los propios átomos generan. A partir de esas fluctuaciones calcula la conductividad térmica. Esta fórmula fue establecida en la década de 1950 por Green y Kubo.

La precisión del cálculo depende de qué tan bien se conozcan las fuerzas entre los átomos. Las reglas que rigen estas fuerzas se denominan potenciales interatómicos. Antiguamente se utilizaban reglas sencillas elaboradas a mano por personas. Estas reglas eran rápidas, pero su precisión era baja. Por el contrario, los cálculos cuánticos eran precisos, pero demasiado lentos.

Aquí es donde el potencial interatómico de aprendizaje automático (MLIP) tiende un puente. La inteligencia artificial aprende los resultados de los cálculos cuánticos para predecir las fuerzas. Ofrece una precisión cercana a la de los cálculos cuánticos a gran velocidad. Este método se aborda en detalle en los capítulos anteriores de este libro. El MLIP también se utiliza en los cálculos de conducción térmica a temperaturas ultraaltas.

Las investigaciones recientes demuestran el poder de esta combinación. En 2021, un equipo de investigación estudió la zirconia mediante MLIP. Calcularon conjuntamente las transiciones de fase y la conducción térmica en función de la temperatura. Extrajeron el flujo de calor mediante el método de Green-Kubo. Este método captura sin omisiones los complejos efectos de las vibraciones atómicas.

Este cálculo tiene ciertas limitaciones que conviene señalar. El método de Green-Kubo calcula la conductividad térmica cuando el material se encuentra en equilibrio. Este valor refleja las propiedades fundamentales del interior del material. Sin embargo, la superficie donde ocurre la ablación no se encuentra en estado de equilibrio. En ese lugar, la radiación y las reacciones químicas superficiales influyen en gran medida sobre el calor. Por tanto, el valor de Green-Kubo por sí solo no puede explicar todo el calor de la superficie. Es apropiado utilizar este valor como entrada para la conducción térmica en el interior del componente.

¿Por qué bajar necesariamente hasta este mundo tan pequeño? ¿No bastaría con medir la conductividad térmica directamente? A temperaturas ultraaltas la medición resulta difícil. No es sencillo medir el interior de una cerámica a 2000 grados Celsius. A medida que avanza la oxidación, el propio material cambia. Por eso se necesita el camino de obtener los valores mediante cálculos.

Los cálculos tienen otra ventaja. Permiten previsualizar composiciones que aún no se han fabricado. Las cerámicas de alta entropía cuentan con innumerables combinaciones. Es imposible fabricarlas y medirlas todas. Los cálculos informan con antelación la conductividad térmica de los candidatos. De ese modo, se seleccionan y fabrican en la práctica únicamente las combinaciones prometedoras.

Esta tendencia continúa hoy en día. Un estudio de 2026 predijo conductividades térmicas extremadamente bajas mediante Green-Kubo y MLIP. En diversos materiales, esta combinación se está consolidando. Sin embargo, el cálculo todavía requiere computadoras de gran escala. También es necesario lidiar con errores como el truncamiento temporal y las fluctuaciones estadísticas. Por ello, se calcula múltiples veces y se obtiene el promedio.

Aquí también rige el mismo principio de validación. Los valores calculados deben contrastarse con los valores experimentales. El cálculo se pone a prueba con algunos materiales conocidos. Si el cálculo coincide bien con el experimento, aporta confianza también para los nuevos materiales. Los valores que no pasan por este proceso de ajuste no van más allá de una referencia. El cálculo y el experimento son dos ojos que se comprueban mutuamente.

5.4.2 Ecuación del flujo de calor en todo el componente

El subtítulo anterior abordó un mundo diminuto. Ahora ampliamos la mirada hacia el componente completo. La temperatura del componente difiere según la posición y cambia con el tiempo. La regla que describe esta variación es la ecuación de conducción térmica. La ecuación se escribe como una fórmula matemática, pero su significado puede explicarse con palabras.

El significado de la ecuación de conducción térmica es simple. La variación de temperatura en un punto determinado se define por la diferencia entre el calor que entra y el que sale. El calor fluye del lado caliente al lado frío. La rapidez de este flujo es la conductividad térmica. El valor calculado anteriormente se introduce aquí.

Existen varias razones por las cuales esta ecuación resulta difícil. En primer lugar, la conductividad térmica varía con la temperatura. Al subir la temperatura, la conductividad térmica cambia. A esto se le llama no linealidad. La no linealidad es una relación en la que la causa y el efecto no guardan una proporción directa. A esto se suma que el material tiene propiedades distintas según la dirección. El calor fluye mejor en la dirección en la que están dispuestas las fibras.

A esto se añade el problema de que el propio material cambia. A medida que avanza la oxidación, la conductividad térmica se modifica. Si la superficie se desgasta, el espesor disminuye. Por consiguiente, los valores de la ecuación varían con el tiempo. La oxidación y la fluencia de la sección anterior se entrelazan con el flujo de calor de esta sección. Por lo tanto, es necesario resolver estos múltiples fenómenos en conjunto, sin separarlos.

Lo que ocurre en los límites también es complejo. Sobre una superficie hipersónica se superponen diversas fuentes de calor. Entra el calor generado por el calentamiento del aire. También existe la radiación, mediante la cual la superficie emite calor en forma de luz. Es el fenómeno por el que un objeto caliente brilla al rojo vivo y pierde calor. A esto se añade además la generación de calor por reacciones químicas.

El calor químico se libera cuando los átomos se reagrupan en la superficie. A velocidades hipersónicas, las moléculas de aire se disocian en átomos. Cuando estos átomos vuelven a emparejarse en la superficie, desprenden calor. A esto se le llama recombinación catalítica. La magnitud de este calor varía según el material de la superficie. Por eso, la elección del material influye enormemente en la temperatura superficial.

Todo este calor se incorpora en las condiciones de contorno de una sola ecuación. Se disponen conjuntamente el calor entrante, la radiación saliente y la generación de calor químico. El tiempo también transcurre simultáneamente. El calor cambia continuamente a lo largo de la trayectoria de vuelo. A esto se le denomina estado no estacionario. Significa que no se aborda un único instante detenido, sino un proceso en evolución.

El método tradicional para resolver esta ecuación es el método de elementos finitos. El componente se divide finamente en pequeñas piezas. Se resuelve la ecuación para cada pieza y se unen. Este método es preciso, pero el cálculo toma mucho tiempo. Requiere cálculos iterativos en cada paso de tiempo. Resulta difícil resolver toda la trayectoria a Mach 7 en tiempo real.

En el desarrollo de cohetes y vehículos hipersónicos se necesitan respuestas rápidas. Cada vez que se modifica el diseño, la temperatura debe recalcularse. Durante las pruebas se requieren predicciones casi en tiempo real. Con cálculos lentos no es posible satisfacer estas demandas. Por esta razón entra en escena el operador de inteligencia artificial del siguiente subtítulo.

5.4.3 Red neuronal que transforma funciones en funciones (DeepONet)

La red neuronal de operadores profundos es una inteligencia artificial que aprende la relación que transforma una función en otra función. A esto se le llama DeepONet. Las redes neuronales ordinarias reciben números y producen números. DeepONet recibe funciones y produce funciones. Aquí, una función se refiere a elementos como una entrada de calor que varía con el tiempo.

Esta idea fue formalizada en 2021 por el equipo de Lu. Establecieron la teoría y la arquitectura de las redes neuronales que aprenden operadores. Esta investigación tiene sus raíces en un teorema matemático. Es el teorema de que un operador puede aproximarse mediante una red neuronal. Este teorema de aproximación constituye el fundamento teórico que sustenta a DeepONet.

DeepONet se compone de dos redes. Una es la red que lee la función de entrada. La otra es la red que lee la posición y el tiempo donde se busca la respuesta. Se multiplican los resultados de ambas redes para obtener la solución. A la primera red se le llama rama y a la segunda se le llama tronco. La rama y el tronco se unen para producir la temperatura.

¿Por qué es importante transformar una función en otra función? La entrada en un problema térmico no es un único número. La entrada es el calentamiento aerodinámico que varía con el tiempo. La salida es también un campo de temperatura que varía según la posición y el tiempo. Este tipo de problemas corresponde a una relación que conecta funciones con funciones. DeepONet aprende esta relación en su totalidad.

La fortaleza de DeepONet radica en su velocidad. Una vez que aprende, entrega la respuesta de inmediato incluso ante una nueva entrada. No resuelve el problema desde cero cada vez, como lo hace el método de elementos finitos. Permite explorar rápidamente múltiples condiciones de vuelo. Resulta idónea para observar la temperatura mientras se modifican los diseños.

El problema radica en los datos de entrenamiento. Para aprender buenas respuestas se necesita una gran cantidad de datos precisos. Sin embargo, generar datos precisos es costoso. Esto se debe a que los cálculos de elementos finitos y de dinámica molecular consumen mucho tiempo. Si los datos son escasos, la inteligencia artificial no aprende adecuadamente. El método de multifidelidad resuelve este problema.

La fidelidad indica qué tan preciso es un cálculo. La alta fidelidad representa cálculos precisos pero costosos. La baja fidelidad representa cálculos aproximados pero económicos. La multifidelidad utiliza ambos enfoques a la vez. Recopila abundantes datos económicos y solo una pequeña cantidad de datos costosos.

La idea de este método se asemeja a trazar un boceto y luego aplicar pintura encima. Primero se dibuja el panorama general con cálculos económicos. Luego, con cálculos costosos, se realiza una aplicación precisa sobre él. La inteligencia artificial aprende la diferencia entre la respuesta de baja fidelidad y la de alta fidelidad. A esta diferencia se le llama residuo. Al aprender el residuo, se incrementa la precisión incluso con pocos datos costosos.

En 2022, el equipo de Lu aplicó este DeepONet de multifidelidad a la transferencia de calor a escala nanométrica. Abordaron el problema del flujo de calor en un mundo diminuto. El método de multifidelidad redujo notablemente el error con la misma cantidad de datos costosos. También disminuyó en gran medida la cantidad necesaria de datos costosos. No obstante, este estudio se centró en la transferencia de calor a escala nanométrica. No constituye una demostración empírica que conecte la oxidación o ablación de cerámicas de ultraalta temperatura con las cargas térmicas a Mach 7.

Conviene explicar también el término homogeneización. Los compuestos cerámicos están formados por una mezcla de fibras y poros. Si se abordara cada detalle de esta compleja estructura, el cálculo resultaría abrumador. Por eso, se convierte una región pequeña en propiedades promedio. A esto se le llama homogeneización. El cálculo de baja fidelidad traza el panorama general con estas propiedades promedio.

También cabe señalar por qué resulta difícil conectar ambos mundos. El mundo atómico se mueve en unidades de nanosegundos. El mundo de los componentes abarca cientos de segundos. La escala de tiempo difiere en más de diez órdenes de magnitud. La escala de tamaño difiere en una proporción similar. Es imposible calcular esta brecha de una sola vez.

Conectar ambos mundos mediante este método en estructuras térmicas cerámicas es la vía que propone este libro. La baja fidelidad se convierte en un modelo de homogeneización simple. La alta fidelidad se

convierte en un análisis preciso que une la dinámica molecular y los elementos finitos. Se plantea un esquema donde la inteligencia artificial conecta ambos niveles. Así, la conducción térmica del mundo diminuto puede vincularse con la temperatura del componente. Esta conexión aún no ha sido ampliamente validada en estructuras térmicas cerámicas.

5.4.4 Aprendizaje con reglas físicas y aumento de velocidad

El hecho de que la inteligencia artificial sea rápida no significa que pueda dar cualquier respuesta. Las respuestas deben respetar las leyes de la física. Por ello, se incorporan reglas físicas en el aprendizaje. Este método comparte las mismas raíces que la predicción de vida útil de la sección anterior.

Primero se realiza un aprendizaje para ajustarse a los datos. Se ajusta tanto a los datos de baja fidelidad como a los de alta fidelidad. Se guía a la inteligencia artificial para que sus respuestas se aproximen a los datos. Se penaliza la diferencia entre las respuestas y los datos. A esta penalización se le llama pérdida.

A esto se suma el aprendizaje para cumplir con las ecuaciones. Se introducen las respuestas de la inteligencia artificial en la ecuación de conducción térmica. Si la respuesta satisface bien la ecuación, la penalización es pequeña. Si se desvía, la penalización es grande. De esta manera, produce respuestas conformes a la física incluso en lugares donde no existen datos.

Al incorporar reglas físicas, el modelo resiste incluso con pocos datos. El aprendizaje puramente basado en datos requiere muchos datos. Si los datos son escasos, rellena los espacios intermedios de forma arbitraria. Las reglas físicas cubren estos vacíos mediante ecuaciones. Resulta adecuado para los problemas cerámicos, donde los datos a temperaturas ultraaltas son escasos. El hecho de resistir aun con pocos datos constituye una gran ventaja del aprendizaje informado por la física.

Las condiciones de contorno se incorporan de la misma manera. Se hace cumplir el balance térmico en la superficie. El calor entrante y la radiación saliente deben coincidir. Si se infringe esta regla, se aplica una penalización. Dado que la temperatura superficial influye en gran medida en la vida útil del material, esta condición es más importante que ninguna otra.

Estas diversas penalizaciones se reúnen en una sola. Se suman la penalización de datos, la de las ecuaciones y la de las condiciones de contorno. Se asigna un peso a cada penalización. La inteligencia artificial aprende a hacer esta penalización total lo más pequeña posible. De este modo, genera predicciones rápidas y conformes a la física.

Para incorporar reglas físicas es necesario calcular derivadas. Se debe medir cuánto varía la temperatura según la posición y el tiempo. Existe una técnica para calcular estas derivadas de forma automática. A esto se le llama diferenciación automática. Sin embargo, este cálculo es pesado y reduce la velocidad.

Por ello, se emplean diversos métodos para aligerar las operaciones. Se divide el cálculo para reducir las partes que se solapan. Se transforman las coordenadas mediante funciones suaves para facilitar la derivación. Gracias a tales optimizaciones, se obtiene una inferencia mucho más rápida que con los cálculos tradicionales. No obstante, el factor concreto de aceleración o los valores de error varían según el problema. Esos valores deben citarse únicamente a partir de casos validados.

El valor de este método se manifiesta en la etapa de diseño. Si se modifica la forma de la tobera, la distribución de temperatura cambia. Con los cálculos tradicionales, se debe esperar mucho tiempo cada vez que se hace un cambio. El operador de inteligencia artificial muestra de inmediato la nueva temperatura. Así, es posible comparar decenas de formas en un solo día. El diseñador llega más rápido a una mejor solución.

También resulta útil en el campo de pruebas. Durante las pruebas de motores se requieren predicciones casi en tiempo real. Es necesario saber de antemano qué lugares alcanzan temperaturas peligrosas. El operador de inteligencia artificial genera estas predicciones con rapidez. Sin embargo, solo se puede confiar en ellas si se calibran continuamente con los datos de las pruebas. La tarea de reducir la diferencia entre las predicciones y las mediciones reales continúa.

Las limitaciones de este método también son claras. Si las reglas físicas se introducen de forma incorrecta, las respuestas empeoran. Fijar los pesos entre las reglas también requiere mucho trabajo. Además, la escasez de datos reales a temperaturas ultraaltas dificulta la validación. Por lo tanto, esta tecnología se encuentra todavía en una fase de investigación y ensayo combinados. Se utiliza más como una herramienta para reducir el número de pruebas que para sustituirlas.

Resumen

Este capítulo abordó la superficie caliente de cohetes y vehículos hipersónicos. Los materiales de este lugar deben soportar el calor, la oxidación, las fuerzas y un tiempo prolongado al mismo tiempo. Los metales no son suficientes, por lo que se necesitan materiales cerámicos. Los compuestos de fibra de carbono y las cerámicas de ultraalta temperatura son los candidatos. Conocer de antemano la vida útil de estos materiales es la clave del desarrollo.

Primero vimos cómo se dañan los materiales. Los compuestos de fibra de carbono se deterioran cuando el oxígeno entra a lo largo de las grietas. El carburo de silicio gana o pierde su película protectora según la temperatura y la presión de oxígeno. En las cerámicas de ultraalta temperatura, la superficie se desprende debido a la evaporación y al cambio de fase de la película de óxido. Este proceso de daño es el objeto de la predicción de la vida útil. Como el daño suele comenzar en silencio dentro del material, hay que leer el interior y no solo la superficie.

A continuación, vimos cómo la inteligencia artificial aborda este daño. Convierte los granos cristalinos en puntos y los límites en líneas para crear un grafo. Al añadir el tiempo a esto, sigue el proceso de acumulación del daño. Al incorporar las leyes de la física en el aprendizaje, prevé incluso la región de ultraalta temperatura donde escasean los datos. La regla de que el daño no retrocede es un ejemplo representativo.

Luego, vimos el diseño inverso. Primero se define el rendimiento deseado y se busca su estructura en orden inverso. El modelo de difusión moldea la estructura a partir de un ruido difuso. Se añade la propiedad de no verse afectada por la dirección para ajustarse a la física. Determina conjuntamente el ángulo de las fibras y el espesor del recubrimiento de la interfaz. El modelo de difusión no da una sola respuesta, sino varios candidatos. Entonces, el ser humano selecciona entre esos candidatos las estructuras fáciles de fabricar.

Por último, vimos la integración multiescala del flujo de calor. Hay que conectar el mundo de las vibraciones atómicas con el del componente completo. La dinámica molecular de Green-Kubo calcula la conducción térmica en el mundo diminuto. La red neuronal de operadores profundos traslada esos valores a la temperatura del componente. El método de multifidelidad combina datos de bajo costo y datos de alto costo. Gracias a esta integración, es posible prever con rapidez la temperatura del componente. Esta predicción rápida es útil tanto en la etapa de diseño como en el lugar de pruebas.

Todas estas herramientas tienen un punto en común. Es el hecho de que la inteligencia artificial no sustituye a las pruebas. La inteligencia artificial reduce los candidatos y adelanta las predicciones. La verificación final queda en manos de las pruebas de túnel de viento, chorro de arco y combustión. Solo cuando la inteligencia artificial y las pruebas reales van de la mano, el desarrollo se acelera verdaderamente.

Otro punto en común es la unión con la física. La inteligencia artificial de este capítulo no aprende únicamente de los datos. Incorpora las leyes de la termodinámica y el equilibrio de fuerzas en su aprendizaje. Por eso puede abordar la región de ultraalta temperatura donde escasean los datos. Se trata de un obstáculo difícil de superar solo con el aprendizaje a partir de datos. El encuentro entre la física y la inteligencia artificial es la clave esencial de este capítulo.

Los métodos de este capítulo están conectados entre sí. La predicción del flujo de calor proporciona la temperatura superficial. La temperatura superficial determina la velocidad de oxidación y fluencia. La oxidación y la fluencia conducen a la predicción de la vida útil. La predicción de la vida útil se convierte en el objetivo del diseño inverso. Están, por así decirlo, unidos en una sola cadena.

También quedan muchas tareas pendientes. Los datos reales a temperaturas ultraaltas siguen siendo insuficientes. Las afirmaciones sobre el rendimiento de los modelos de diseño inverso requieren una validación con piezas reales. La forma de integrar las reglas físicas también debe perfeccionarse aún más. A medida que se resuelvan estos desafíos, el diseño de los sistemas de protección térmica de próxima generación irá avanzando. Un sistema de protección térmica se refiere a toda la estructura que protege al vehículo de su superficie caliente.

Corea también está dando pasos en este campo. El cohete Nuri y los vehículos de lanzamiento de próxima generación manejan componentes calientes. Los cohetes reutilizables deben usar las piezas varias veces. Por ello, la predicción de la vida útil se vuelve tanto más importante. Estos métodos que utilizan conjuntamente la inteligencia artificial y la física forman esa base. Los ojos que leen el mundo diminuto de los materiales protegen la seguridad de los grandes vuelos.

Fuentes y referencias relacionadas

Binner, J., Porter, M., Baker, B., Zou, J., Venkatachalam, V., Rubio Diaz, V., D'Angio, A., Ramanujam, P., Zhang, T., & Murthy, T. S. R. C. (2020). Selection, processing, properties and applications of ultra-high temperature ceramic matrix composites. International Materials Reviews, 65(7), 389-444. https://doi.org/10.1080/09506608.2019.1652006

Fahrenholtz, W. G., & Hilmas, G. E. (2017). Ultra-high temperature ceramics: Materials for extreme environments. Scripta Materialia, 129, 94-99. https://doi.org/10.1016/j.scriptamat.2016.10.018

DiCarlo, J. A., Yun, H. M., Morscher, G. N., & Bhatt, R. T. (2004). SiC/SiC composites for 1200°C and above. NASA Technical Reports Server, 20040191405. https://ntrs.nasa.gov/citations/20040191405

NASA. (2008). Ceramic Matrix Composite (CMC) thermal protection systems and hot structures for hypersonic vehicles. NASA Technical Reports Server, 20080017096. https://ntrs.nasa.gov/citations/20080017096

Wu, T., Wang, Y., Qi, W., Luo, X., Luo, P., Gao, X., & Song, Y. (2026). Predictive constitutive modelling of oxidation-induced degradation in 2.5D woven C/SiC composites. Materials, 19(2), 307. https://doi.org/10.3390/ma19020307

Qu, N., Zhou, W., Zhang, W., Liu, Y., Zheng, L., Cao, D., Tan, M., Zhu, J., & Zhang, X. (2026). Research progress and prospects of ultra-high-temperature ceramics: Experimentation, multiscale simulation and data-driven design. Nanomaterials, 16(11), 693. https://doi.org/10.3390/nano16110693

Baier, L., Frieß, M., Hensch, N., & Leisner, V. (2024). Development of ultra-high temperature ceramic matrix composites for hypersonic applications via reactive melt infiltration and mechanical testing under high temperature. CEAS Space Journal, 17(4), 635-644. https://doi.org/10.1007/s12567-024-00562-y

Peters, A. B., Zhang, D., Chen, S., Ott, C., Oses, C., Curtarolo, S., McCue, I., Pollock, T. M., & Eswarappa Prameela, S. (2024). Materials design for hypersonics. Nature Communications, 15, 3328. https://doi.org/10.1038/s41467-024-46753-3

Bianco, G., Nisar, A., Zhang, C., Boesl, B., & Agarwal, A. (2022). Predicting oxidation damage in ultra high-temperature borides: A machine learning approach. Ceramics International, 48(20), 29763-29769. https://doi.org/10.1016/j.ceramint.2022.06.236

Meng, H., Liu, Y., Yu, H., Zhuang, L., & Chu, Y. (2025). Machine-learning-potential-driven prediction of high-entropy ceramics with ultra-high melting points. Cell Reports Physical Science, 6, 102449. https://doi.org/10.1016/j.xcrp.2025.102449

Liu, Y., Meng, H., Zhu, Z., Yu, H., Zhuang, L., & Chu, Y. (2024). Exploring mechanical and thermal properties of high-entropy ceramics via general machine learning potentials. arXiv, 2406.08243. Preimpresión sin revisión por pares. https://arxiv.org/abs/2406.08243

Thomas, A., Durmaz, A. R., Alam, M., Gumbsch, P., Sack, H., & Eberl, C. (2023). Materials fatigue prediction using graph neural networks on microstructure representations. Scientific Reports, 13, 12562. https://doi.org/10.1038/s41598-023-39400-2

Zheng, B., Wen, J., Choy, Y.-S., & Fei, C. (2026). Physics-constrained EBSD image inpainting via adversarial graph learning: Bridging crystallographic rules and multimodal deep learning. Expert Systems with Applications, 298, 129667. https://doi.org/10.1016/j.eswa.2025.129667

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Chen, L., Li, Y., Ma, Y., Gao, L., & Yu, L. (2025). Multiscale graph equivariant diffusion model for 3D molecule design. Science Advances, 11(16), eadv0778. https://doi.org/10.1126/sciadv.adv0778

Xu, P., Wu, Y., Zarei, A., Ahmed, S., Pilla, S., Li, G., & Luo, F. (2025). Physically constrained 3D diffusion for inverse design of fiber-reinforced polymer composite materials. Composites Part B: Engineering, 303, 112515. https://doi.org/10.1016/j.compositesb.2025.112515

Dai, B., Wang, H., & Gui, J. (2026). Property-informed diffusion-based text-to-microstructure generation. arXiv, 2606.08150. Preimpresión sin revisión por pares. https://arxiv.org/abs/2606.08150

Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence, 3, 218-229. https://doi.org/10.1038/s42256-021-00302-5

Lu, L., Pestourie, R., Johnson, S. G., & Romano, G. (2022). Multifidelity deep neural operators for efficient learning of partial differential equations with application to fast inverse design of nanoscale heat transport. Physical Review Research, 4, 023210. https://doi.org/10.1103/PhysRevResearch.4.023210

Green, M. S. (1954). Markoff random processes and the statistical mechanics of time-dependent phenomena. Journal of Chemical Physics, 22, 398-413. https://doi.org/10.1063/1.1740082

Kubo, R. (1957). Statistical-mechanical theory of irreversible processes. Journal of the Physical Society of Japan, 12, 570-586. https://doi.org/10.1143/JPSJ.12.570

Verdi, C., Karsai, F., Liu, P., Jinnouchi, R., & Kresse, G. (2021). Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials. npj Computational Materials, 7, 156. https://doi.org/10.1038/s41524-021-00630-5

Zhang, X., Liu, N., Guo, Q., Shi, G., & Wang, Y. (2026). Beyond Boltzmann transport: Green-Kubo prediction of lattice thermal conductivity with machine-learned potentials. Journal of Chemical Physics, 164(7), 074101. https://doi.org/10.1063/5.0297462

Mandal, S., Srivastava, A., Das, T., Singh, A. K., & Maiti, P. K. (2025). Probing lattice anharmonicity and thermal transport in ultralow-κ materials using machine learning interatomic potentials. Small, 22(8), 2513476. https://doi.org/10.1002/smll.202513476

Arabha, S., Shokri Aghbolagh, Z., Ghorbani, K., Hatam-Lee, S. M., & Rajabpour, A. (2021). Recent advances in lattice thermal conductivity calculation using machine-learning interatomic potentials. Journal of Applied Physics, 130(21), 210903. https://doi.org/10.1063/5.0069443

Capítulo 6. Modelado del estrés térmico y reconocimiento in situ (In-situ) de defectos en procesos de fabricación aditiva de metales (LPBF/DED)

Introducción

Pensemos en hornear un pastel en la cocina. Se vierte la masa en un molde, se mete al horno y sale una sola pieza. Este método de verter material en un molde para fabricar algo de una sola vez es una técnica de manufactura tradicional. Las piezas de cohetes también se fabricaron de esta manera durante mucho tiempo. Se desbastaba un gran bloque de metal o se fabricaban varias piezas por separado y luego se unían.

Sin embargo, en el interior de la cámara de combustión y la tobera del motor cohete pasan cientos de canales de refrigeración muy estrechos. Para proteger las paredes de las llamas calientes, es necesario hacer circular combustible frío por esos canales. Una estructura interna tan compleja es difícil de fabricar mediante mecanizado. Si se unen varias piezas, las juntas se convierten en puntos débiles.

La fabricación aditiva de metales abre un camino diferente. Este método fabrica piezas acumulando metal capa por capa. Al igual que una impresora 3D extruye y deposita plástico línea por línea, funde y acumula polvo metálico o hilo metálico. Incluso los canales internos complejos y huecos pueden fabricarse en un solo proceso. Al no tener juntas, también disminuyen los puntos débiles.

Este método otorga libertad, pero también genera nuevas preocupaciones. Mientras se funde y acumula el metal capa por capa, pueden formarse poros microscópicos. Al calentarse y enfriarse en cada capa, la pieza también puede deformarse. La forma de los cristales puede variar según la dirección, lo que hace que la resistencia sea desigual. Estos defectos no se aprecian fácilmente a simple vista. Sin embargo, en los componentes de cohetes, un solo defecto pequeño puede provocar un gran accidente.

En este capítulo, examinamos cómo la inteligencia artificial ayuda a resolver este problema. La inteligencia artificial realiza tres tareas: detecta rápidamente defectos durante la fabricación; predice con rapidez el estrés térmico y la deformación en lugar de los cálculos lentos; y ayuda a ajustar las condiciones del proceso en tiempo real. Al seguir estos tres puntos en orden, se vislumbra el esquema de una fábrica que produce de manera autónoma piezas de cohetes casi perfectas.

En la base de las tecnologías que aborda este capítulo subyace una idea: la inteligencia artificial por sí sola es insuficiente, y la física por sí sola tampoco basta. Una inteligencia artificial que solo ha aprendido de los datos es vulnerable al ruido. Si solo se utilizan cálculos físicos, el proceso es demasiado lento. Al combinar ambas, se complementan las debilidades mutuas. Esta integración es la esencia del aprendizaje automático informado por la física (PIML), y las diversas tecnologías de este capítulo son ejemplo de ello.

6.1 Dos métodos para fabricar piezas de cohetes capa por capa

Contenido de esta sección

Existen dos métodos principales para fabricar componentes de motores cohete mediante la adición de metal. Uno consiste en disparar un láser sobre una fina capa de polvo metálico. El otro consiste en introducir polvo o hilo metálico en el punto focal del láser para fundirlo y acumularlo. Ambos métodos difieren en el tamaño y la precisión de las piezas que pueden fabricar.

Asimismo, esta sección presenta tres metales utilizados en cohetes. Los tres metales tienen propiedades diferentes y requieren precauciones distintas durante su fabricación. Es necesario saber qué ocurre en su interior mientras el metal se funde y se solidifica. Solo así se puede comprender por qué son necesarias las tecnologías de inteligencia artificial que se describen más adelante. Esta sección sienta las bases al presentar las propiedades de estos tres metales.

6.1.1 Sistemas de propulsión de cohetes de próxima generación y fabricación aditiva (Additive Manufacturing) de metales

Los motores cohete operan en condiciones extremas. La presión dentro de la cámara de combustión alcanza cientos de veces la presión de los neumáticos de un automóvil. La temperatura donde se quema el combustible supera los 3000 grados. Por el contrario, en los canales de refrigeración fluye combustible líquido a casi 200 grados bajo cero. Una sola pieza debe soportar simultáneamente este calor y este frío extremos.

Fabricar este tipo de piezas con métodos tradicionales requiere mucho trabajo manual. Se mecaniza cada canal de refrigeración uno por uno y se unen varias piezas mediante un método llamado soldadura fuerte (brazing). La soldadura fuerte es un método de unión en el que se funde e introduce otro metal entre las piezas metálicas. Cuantas más piezas haya, más juntas se generan y la fabricación tarda varios meses. Con el tiempo, estas juntas se convierten fácilmente en lugares propensos a fugas o grietas.

La fabricación aditiva de metales reduce drásticamente este proceso. La fusión por lecho de polvo láser (Laser Powder Bed Fusion; LPBF) extiende una capa fina de polvo metálico y dispara un láser sobre ella. Funde y une solo las zonas necesarias, vuelve a extender otra capa fina de polvo y funde de nuevo. Al repetir este proceso miles de veces, se completa la pieza. Se adapta muy bien a componentes pequeños y de alta precisión, como los revestimientos de cámaras de combustión con canales densos.

La deposición bajo energía dirigida (Directed Energy Deposition; DED) funciona de una manera un poco diferente. Introduce polvo o hilo metálico en el punto focal del láser mientras lo funde y deposita. Es similar a aplicar capas sucesivas de pintura con un pincel. La velocidad de deposición es rápida y resulta adecuada para piezas grandes. Este método se utiliza para fabricar faldones de tobera con diámetros superiores a 1 metro.

Ambos métodos tienen sus ventajas y desventajas. La fusión por lecho de polvo láser es muy precisa pero lenta, debido a que requiere extender miles de capas finas. Además, el tamaño de la pieza que se puede fabricar está limitado por las dimensiones de la cámara. Por el contrario, la deposición bajo energía dirigida es rápida y permite fabricar piezas de gran tamaño, pero la superficie resultante es rugosa. Requiere un trabajo adicional de mecanizado y acabado posterior a la fabricación.

La calidad del polvo también determina el resultado. Las partículas de polvo deben ser uniformemente esféricas para extenderse bien en capas finas. Si el tamaño de las partículas es irregular, se forman huecos. Si el polvo absorbe humedad, se generan burbujas de gas durante la fabricación. Por ello, el polvo se almacena y gestiona en condiciones secas. Una buena pieza comienza con un buen polvo.

Ambos métodos poseen la capacidad de fabricar piezas como un solo bloque integral. Permiten conformar sin juntas revestimientos de cámaras de combustión, cabezales de inyectores, impulsores de turbobombas y grandes toberas. El proyecto de tecnologías de propulsión de cohetes (RAMPT) de la Administración Nacional de Aeronáutica y el Espacio (NASA) ha estado fabricando toberas a gran escala con este método. Según datos públicos, se fabricó una tobera con canales de refrigeración de aproximadamente 1,5 metros de diámetro y 1,8 metros de longitud en solo 90 días. Esta tobera equivale al 65 por ciento del tamaño de la tobera de un motor grande para vehículos de lanzamiento espacial.

El mismo proyecto también informa sobre logros en pruebas. En pruebas de desarrollo que integraron múltiples inyectores, cámaras de combustión y toberas, los componentes soportaron combustiones de larga duración. El tiempo acumulado de combustión superó con creces los 10 000 segundos y se repitieron cientos de encendidos. Esto significa que las piezas fabricadas mediante adición capa por capa pueden soportar la combustión real de un cohete. Este logro demuestra que la fabricación aditiva está pasando del laboratorio a los componentes de vuelo.

El beneficio que ofrece este método radica en el tiempo y en la cantidad de piezas. El método tradicional fabricaba y unía cientos de piezas por separado. La fabricación aditiva reduce esos cientos de piezas a solo unas pocas. Al reducirse las piezas, disminuyen las juntas y también el tiempo de fabricación. Cuando se desea modificar el diseño, se puede volver a fabricar rápidamente. Esta rapidez acelera el ritmo de desarrollo de nuevos motores.

6.1.2 Propiedades metalúrgicas y comportamiento de transformación de fase por aleación

Los metales utilizados en los componentes de cohetes tienen misiones diferentes. Esta sección examina tres metales representativos: Inconel 718, GRCop-42 y C-103. La metalurgia es la disciplina que estudia la estructura interna de los metales y sus transformaciones. Incluso el mismo metal adquiere una estructura interna diferente según cómo se enfríe.

Los tres metales desempeñan funciones distintas. El Inconel 718 se encarga de las piezas resistentes que soportan grandes esfuerzos mecánicos. El GRCop-42 se encarga de las paredes de la cámara de combustión, que deben evacuar el calor con rapidez. El C-103 se encarga del extremo de la tobera, que entra en contacto directo con las llamas. Dentro de un solo motor, los tres metales se reparten las tareas. Por ello, se requiere la metalurgia adecuada para procesar con éxito cada metal.

El Inconel 718 es una superaleación a base de níquel. Una superaleación es una aleación diseñada para mantener su resistencia incluso a altas temperaturas. Este metal se utiliza ampliamente en turbobombas, válvulas y tuberías de alta presión. Es resistente, se suelda bien y soporta elevadas temperaturas. Por ello, es adecuado para los componentes del motor sometidos a grandes cargas mecánicas.

El problema surge con el enfriamiento rápido. El metal fundido por el láser se solidifica en un abrir y cerrar de ojos. En este momento, elementos como el niobio y el molibdeno se concentran en los límites de grano. Esta concentración se denomina microsegregación. Si la segregación es severa, en ese lugar se forma un compuesto frágil llamado fase de Laves (Laves phase). Este compuesto arrebata los elementos que confieren resistencia al metal, reduciendo así su solidez.

Por ello, el Inconel 718 requiere un tratamiento térmico posterior a la fabricación. La pieza se vuelve a calentar a alta temperatura para disolver y eliminar los compuestos frágiles. A continuación, se reduce la temperatura para inducir la formación de partículas de refuerzo muy pequeñas y uniformes. Solo tras este proceso se obtiene una pieza resistente a la fatiga. La fatiga es el fenómeno por el cual un material se agrieta progresivamente al someterse a cargas repetidas.

Estas partículas de refuerzo tienen un nombre: son una fase muy pequeña formada por níquel y niobio, denominada gamma doble prima. En su fórmula química, se une un átomo de niobio por cada tres de níquel. Esta fase es la responsable de la mayor parte de la resistencia del Inconel 718. Es una fase diferente de la gamma prima, común en otras superaleaciones de níquel. Por tanto, controlar adecuadamente el niobio es la clave de este metal.

El GRCop-42 es una aleación de cobre desarrollada por la NASA para cámaras de combustión de cohetes. El cobre es un excelente conductor térmico. Esta propiedad es necesaria para disipar rápidamente el

intenso calor de las paredes de la cámara de combustión. Sin embargo, el cobre puro se ablanda al calentarse. Por ello, se añaden pequeñas cantidades de cromo y niobio al cobre. Estos dos elementos forman partículas microscópicas que refuerzan y sostienen la matriz de cobre.

Existen dificultades al fabricar GRCop-42 con láser. El cobre refleja intensamente la luz, rebotando gran parte de la radiación láser como si fuera un espejo. Por tanto, para fundirlo adecuadamente se requiere un láser de alta potencia o un láser verde. Tras la fabricación, se somete a un proceso de prensado a alta presión y temperatura para eliminar los poros internos. Al superar este proceso, mantiene una excelente conductividad térmica y conserva su resistencia incluso en entornos con temperaturas superiores a 800 grados.

Examinemos por qué la pared de la cámara de combustión debe evacuar el calor con tanta eficacia. En un lado de la pared inciden llamas a más de 3000 grados. En el otro lado fluye combustible frío. Esta delgada pared separa ambas temperaturas. Si el calor no se disipa con rapidez, la pared se funde o se agrieta. Por esta razón, se emplean aleaciones de cobre con alta conductividad térmica como material de pared.

Esta aleación se combina con los canales de refrigeración. Dentro de la pared se forman canales delgados dispuestos de manera densa. El combustible fluye por esos canales y evacua el calor de la pared. Cuanto más complejos sean los canales, más uniformemente se disipa el calor. La fabricación aditiva tiene la capacidad de crear estos canales complejos en un solo paso. El material y el método de fabricación se complementan mutuamente.

El C-103 es una aleación para altas temperaturas a base de niobio. El punto de fusión del niobio supera con creces los 2000 grados. Por ello, se utiliza en los extremos de toberas en contacto directo con las llamas o en propulsores de control de actitud. A esta aleación se le añaden pequeñas cantidades de hafnio, titanio y circonio para ajustar sus propiedades. Para fabricar toberas de chapa fina de gran tamaño se utiliza el método DED.

Al fabricar C-103, se debe tener especial cuidado con el aire. El niobio fundido absorbe con rapidez el oxígeno y el nitrógeno del aire. Si el nivel de oxígeno supera cierta cantidad, el metal se vuelve quebradizo y adquiere una fragilidad similar a la del vidrio, fracturándose con facilidad. Por esta razón, la cámara de fabricación se llena con un gas de protección como el argón para mantener el oxígeno en niveles extremadamente bajos.

La oxidación también representa un problema importante durante su uso. El oxígeno presente en los gases de escape calientes ataca continuamente al niobio. Por ello, se aplica en la superficie un recubrimiento protector de siliciuro a base de silicio. Esta capa impide que el oxígeno entre en contacto con el metal. Se requiere tanto el control atmosférico durante la fabricación como el recubrimiento protector durante el servicio. Estos dos aspectos determinan el éxito o fracaso de las piezas de C-103.

Las investigaciones recientes profundizan en los procesos de fabricación de este metal. Un estudio de 2025 fabricó láminas delgadas de C-103 mediante deposición por hilo láser y examinó sus propiedades. Otro estudio trazó mapas de defectos utilizando modelos de baño de fusión que resuelven múltiples fenómenos físicos simultáneamente. Esto permite predecir y cartografiar de antemano bajo qué condiciones se generan defectos. Con este tipo de mapas, es posible seleccionar y aplicar condiciones óptimas, encontrando la forma de evitar defectos antes de la fabricación.

6.2 Mecanismos de reconocimiento de defectos in situ (In-situ) y modelos de imágenes informados por la física

Contenido de esta sección

In situ significa observar en el mismo lugar y momento en que se está fabricando. Si se detecta un defecto después de terminar la pieza, ya es demasiado tarde. Si se descubren los defectos durante la fabricación, es posible corregirlos en el acto o detener el proceso. Esta sección aborda los métodos para observar y detectar defectos durante la fabricación.

En primer lugar, se analiza por qué y cómo se generan los defectos. A continuación, se presentan los dispositivos que observan el pequeño pozo formado por el láser. Por último, se describe cómo la inteligencia artificial interpreta las imágenes observadas. Aquí surge el método de combinar las leyes de la física con la inteligencia artificial, lo cual constituye un ejemplo del aprendizaje automático informado por la física que articula todo este libro.

6.2.1 Mecanismos termofluidodinámicos de formación de defectos en el baño de fusión

Cuando el láser funde el polvo metálico, se forma un pequeño charco. Este pozo se denomina baño de fusión (Melt pool). Su diámetro es tan reducido como el grosor de unos pocos cabellos. A medida que el láser avanza, este pozo también se desplaza. La calidad de la pieza depende en gran medida del estado de este pequeño baño de fusión.

Dentro de este pozo, el metal se mueve sin cesar. La tensión superficial varía entre las zonas más calientes y las menos calientes. Esa diferencia empuja el metal fundido en diversas direcciones. Este flujo se denomina convección de Marangoni. Se asemeja al remolino que se forma en la superficie de una olla con caldo. Este remolino puede tanto generar defectos como eliminarlos.

Un tipo de defecto es la porosidad por ojo de cerradura (Keyhole porosity). Se produce cuando el láser es demasiado potente o avanza demasiado despacio. La intensa luz penetra profundamente en el metal y crea una cavidad estrecha y profunda, similar al ojo de una cerradura. Esta cavidad es inestable y colapsa de forma repentina. Al colapsar, el gas de su interior queda atrapado en el metal y forma burbujas redondeadas. Estas burbujas atrapadas constituyen la porosidad por ojo de cerradura.

El ojo de cerradura no es del todo perjudicial. Con una profundidad adecuada, funde el metal a suficiente profundidad y lo une con firmeza. El problema surge cuando es demasiado profundo e inestable. Por ello, en el proceso se procura mantener una profundidad adecuada del ojo de cerradura. Se busca una ventana operativa óptima sin aumentar ni reducir en exceso la potencia del láser. Hallar esta ventana adecuada es la clave del control del proceso.

Otro defecto es la falta de fusión (Lack-of-Fusion; LoF). A diferencia del ojo de cerradura, se produce cuando el láser es débil o la separación entre pasadas es demasiado amplia. Las pasadas adyacentes del láser o la capa inferior no se funden ni se unen adecuadamente entre sí. Entre ellas queda un espacio vacío. Este hueco tiene bordes afilados y contiene polvo sin fundir en su interior.

Ambos defectos presentan una naturaleza de riesgo distinta. La porosidad por ojo de cerradura es redondeada, por lo que resulta relativamente menos peligrosa. La falta de fusión presenta bordes afilados, lo que concentra las tensiones en sus extremos. Responde al mismo principio por el cual la punta de un fragmento de vidrio puntiagudo se fractura con facilidad. Por lo tanto, la falta de fusión tiende a convertirse en el punto de inicio de grietas por fatiga. Dado que los componentes de cohetes soportan innumerables vibraciones, este defecto representa un riesgo considerable.

Las investigaciones recientes clasifican los defectos de manera más detallada. Un estudio de 2026 dividió los poros en tres categorías: sin poros, poros menores de 15 micrómetros y poros mayores que ese tamaño. Estos 15 micrómetros corresponden al límite de clasificación fijado en dicha investigación. En ese estudio, los poros grandes por ojo de cerradura resultaron más perjudiciales. Los poros pequeños de gas fueron menos dañinos bajo esas condiciones. Dicha investigación entrenó esta clasificación combinando señales de fotodiodos e imágenes precisas de rayos X.

Sin embargo, este umbral no debe adoptarse de inmediato como criterio de aceptación. El peligro de un poro no se determina únicamente por su tamaño. Su ubicación, forma, las fuerzas aplicadas, el material y la vida útil requerida influyen de manera conjunta. Incluso con el mismo tamaño, resulta más peligroso si se encuentra cerca de la superficie. Por ello, es necesario distinguir entre los límites de clasificación y los criterios de seguridad estructural.

Existe una razón para desglosar los defectos de forma tan minuciosa. Si todos los poros se trataran por igual, se descartarían piezas en perfecto estado considerándolas defectuosas. Por el contrario, si se ignoran todos, se pasarían por alto poros peligrosos. Es necesario distinguir el tipo y el tamaño de los defectos para responder de manera adecuada. La inteligencia artificial se encarga de realizar esta distinción sutil.

Los defectos dejan señales precursoras visibles. Justo antes de que se forme un ojo de cerradura, la forma del baño de fusión oscila. Aparecen patrones en los que la poza se vuelve más profunda y estrecha. El ojo humano no puede seguir estos cambios tan rápidos. Cámaras capaces de capturar decenas de miles de fotogramas por segundo y la inteligencia artificial leen estas señales precursoras. El objetivo es detectar la anomalía antes de que se produzca el defecto.

Las salpicaduras de polvo también están vinculadas a los defectos. Las salpicaduras son un fenómeno en el que el metal fundido es proyectado cuando el láser impacta con fuerza. Las gotas proyectadas caen sobre el polvo que aún no se ha fundido. Cuando la siguiente pasada del láser incide en esa zona, no logra fundirse adecuadamente. Por lo tanto, una alta frecuencia de salpicaduras incrementa la falta de fusión. El grado de salpicadura también actúa como una señal temprana que advierte de posibles defectos.

6.2.2 Ojos para medir la temperatura del baño de fusión y líneas de base físicas

Para observar el baño de fusión se requieren ojos especiales. La poza es sumamente brillante y cambia con gran rapidez. La pirometría (Pyrometry) es un método para medir la temperatura a partir de la luz emitida por un objeto caliente. Se basa en el principio de que los objetos emiten una luz diferente a medida que aumenta su temperatura. Este método mide la temperatura sin entrar en contacto con el objeto.

La pirometría coaxial (Coaxial) capta la luz en la misma dirección por la que pasa el láser. Es como si el láser y la cámara compartieran la misma lente. Por ello, sin importar por dónde se desplace el láser, siempre es posible observar la poza de frente. Una cámara de alta velocidad que captura decenas de miles de fotogramas por segundo registra las variaciones de brillo y forma de la poza. Al comparar la luz en dos longitudes de onda distintas, se obtiene una temperatura menos afectada por las condiciones de la superficie.

Este método de dos longitudes de onda tampoco es perfecto. La emisividad se refiere al grado en que un objeto emite radiación luminosa. El método asume que la relación de emisividad entre las dos longitudes de onda es constante. Si esta suposición no se cumple, se generan errores en la temperatura medida. Los vapores metálicos desprendidos o una lente sucia también alteran las lecturas. Por ello, los valores medidos deben calibrarse periódicamente.

La cámara por sí sola no es suficiente. Debe existir una referencia con la cual contrastar qué es normal y qué es anómalo. Como una de esas referencias, puede citarse el modelo de conducción térmica de Rosenthal (Rosenthal). Se trata de una teoría física clásica formulada en 1946. Es una fórmula analítica resuelta a mano que describe la distribución de temperatura cuando una fuente de calor móvil calienta el metal.

El modelo de Rosenthal es solo un esquema aproximado. Trata el metal como un medio semiinfinito y considera la fuente de calor como un punto puntual. Asume que las propiedades del metal son constantes independientemente de la temperatura. También ignora el calor latente que entra y sale durante la fusión y la ebullición. Asimismo, no contempla el flujo del metal fundido ni la evaporación.

Por ello, esta fórmula solo proporciona una temperatura de referencia aproximada. La rápida dinámica del colapso del ojo de cerradura no puede explicarse mediante esta ecuación. Si se requiere una referencia más precisa, se recurre al análisis por elementos finitos o a la dinámica de fluidos computacional. Los casos en que se utilizan directamente los valores de Rosenthal como línea de base para el monitoreo en tiempo real aún son poco frecuentes. Aquí se presenta como un ejemplo del concepto de establecer una referencia física.

La inteligencia artificial puede recibir estas discrepancias como datos de entrada. Aprende la diferencia entre el campo de temperatura real y la referencia física para predecir defectos. Disponer de una referencia física aligera la carga de aprendizaje de la inteligencia artificial. Incluso con pocos datos, la física actúa como guía. En consecuencia, la referencia física reduce la carga de entrenamiento de la inteligencia artificial.

¿Por qué es tan crucial el tiempo real? El láser se desplaza a varios metros por segundo. El tiempo necesario para depositar una capa es muy breve. La zona donde se produjo un defecto pronto queda cubierta por la capa siguiente. Es imprescindible detectarlo antes de que quede cubierto para poder repararlo en ese mismo lugar. Por lo tanto, las decisiones deben tomarse con rapidez, en escalas de milisegundos.

Los dispositivos de observación también continúan evolucionando. Investigaciones recientes han propuesto un método para capturar simultáneamente dos longitudes de onda con una sola cámara. Esta técnica registra las variaciones de temperatura del baño de fusión a velocidades de hasta 100 000 fotogramas por segundo. Las imágenes térmicas precisas obtenidas de este modo sirven de base para discriminar los poros mediante diversos enfoques. Disponer de estos ojos precisos es esencial para identificar los poros de manera adecuada.

6.2.3 Principios de diseño de modelos de visión informados por la física

Ha llegado el momento de interpretar las imágenes observadas. Esta sección propone los principios de diseño de modelos de visión que incorporan la física. En primer lugar, se examinan las redes neuronales de visión ampliamente utilizadas. El transformador de visión (Vision Transformer) es una red neuronal que divide una imagen en pequeños parches y analiza sus relaciones. Corta la imagen como si fuera un mosaico y examina cómo se conecta cada fragmento con los demás. En las imágenes del baño de fusión, la zona brillante central, la cola que se extiende hacia atrás y el brillo del polvo salpicado son diferentes entre sí. Al interpretar estas relaciones, el modelo discrimina los defectos.

La fortaleza de esta red neuronal reside en su capacidad para vincular partes distantes entre sí. La temperatura en la parte frontal del baño de fusión y la cola posterior están interconectadas. El modo en que fluye el calor no puede comprenderse observando solo un fragmento aislado. El transformador de visión compara todos los parches simultáneamente. De este modo, interpreta el flujo térmico de toda la poza. Este amplio campo de visión contribuye a captar las señales precursoras de los defectos.

Este libro propone integrar aquí las leyes de la conducción térmica. Se busca que la inteligencia artificial verifique por sí misma si sus respuestas concuerdan con la física al generarlas. La función de pérdida (Loss function) es una puntuación que mide el grado de error de la inteligencia artificial. La propuesta consiste en incluir en esta puntuación el error de clasificación de defectos, el error del campo de temperatura y los residuos físicos. El residuo es un valor que indica el grado de discrepancia respecto a las ecuaciones.

Incorporar la física de este modo aporta ventajas significativas. Las imágenes de las cámaras siempre contienen ruido. Además, el polvo salpicado puede bloquear la luz. Un modelo entrenado exclusivamente con datos se desestabiliza fácilmente ante tales perturbaciones. En cambio, un modelo que también ha aprendido la física cuenta con el respaldo de las leyes de conducción térmica. Aunque la imagen se vuelva algo borrosa, mantiene juicios acordes con los principios físicos.

Las investigaciones experimentales también respaldan el potencial de las características físicas. Un estudio extrajo características basadas en la física a partir de señales de cámaras infrarrojas. Mediante estas características, identificó poros en componentes de Inconel 718. Esto se debe a que el propio patrón de variación de la temperatura contiene pistas sobre los defectos. Las características derivadas de la física son robustas frente al ruido y poseen un significado inequívoco. Por ello, también resulta más sencillo explicar los fundamentos de las decisiones tomadas.

La velocidad también es fundamental. Para emitir juicios durante la fabricación, las respuestas deben generarse en un abrir y cerrar de ojos. Un estudio de 2026 combinó redes neuronales preentrenadas con clasificadores tradicionales. Este método discriminó imágenes del baño de fusión de aleaciones de níquel con una precisión de aproximadamente 0,95. El tiempo empleado para evaluar una imagen fue de 1,15 milisegundos. Aunque se trata de un estudio preliminar que aún no ha pasado por la revisión por pares, demuestra la viabilidad de realizar inferencias lo bastante rápidas para su aplicación en entornos de planta.

El enfoque de incorporar la física requiere un ajuste de ponderaciones. Si se penaliza con excesiva fuerza el residuo físico, el modelo aprende menos de los datos. Si se siguen exclusivamente los datos, se desvía de las leyes físicas. Un buen modelo requiere equilibrar adecuadamente ambas fuerzas. Encontrar este punto de equilibrio constituye el ajuste fino del aprendizaje informado por la física. La ponderación óptima varía en función de la pieza y del equipo.

Conviene distinguir aquí entre lo validado y lo propuesto. La detección de poros mediante características infrarrojas y los modelos híbridos rápidos son resultados demostrados por investigaciones reales. Por su parte, la arquitectura que integra residuos de conducción térmica en el transformador de visión es la dirección conceptual planteada en este libro. Aún no es un método validado en componentes reales. Las decisiones tomadas por la inteligencia artificial deben verificarse nuevamente mediante secciones transversales y tomografía computarizada.

6.3 Reducción de la distorsión en toberas de gran tamaño

La falda de la tobera es un componente de gran tamaño con forma de campana que se acopla a la parte trasera del motor. Se fabrica apilando láminas delgadas en forma de cono redondeado. Estas piezas tan grandes y delgadas tienden a distorsionarse con facilidad durante la fabricación. Esta sección aborda las causas de dicha distorsión y las estrategias para prevenirla.

En primer lugar, se analiza por qué se acumulan esfuerzos en el interior de la pieza. A continuación, se presenta un modelo rápido de inteligencia artificial que sustituye a los cálculos lentos. Por último, se examina cómo regular el movimiento del láser con dicho modelo para mitigar la distorsión. Cuanto mayor es el componente, más complejo y crucial resulta el control de esta distorsión.

6.3.1 Acumulación de tensiones térmicas y distorsión por pandeo en la fabricación aditiva de láminas delgadas de gran tamaño

La tensión residual (Residual stress) es la fuerza interna de tracción y compresión que permanece en la pieza. Aunque no es visible a simple vista, actúa constantemente en su interior. Imaginemos cómo se genera esta fuerza. Cuando el láser calienta una línea de metal, esa zona se expande debido a la dilatación del metal calentado.

Cuando el láser pasa, esa línea vuelve a enfriarse. Al enfriarse, tiende a contraerse a su estado original. Sin embargo, la capa inferior ya se encuentra solidificada y restringe este movimiento. Al enfrentarse la fuerza que tiende a contraerse con la fuerza de restricción, permanecen tensiones en el interior. En la parte superior se acumulan esfuerzos de tracción y en la inferior, esfuerzos de compresión.

En una falda de tobera delgada y de gran tamaño, estas fuerzas causan graves problemas. Es similar a tirar de un solo extremo de una hoja de papel, lo que hace que se arrugue y se curve. La pieza se distorsiona en dirección radial o la lámina delgada se flexiona de manera repentina. Esta flexión abrupta se denomina pandeo (Buckling). En casos severos, la pieza puede llegar a desprenderse de la base de soporte.

El control resulta aún más difícil a medida que la pieza es mayor, ya que el calor se acumula de forma continua al alcanzar cientos de capas. El estado térmico difiere entre la deposición de las primeras capas y la de las capas posteriores. Aunque se apliquen los mismos parámetros de láser, los resultados varían. Por lo tanto, en componentes grandes es imprescindible monitorizar el estado térmico e ir ajustando las condiciones del proceso.

Las aleaciones de niobio como el C-103 presentan este problema de forma aún más crítica. Esto se debe a que su elevado punto de fusión requiere un aporte térmico considerable para fundirse. Un alto aporte térmico se traduce en mayores tensiones. Al fabricar toberas de gran tamaño con este metal, el control de la distorsión determina el éxito o el fracaso del proceso. Es necesario integrar el análisis térmico y el diseño de trayectorias desde el principio.

La distorsión es difícil de corregir una vez finalizada la fabricación. Si se intenta enderezar a la fuerza una pieza ya endurecida, se generan nuevas grietas. Por ello, es preferible evitar que las tensiones se acumulen durante la fabricación. Es análogo a prevenir una enfermedad antes de contraerla en lugar de curarla después. Esta prevención es la tarea encomendada a la predicción por inteligencia artificial y al control activo. Los modelos sustitutos que se presentarán más adelante facilitan dicha prevención.

6.3.2 Modelos sustitutos basados en termodinámica (Surrogate Model)

Para conocer con antelación las tensiones dentro de una pieza se requieren cálculos. El análisis por elementos finitos (Finite Element Analysis) es un método que divide la pieza en pequeñas subdivisiones para calcular temperaturas y fuerzas. Es sumamente preciso, pero en componentes grandes requiere mucho tiempo, ya que exige resolver millones de pequeños elementos uno por uno. Un solo análisis de una tobera de gran tamaño puede tardar varios días.

El modelo sustituto (Surrogate model) es el encargado de reemplazar estos cálculos lentos. Un modelo sustituto es un modelo rápido que aprende con antelación los resultados de cálculos lentos. Con frecuencia se utiliza el perceptrón multicapa (Multilayer Perceptron; MLP). Se trata de una red neuronal básica que transforma entradas en salidas al pasar por múltiples capas. Tras entrenarse con numerosos resultados de análisis de alta fidelidad, estima los resultados bajo nuevas condiciones.

¿Qué recibe este modelo y qué produce? Las entradas son la potencia del láser, la velocidad de desplazamiento, la trayectoria recorrida por el láser y el tiempo de espera entre capas. Las salidas son la

distribución de tensiones residuales en toda la pieza y la forma de la deformación final. Lo que a un análisis riguroso le llevaría varios días, este modelo lo estima en poco tiempo. Esto permite comparar diversas condiciones con rapidez.

La velocidad abre nuevos caminos. Si hay que esperar varios días cada vez que se cambia una sola condición, resulta difícil encontrar un buen diseño. Aunque sea una estimación, si es rápida permite filtrar miles de condiciones de antemano. Se seleccionan únicamente los candidatos prometedores para verificarlos de nuevo mediante análisis rigurosos. El modelo subrogado actúa como una red que reduce el amplio abanico de candidatos.

Para que un modelo subrogado aprenda bien, necesita una gran cantidad de datos de calidad. Se ejecutan previamente análisis rigurosos bajo diversas condiciones para recopilar datos. También se incorporan los valores medidos en piezas reales fabricadas. Cuanto más amplios y uniformes sean los datos, más precisa será la estimación. Ante condiciones desconocidas sobre las que no ha aprendido, la estimación se vuelve inestable. Por ello, el modelo subrogado solo se utiliza con confianza dentro del rango en el que ha sido entrenado.

Las investigaciones recientes han obtenido resultados en esta dirección. Un estudio de 2025 estimó con rapidez las tensiones residuales mediante aprendizaje automático y redujo las deformaciones. En dicho estudio se informó que la curvatura de probetas con forma de puente disminuyó notablemente. La diferencia entre las tensiones predichas y el análisis riguroso también fue pequeña. Esto significa que el modelo rápido se aproximó bastante a la precisión del análisis riguroso.

Al incorporar información física, el resultado mejora aún más. Otro estudio de 2025 introdujo leyes físicas en una red neuronal para predecir la deformación de piezas grandes casi en tiempo real. Utilizó una estructura llamada operador neuronal (Neural Operator) para aprender a adaptarse a diversas condiciones. Aunque este estudio es preliminar y aún no ha pasado por la revisión por pares, muestra una vía para lograr simultáneamente predicciones rápidas y coherencia física.

6.3.3 Optimización activa de la trayectoria de escaneo y del tiempo de enfriamiento entre capas

Ahora es el momento de modificar el proceso real basándose en las predicciones. La trayectoria de escaneo (Scan path) es la dirección y el orden en que se desplaza el láser. Si se hace pasar el láser continuamente en una sola dirección, el calor se concentra en un solo lado. Ese calor concentrado deja grandes tensiones. Por eso se cambia de dirección para dispersar el calor de manera uniforme.

Han surgido diversas trayectorias ingeniosas. Existe un método que divide una superficie amplia en zonas pequeñas para rellenarlas de forma alterna. También hay un método que rota ligeramente la dirección del láser en cada zona. Asimismo, se emplean trayectorias curvas que rellenan el espacio de manera densa a medida que avanzan. El objetivo es uno solo: dispersar el flujo de calor uniformemente en todas direcciones para evitar que se concentre en un solo lado.

El tiempo de espera entre capas también es fundamental. Si se deposita la siguiente capa cuando la anterior aún está caliente, el calor se acumula capa sobre capa. El calor superpuesto provoca grandes deformaciones. Por eso se espera a que la capa anterior se enfríe hasta cierto punto. A este tiempo de espera se le llama tiempo de enfriamiento entre capas.

La optimización activa (Active optimization) es un método que modifica estas condiciones durante el proceso. La pirometría indica cuán caliente está la pieza en ese momento. El modelo subrogado predice rápidamente el riesgo de tensiones en la siguiente trayectoria. A partir de ambas informaciones, se ajustan en tiempo real la potencia del láser, la velocidad y el tiempo de espera. Es un proceso que traslada la dilatada experiencia humana a números y reglas.

Los métodos de ajuste también se vuelven más inteligentes. La optimización bayesiana (Bayesian Optimization) es un método para encontrar buenas condiciones con pocos intentos. Selecciona inteligentemente la siguiente condición a probar a partir de los resultados ya obtenidos. El algoritmo genético (Genetic Algorithm) es un método que combina buenas condiciones para encontrar otras aún mejores. Mediante estos métodos, el tiempo de espera se ajusta al estado térmico de la capa anterior.

Este ajuste se realiza de forma diferente en cada capa. Las capas inferiores se enfrían con rapidez porque la plataforma de soporte absorbe el calor. A medida que se asciende hacia las capas superiores, el calor puede disiparse con mayor dificultad. En tales casos, se incrementa el tiempo de espera en las capas superiores. No obstante, esto es solo un ejemplo. El grado de acumulación térmica varía según la forma de la pieza, la base de soporte, la superficie de la capa y la secuencia de deposición. Por ello, se determina el tiempo de espera observando el estado térmico de cada capa mediante sensores.

Cuando este método funciona adecuadamente, la deformación se reduce de forma notable. Las tensiones se distribuyen uniformemente y la tensión máxima disminuye. La pieza se mantiene dentro de las tolerancias permitidas. Sin embargo, este ajuste siempre debe realizarse dentro de una ventana de proceso segura. Es necesario considerar conjuntamente el retardo de los sensores, los errores del modelo y los límites de respuesta del equipo. Solo después de verificarlo primero con probetas de prueba se puede aplicar a piezas reales.

El concepto de gemelo digital también se nutre de aquí. Un gemelo digital consiste en tener un duplicado virtual de la pieza real dentro de la computadora. Los valores de los sensores del proceso real se introducen continuamente en este duplicado. El gemelo refleja lo que ocurre en el interior de la pieza en ese instante. Asimismo, muestra de antemano qué riesgos podrían presentarse en la siguiente capa. El modelo subrogado es el cerebro que ayuda a este gemelo a pensar con rapidez.

6.4 El control de la morfología cristalina

Cuando el metal se solidifica, en su interior crecen cristales. La forma de estos cristales determina la resistencia de la pieza. Si los cristales crecen alargados únicamente en una dirección, la resistencia varía según la orientación. Esta sección aborda los métodos para controlar la forma de los cristales según se desee.

Primero se examinan las dos formas de los cristales y sus diferencias. A continuación, se presenta un modelo generativo que traza la morfología cristalina a partir de las condiciones del proceso. Por último, se analiza el control mediante aprendizaje por refuerzo para regular de forma autónoma la forma de los cristales durante la fabricación. El mundo invisible de los cristales decide el destino de la pieza.

6.4.1 Transición de grano columnar a equiaxial (CET) y anisotropía mecánica

Cuando el metal se solidifica, los cristales crecen con dos morfologías. Los granos columnares (Columnar grain) son cristales alargados en una sola dirección, semejantes a columnas; se asemejan a varios troncos de árboles puestos en pie. Los granos equiaxiales (Equiaxed grain) son cristales que crecen con un tamaño similar en todas direcciones; su aspecto se parece a una capa uniforme de grava.

Esta diferencia determina las propiedades de la pieza. Si abundan los granos columnares, la pieza es fuerte en la dirección de las columnas pero débil transversalmente. La resistencia y la vida a fatiga varían según la dirección. A esta propiedad se le denomina anisotropía (Anisotropy). Responde al mismo principio de la madera, que se parte con facilidad a lo largo de la veta, pero difícilmente en sentido transversal.

En los componentes de cohetes, la anisotropía puede ser motivo de preocupación, ya que las piezas reciben fuerzas desde múltiples direcciones. Si solo son resistentes en una dirección, la dirección más

débil se fractura primero. Los granos equiaxiales alivian este problema: ofrecen una resistencia uniforme en todas direcciones y una alta resistencia al agrietamiento durante la solidificación. Por ello, los granos equiaxiales resultan ventajosos en componentes sometidos a cargas multidireccionales.

Sin embargo, los granos equiaxiales no siempre son la mejor opción. En piezas donde la fluencia lenta o la fatiga térmica representan un problema, los cristales columnares pueden ser preferibles. La fluencia es el fenómeno por el cual un material se deforma lentamente al soportar fuerzas de manera prolongada. En piezas donde el calor fluye en una sola dirección, los cristales con orientación preferencial también pueden ser ventajosos. La morfología cristalina deseada varía en función de las fuerzas y solicitaciones a las que se expone la pieza.

En la fabricación aditiva, los granos columnares se forman con facilidad. Esto ocurre porque el calor se disipa en la dirección de deposición del láser. Los cristales crecen en forma de columnas siguiendo la dirección en la que escapa el calor. Como resultado, los cristales se alinean en la dirección de acumulación de capas. A este conjunto de cristales cuya orientación se inclina hacia una dirección se le denomina textura cristalográfica. El objetivo de esta sección es modificar dicha orientación preferencial según sea necesario.

La forma de los cristales depende en gran medida de dos valores. El gradiente de temperatura (Temperature gradient; G) indica la inclinación térmica, es decir, con qué rapidez varía la temperatura según la posición. La velocidad de solidificación (Solidification rate; R) es la rapidez con la que se desplaza el límite entre el líquido y el sólido. Ambos valores desempeñan funciones distintas.

La relación entre ambos, G dividido por R, determina la morfología de los cristales. Si este cociente es elevado, tiende a ser columnar; si es bajo, se inclina hacia la forma equiaxial. El producto de ambos, G multiplicado por R, representa la velocidad de enfriamiento. Cuanto mayor sea este valor, más finos serán los cristales. Por tanto, es posible controlar la morfología y el tamaño por separado. Al modificar la potencia y la velocidad del láser, estos dos valores cambian conjuntamente.

No solo estos dos valores determinan la morfología. Las partículas microscópicas que actúan como semillas nucleadoras, el flujo del metal fundido y el grado de refundición también influyen. Además, se hereda la orientación de los cristales que ya han crecido debajo. Por ello, aunque los granos equiaxiales tienden a formarse mejor cuando el gradiente térmico es bajo y la solidificación es rápida, no siempre ocurre así. Es necesario considerar múltiples parámetros en conjunto para controlar adecuadamente la morfología cristalina.

Existe un punto en el que se produce el cambio de grano columnar a equiaxial. A esta transformación se la conoce como transición de columnar a equiaxial (Columnar-to-Equiaxed Transition; CET). Cuando el gradiente de temperatura y la velocidad de solidificación superan ciertas condiciones, se desencadena dicho cambio. Conocer estas condiciones permite aproximarse a la morfología cristalina deseada. Teorías clásicas han delineado este límite mediante fórmulas matemáticas.

Las investigaciones recientes refinan este límite con mayor precisión. Un estudio de 2025 estableció criterios para predecir la CET a partir de la composición de la aleación. Mediante aprendizaje automático, reveló qué elementos influyen de manera determinante en la forma de los cristales. Esto significa que es posible modificar la morfología cristalina con solo variar levemente la composición. Tratar conjuntamente el proceso y la composición amplía el margen de control cristalino.

El aprendizaje automático orientado al análisis de la composición sigue avanzando. A través de los datos, identifica qué elementos influyen significativamente en la morfología cristalina. Existen resultados que demuestran que la cantidad de elementos como el silicio o el carbono altera la forma de los cristales. Estas investigaciones van completando progresivamente el mapa del control cristalino.

6.4.2 Modelos generativos para trazar la morfología cristalina a partir de las condiciones del proceso

Se desea previsualizar la morfología cristalina a partir de las condiciones del proceso. Para ello, se necesita una herramienta que genere imágenes de la microestructura cuando se introducen dichas condiciones. Un ejemplo de tales herramientas son las redes generativas antagónicas condicionales (Conditional GAN). Esta red neuronal es un modelo generativo que produce la imagen deseada a partir de ciertas condiciones dadas. Aquí, las condiciones son valores como la potencia del láser, la velocidad, el gradiente de temperatura y la velocidad de solidificación. El uso de este modelo para el control microestructural se encuentra aún en fase de propuesta.

Este modelo aprende mediante la competencia entre dos partes. El generador (Generator) crea las imágenes de la microestructura. El discriminador (Discriminator) determina si la imagen es real o falsa. El generador se esfuerza por engañar al discriminador, y este se esfuerza por no ser engañado. Al repetir esta confrontación, el generador produce imágenes cada vez más realistas. Es una dinámica similar a la de un falsificador y un perito que perfeccionan mutuamente sus habilidades.

En este contexto, las imágenes reales corresponden a capturas de difracción de electrones retrodispersados (EBSD). Este método lee la orientación de los cristales mediante un microscopio electrónico. Muestra un mapa en el que se asignan colores a la orientación de cada cristal. El generador aprende a imitar estos mapas. Al introducir las condiciones, dibuja el mapa cristalográfico que resultaría de ellas.

Si este modelo aprende adecuadamente, resulta de gran utilidad. Permite al diseñador prever cómo cambiarán los cristales cuando se modifiquen las condiciones del proceso. Es posible estimarlo sin necesidad de fabricar físicamente la pieza y cortarla. Esto ahorra un tiempo y unos materiales considerables. Además, permite comparar múltiples condiciones en pantalla y seleccionar la más favorable.

Obtener un mapa cristalográfico real requiere un trabajo laborioso. Es preciso cortar la pieza y pulir finamente la superficie. Luego, se debe escanear punto por punto la orientación de los cristales con el microscopio electrónico. Analizar una superficie amplia toma mucho tiempo. El modelo generativo se convierte en un atajo que reduce notablemente este lento proceso. Por supuesto, dicho atajo debe verificarse de vez en cuando mediante mediciones reales.

También existen aspectos con los que se debe tener precaución. Los modelos generativos son muy eficaces creando imágenes verosímiles. Sin embargo, la verosimilitud no equivale a la verdad física. Aunque la imagen luzca impecable, puede diferir de la realidad. Por ello, se evalúan conjuntamente indicadores físicos como el tamaño del grano, la distribución de orientaciones y la proporción de fases. Solo cuando estos indicadores coinciden con la realidad se puede confiar en la imagen.

Incorporar la física en la función de pérdida incrementa la fiabilidad. Si el tamaño de los cristales en la imagen generada se desvía de la distribución real, se aplica una penalización. De este modo, se guía al generador para que dibuje imágenes conformes a la física. Se reduce así la brecha entre una imagen estéticamente atractiva y una físicamente correcta. Los modelos generativos solo pueden aplicarse a componentes de cohetes cuando se miden con la regla de la física.

6.4.3 Aprendizaje por refuerzo para controlar la morfología cristalina durante la fabricación

Ahora es el momento de controlar de forma autónoma la morfología de los cristales durante la fabricación. El aprendizaje por refuerzo (Reinforcement Learning) es un método para aprender acciones óptimas mediante intentos y recompensas. Es similar al proceso por el cual un bebé aprende a caminar cayéndose y levantándose. Si actúa bien, recibe un premio; si lo hace mal, recibe un castigo. Así, incrementa progresivamente las acciones que le reportan mayores recompensas.

En la fabricación aditiva, la acción corresponde al ajuste del láser. La acción consiste en cuánto aumentar o reducir la potencia del láser, y en cuánto variar su velocidad. Sin embargo, la forma de los cristales no se puede observar directamente durante la fabricación. Por ello, se observa en su lugar la temperatura y la forma del baño de fusión. A partir de estos valores, se infiere la morfología cristalina. La pirometría informa de este estado en tiempo real.

La recompensa se determina según la proximidad al objetivo. Cuanto más se acerque a la morfología cristalina deseada, mayor será el premio. Si la proporción de granos equiaxiales alcanza el objetivo, la recompensa aumenta. Por el contrario, si se produce ojo de cerradura (keyhole) o falta de fusión, se aplica una penalización. Guiada por estas recompensas, la inteligencia artificial aprende a regular el láser.

Cómo estructurar la recompensa determina el éxito o el fracaso. Si la recompensa se diseña de manera descuidada, se desvía por el camino equivocado. A veces se persigue únicamente la forma de los granos cristalinos y se desperdicia energía. Por eso, la recompensa se diseña incluyendo la forma cristalina, los defectos y la energía al mismo tiempo. Se trata de poner múltiples objetivos en la balanza y equilibrarlos. Diseñar adecuadamente esta balanza es el núcleo del diseño del aprendizaje por refuerzo.

¿Por qué es necesario este control en tiempo real? Porque la forma de la pieza varía según la zona. Las esquinas, las paredes delgadas y las partes salientes tienen diferentes velocidades de disipación del calor. Aunque se apliquen las mismas condiciones de láser, la forma cristalina cambia según el lugar. El control en tiempo real busca mantener uniforme la forma cristalina modificando las condiciones en cada punto.

Esta dirección aún no es un estándar industrial. Abundan las propuestas en etapas de simulación y probetas de ensayo. Son escasos los casos en los que se ha controlado la forma cristalina en bucle cerrado y en tiempo real. Esto se debe a que es difícil medir la forma cristalina directamente durante el proceso. Además, se deben superar el retardo de los sensores, los errores del modelo y los límites de respuesta del equipo.

El aprendizaje por refuerzo es excelente para experimentar de antemano intentos arriesgados. Si se prueban malas condiciones en equipos reales, la pieza se echa a perder. En las simulaciones por computadora, es posible fallar todo lo que se quiera. A partir de innumerables fallos, se aprende previamente una buena regulación. Tras aprender, se traslada con cautela al equipo real dentro de un rango seguro. Es como si la computadora pagara el costo del fracaso en nuestro lugar.

Por eso el orden es fundamental. Primero se aprende el control en la simulación. Luego se comprueba en probetas de ensayo limitadas. Se utiliza como herramienta auxiliar dentro de una ventana de proceso verificada. Solo al seguir estos pasos se puede introducir con precaución el control por aprendizaje por refuerzo en piezas reales. Una vez fabricada la pieza, la forma cristalina se vuelve a verificar mediante EBSD para comprobar si el control fue correcto.

6.5 Monitorización de defectos escuchando el interior mediante sonido

Hasta ahora se buscaban defectos mediante la luz. Sin embargo, la luz solo observa la superficie. Las grietas que se originan en lo profundo del interior de la pieza son difíciles de captar con cámaras. Esta sección aborda los métodos para escuchar el interior a través del sonido.

Primero examinamos qué tipo de sonido emiten los defectos. Luego vemos cómo convertir ese sonido en una imagen. Por último, analizamos cómo juzgar mejor utilizando tanto el sonido como la luz juntos. Al combinar sentidos distintos, se compensa lo que uno de ellos pasa por alto.

6.5.1 Mecanismo físico de las señales de emisión acústica (AE) de banda ancha

La emisión acústica (Acoustic Emission; AE) es el sonido que se produce cuando ocurre un cambio repentino dentro de un material. En rigor, se trata de ondas elásticas diminutas. Sigue el mismo principio por el cual un vaso de vidrio emite un chasquido cuando se agrieta. También es como el crujido que se oye cuando la madera se dobla y se quiebra. Los sucesos en el interior del material se filtran al exterior en forma de sonido.

En la fabricación aditiva, diversos sucesos emiten sonidos diferentes entre sí. El agrietamiento en caliente libera una gran cantidad de energía en un instante. Por eso emite un sonido agudo de alta frecuencia. Es similar a un chasquido corto y seco. Este tipo de sonido tiene una gran amplitud y surge repentinamente.

Cuando el ojo de cerradura (keyhole) colapsa, también se produce un sonido. El espacio vacío dentro del baño de fusión se hunde repentinamente y genera un impacto. Este sonido oscila de forma continua a una frecuencia algo más baja que la de las grietas. Se asemeja a la turbulencia que se produce cuando colapsa una burbuja en el agua. La textura del sonido es distinta a la de una grieta.

También se genera sonido cuando el polvo salpica o el equipo roza. Este sonido se sitúa en frecuencias aún más bajas. Los ruidos normales que no son defectos también se mezclan en el sonido. Por eso es necesario filtrar y escuchar con atención. La clave reside en distinguir entre el sonido de un defecto y el simple ruido.

Existe una razón por la cual cada sonido tiene una textura diferente. Si se libera repentinamente una gran energía, se transmiten frecuencias altas. Si se libera lentamente, se transmiten frecuencias bajas. Las grietas se abren en un abrir y cerrar de ojos, por lo que emiten un sonido de tono alto. Las salpicaduras de polvo son relativamente lentas, por lo que producen un sonido de tono bajo. No obstante, las frecuencias captadas en la práctica varían según los sensores, las vías de propagación y los métodos de filtrado. Por lo tanto, este orden debe entenderse como una tendencia observada en ensayos específicos.

El método de escucha acústica ofrece una gran ventaja. La luz solo ve la superficie, pero el sonido atraviesa el interior. Incluso las grietas en lo profundo del cuerpo fabricado pueden captarse mediante el sonido. Detectar el momento exacto en que surge una grieta permite responder con rapidez. Se puede detener el proceso en el acto, marcar la zona e inspeccionarla de nuevo. También se podría pensar en una reparación volviendo a aplicar el láser para refundir el material. Sin embargo, la refundición no siempre elimina las grietas. Incluso puede inducir nuevas tensiones térmicas y nuevas fisuras. Hay que diferenciar entre un proceso de reparación validado y una simple refundición.

Existen dos formas de escuchar el sonido. Una consiste en escuchar mediante sensores piezoeléctricos en contacto con la pieza. Un sensor piezoeléctrico es un dispositivo que genera electricidad al recibir una fuerza. Se adhiere a la placa de soporte de la pieza para captar las ondas elásticas que han

atravesado el interior. La otra consiste en captar el sonido en el aire mediante un micrófono. Ambos métodos difieren en las trayectorias de propagación del sonido y en sus frecuencias.

6.5.2 Transformada wavelet de tiempo-frecuencia

Analizar la señal acústica en su forma original dificulta el diagnóstico. Resulta complicado saber qué tipo de defecto es basándose únicamente en la oscilación a lo largo del tiempo. Esto se debe a que sonidos de diversas frecuencias se superponen en el mismo instante. Es necesario examinar conjuntamente cuándo ocurrió un sonido y con qué tono o frecuencia. Solo así se puede leer la huella dactilar del defecto.

La transformada wavelet (Wavelet Transform) se encarga de esta tarea. Este método representa conjuntamente el tiempo y la frecuencia. Convierte la señal de sonido en una imagen donde el eje horizontal es el tiempo y el eje vertical es la frecuencia. Esta imagen se denomina escalograma (Scalogram). Se parece a una partitura musical que muestra el tiempo y la altura de las notas a la vez.

Esta imagen dibuja patrones diferentes para cada defecto. El sonido agudo de una grieta queda registrado como un punto brillante en la zona superior. Las fluctuaciones del ojo de cerradura aparecen como un patrón continuo en la zona media. El ruido de fondo se extiende de manera difusa en la parte inferior. De este modo, al convertirse el sonido en una imagen, resulta fácil compararlo visualmente.

Al leer los patrones, es fundamental elegir bien la wavelet madre. La wavelet madre es una forma de onda corta que sirve de patrón de referencia para comparar el sonido. Se debe utilizar una referencia cuya textura se asemeje a la del sonido para que el patrón sea nítido. Si se utiliza una referencia con una textura diferente, el patrón se vuelve borroso. Elegir el patrón adecuado también forma parte del preprocesamiento. Solo si se selecciona la referencia correcta se obtiene una imagen con patrones nítidos.

La inteligencia artificial aprende los patrones a partir de estas imágenes. Redes neuronales diseñadas para procesar imágenes leen los escalogramas. Aprenden qué patrón corresponde a cada defecto a través de numerosos ejemplos. Tras aprender, identifican los defectos al analizar imágenes nuevas. Gracias a la conversión del sonido en imagen, se puede aprovechar la potencia de la inteligencia artificial de visión por computadora.

Estudios recientes han logrado buenos resultados con este método. Una investigación estimó los poros por ojo de cerradura únicamente a partir del sonido en el aire y la velocidad del láser. Se entrenó el modelo convirtiendo el sonido captado por el micrófono en escalogramas. El valor que indica la fiabilidad de los resultados superó 0,8. Este estudio corresponde a resultados publicados en revistas académicas. Demostró que es posible estimar defectos internos en milisegundos utilizando únicamente el sonido.

La monitorización por sonido tiene también la ventaja de ser económica. Las cámaras de alta velocidad y los equipos de rayos X de precisión son costosos y difíciles de manejar. Los sensores piezoeléctricos son relativamente baratos y fáciles de instalar. Si se colocan varios, también es posible estimar de dónde proviene el sonido. Por ello, la monitorización basada en sonido destaca como una vía práctica en el campo de trabajo.

El principio de localización mediante múltiples sensores es sencillo. El mismo sonido llega a cada sensor en instantes diferentes. Llega primero al sensor más cercano y más tarde al más lejano. Al comparar estas diferencias en el tiempo de llegada, se estima el punto de origen del sonido. Es similar a determinar una dirección a partir del desfase temporal con que el sonido llega a varios micrófonos. La posición hallada de este modo nos indica dónde se encuentra el defecto.

6.5.3 Detección de defectos mediante el uso combinado de luz y sonido

Multimodal (Multimodal) significa utilizar varios tipos de datos de forma combinada. Es similar a cuando una persona utiliza los ojos y los oídos al mismo tiempo. Si solo se escucha el sonido, hay aspectos que se pasan por alto; si solo se observa la luz, también se pierden detalles. Al combinar ambos, se cubren las carencias mutuas. Esta sección aborda los métodos para integrar el sonido y la luz.

La luz y el sonido observan cosas distintas. La pirometría muestra claramente la temperatura de la superficie. La emisión acústica capta los cambios repentinos que se originan en el interior. Las anomalías térmicas en la superficie las detecta primero la cámara. Las grietas internas las percibe primero el sonido. Al fusionar ambas señales, se examinan simultáneamente la superficie y el interior.

El método de integración es el siguiente. La señal acústica es leída por una red neuronal dedicada al sonido. Las imágenes ópticas son procesadas por una red neuronal dedicada a imágenes. Las características extraídas por ambas redes se reúnen en un solo lugar para emitir el juicio final. Se trata de asignar a cada una la tarea en la que destaca y luego aunar fuerzas en la conclusión.

Investigaciones reales han demostrado la potencia de esta integración. Un estudio recopiló conjuntamente datos de pirometría y sonido. Al combinar ambas señales, se identificaron mejor los poros por ojo de cerradura. Se redujo el número de poros omitidos en comparación con el uso de una sola señal. Lograr detectar también las grietas de esta manera combinada constituye un desafío para el futuro.

Al combinarlas, el ajuste temporal es más importante que cualquier otra cosa. El instante en que se produce el sonido y el momento en que cambia la luz deben sincronizarse con exactitud. Si los tiempos no coinciden, se pueden interpretar sucesos distintos como si fueran el mismo. Por ello, ambos sensores se sincronizan con altísima precisión. Dado que el ruido difiere según el sensor, también se requiere un acondicionamiento previo de las señales.

Existen diversas formas de fusionar ambas señales. Hay un enfoque en el que las señales se mezclan desde el principio para aprender juntas. También existe el método de aprender por separado y fusionar los resultados únicamente en la conclusión. Los estudios recientes suelen emplear el método de combinar las características aprendidas por separado mediante un clasificador ligero. Este método es computacionalmente liviano y se adapta bien a nuevos defectos. La idoneidad de cada método depende de la pieza y del equipo.

La tendencia a combinar múltiples sentidos se ampliará aún más en el futuro. Además de la luz y el sonido, también se utiliza la señal de reflexión del láser. Se recopilan conjuntamente fotografías de las capas fabricadas y mapas de temperatura. Múltiples ojos y oídos examinan una misma pieza capa por capa. Cuantos más sentidos haya, menos defectos se pasarán por alto. La tarea de articular todas estas señales en un conjunto integrado le corresponde a la inteligencia artificial.

Este enfoque también presenta limitaciones evidentes. Es más seguro emplear el criterio de la inteligencia artificial para emitir alertas durante la producción y para preseleccionar candidatos defectuosos. El dictamen definitivo de aceptación se emite mediante ensayos no destructivos, observación de secciones transversales y tomografía computarizada. El desfase temporal entre sensores y el ruido también pueden provocar diagnósticos erróneos. La inteligencia artificial es un guía que acota las zonas que el ser humano debe inspeccionar.

Resumen

Este capítulo ha examinado la tecnología de fabricación de componentes de cohetes mediante la deposición y superposición de metales. Hemos visto los métodos de fusión de polvo mediante láser y de fusión de hilo metálico. Se han analizado las propiedades y precauciones de tres metales: Inconel 718, GRCop-42 y C-103. Cada metal tiene misiones distintas y plantea diferentes problemas que controlar durante su fabricación.

La fabricación aditiva ofrece libertad, pero también entraña el riesgo de defectos. Los poros por ojo de cerradura y la falta de fusión se convierten en los puntos de inicio de las grietas por fatiga. Las tensiones residuales deforman las piezas de gran tamaño. La orientación preferencial de los cristales sesga la resistencia mecánica hacia una sola dirección. Estos defectos se relacionan de forma directa con la seguridad del motor.

Estos defectos están entrelazados entre sí. Una sola condición de láser inadecuada provoca múltiples defectos simultáneos. Al intentar evitar el ojo de cerradura, se puede provocar una falta de fusión. Al intentar reducir las tensiones, la forma cristalina puede deteriorarse. Por eso no se deben cambiar las condiciones fijándose en un único defecto. Es necesario sopesar múltiples defectos a la vez y encontrar el equilibrio.

La inteligencia artificial ayuda a resolver este problema desde tres direcciones. La primera consiste en detectar rápidamente los defectos durante la fabricación. Modelos que incorporan leyes físicas en imágenes de pirometría identifican los poros. Modelos que leen el sonido tras convertirlo en imagen captan las grietas internas. Al combinar luz y sonido, se examinan la superficie y el interior al mismo tiempo.

La siguiente consiste en sustituir los cálculos lentos para predecir con rapidez las tensiones térmicas y las deformaciones. Los modelos sustitutos reducen a poco tiempo análisis que requerirían varios días. Con esta predicción rápida se preseleccionan innumerables condiciones de proceso. Al incorporar la física, las predicciones no se apartan de las leyes físicas.

La última consiste en ayudar a ajustar las condiciones del proceso en tiempo real. Se reducen las deformaciones regulando la trayectoria de escaneo y el tiempo de enfriamiento. Se ajusta el láser para mantener uniforme la forma cristalina. Los modelos generativos dibujan de antemano la forma cristalina a partir de las condiciones dadas. El aprendizaje por refuerzo controla de manera autónoma las condiciones durante la fabricación.

Debe quedar algo claro: la inteligencia artificial no puede sustituir todas las inspecciones. Los diversos modelos examinados en este capítulo son rápidos e inteligentes, pero no son perfectos. Abundan los estudios iniciales y son pocos los que se han consolidado como estándares industriales. Es más seguro utilizar el juicio de la inteligencia artificial para emitir alertas y preseleccionar candidatos. La garantía de calidad final se apoya conjuntamente en probetas de ensayo, ensayos no destructivos y pruebas de combustión.

Existe también una nueva tendencia que utiliza modelos de lenguaje. En la fabricación aditiva se acumulan montañas de artículos e informes. Los modelos de lenguaje grandes (Large Language Model; LLM) leen estos textos y organizan el conocimiento. Es posible consultarles en qué condiciones suelen originarse determinados defectos. Un estudio de 2025 intentó predecir las condiciones de aparición de defectos mediante modelos de lenguaje. Estas herramientas se convierten en un puente entre la experiencia humana y la literatura científica.

El panorama general es el siguiente. Los sensores actúan como ojos y oídos, los modelos sustitutos como el cerebro y los controladores como las manos. Cuando estos tres elementos se conectan como uno solo,

surge una fábrica que produce piezas supervisándose a sí misma. El hecho de que la NASA haya fabricado una tobera de gran tamaño mediante fabricación aditiva y superado prolongadas pruebas de combustión es el punto de partida. En el siguiente capítulo examinaremos cómo dar acabado a las piezas así fabricadas y cómo evaluar si resistirán a largo plazo.

Fuentes y referencias relacionadas

DebRoy, T., Wei, H. L., Zuback, J. S., Mukherjee, T., Elmer, J. W., Milewski, J. O., Beese, A. M., Wilson-Heid, A., De, A., & Zhang, W. (2018). Additive manufacturing of metallic components: Process, structure and properties. Progress in Materials Science, 92, 112-224. https://doi.org/10.1016/j.pmatsci.2017.10.001

King, W. E., Barth, H. D., Castillo, V. M., Gallegos, G. F., Gibbs, J. W., Hahn, D. E., Kamath, C., & Rubenchik, A. M. (2014). Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing. Journal of Materials Processing Technology, 214, 2915-2925. https://doi.org/10.1016/j.jmatprotec.2014.06.005

Rosenthal, D. (1946). The theory of moving sources of heat and its application to metal treatments. Transactions of the ASME, 68, 849-866.

Hunt, J. D. (1984). Steady state columnar and equiaxed growth of dendrites and eutectic. Materials Science and Engineering, 65, 75-83. https://doi.org/10.1016/0025-5416(84)90201-5

Mukherjee, T., Manvatkar, V., De, A., & DebRoy, T. (2017). Mitigation of thermal distortion during additive manufacturing. Scripta Materialia, 127, 79-83. https://doi.org/10.1016/j.scriptamat.2016.09.001

Gradl, P. R., Protz, C. S., Cooper, K., Ellis, D., Evans, L. J., & Garcia, C. (2019). GRCop-42 Development and Hot-fire Testing Using Additive Manufacturing Powder Bed Fusion for Channel-cooled Combustion Chambers. AIAA Propulsion and Energy 2019 Forum. https://doi.org/10.2514/6.2019-4228

Gradl, P., Mireles, O. R., Katsarelis, C., Smith, T. M., Sowards, J., Park, A., et al. (2023). Advancement of extreme environment additively manufactured alloys for next generation space propulsion applications. Acta Astronautica, 211, 483-497. https://doi.org/10.1016/j.actaastro.2023.06.035

NASA. (2023). Rapid Analysis and Manufacturing Propulsion Technology (RAMPT). https://www.nasa.gov/rapid-analysis-and-manufacturing-propulsion-technology/

NASA. (2023). GRCop-42 additive manufacturing for high heat flux thrust chamber liners (Document ID 20230007103). NASA Technical Reports Server. https://ntrs.nasa.gov/citations/20230007103

NIST. (2024). Additive Manufacturing Benchmark Test Series (AM Bench). National Institute of Standards and Technology. https://www.nist.gov/ambench

NIST. (2022). AM Bench 2022 measurement results data: in-situ thermography and scan strategy laser scans. National Institute of Standards and Technology. https://www.nist.gov/data-publications/am-bench-2022-measurement-results-data-situ-thermography-and-scan-strategy-laser

Levine, L. E., Lane, B. M., Heigel, J. C., et al. (2024). Outcomes and conclusions from the 2022 AM Bench measurements, challenge problems, and conference. Integrating Materials and Manufacturing Innovation, 13, 154-193. https://doi.org/10.1007/s40192-024-00372-4

Aydogan, B., & Chou, K. (2024). Review of in situ detection and ex situ characterization of porosity in laser powder bed fusion metal additive manufacturing. Metals, 14(6), 669. https://doi.org/10.3390/met14060669

Chen, L., Bi, G., Yao, X., Su, J., Tan, C., Feng, W., Benakis, M., Chew, Y., & Moon, S. K. (2024). In-situ process monitoring and adaptive quality enhancement in laser additive manufacturing: a critical review. arXiv:2404.13673. https://arxiv.org/abs/2404.13673

Emmanuel, I., Yang, Z., Yeung, H., & Zhang, X. (2026). Machine learning modeling for real-time melt pool monitoring in laser powder bed fusion additive manufacturing: a hybrid approach. arXiv:2606.23851. https://arxiv.org/abs/2606.23851

Liu, H., Guirguis, D., Zeng, X., Maurer, L., Rajan, V., Sanaei, N., Yang, C.-T., Beuth, J. L., Rollett, A. D., & Kara, L. B. (2026). Airborne acoustic emission enables sub-scanline keyhole porosity quantification and effective process characterization for metallic laser powder bed fusion. Additive Manufacturing, 116, 105062. https://doi.org/10.1016/j.addma.2025.105062

Tian, M., Mu, H., Ding, D., Li, M., Ding, Y., & Zhao, J. (2025). Real-time distortion prediction in metallic additive manufacturing via a physics-informed neural operator approach. arXiv:2511.13178. https://arxiv.org/abs/2511.13178

Zhang, R., Strickland, J., Hou, X., Yang, F., Li, X., de Oliveira, J. A., Li, J., & Zhang, S. (2025). Rapid residual stress simulation and distortion mitigation in laser additive manufacturing through machine learning. Additive Manufacturing, 102, 104721. https://doi.org/10.1016/j.addma.2025.104721

Bostan, B., Hinnebusch, S., Anderson, D., & To, A. C. (2025). Accurate detection of local porosity in laser powder bed fusion through deep learning of physics-based in-situ infrared camera signatures. Additive Manufacturing, 101, 104701. https://doi.org/10.1016/j.addma.2025.104701

Giri, S., Liu, S., Gorgannejad, S., Thampy, V., Quan, P., Nicolino, J. W., Strantza, M., Forien, J.-B., Martin, A. A., Calta, N. P., & Tassone, C. J. (2026). In-situ sensor monitoring of multi-class gas porosity formation in laser powder bed fusion using convolutional neural network. Journal of Intelligent Manufacturing. https://doi.org/10.1007/s10845-026-02861-z

Zhang, H., Caputo, A. N., Neu, R. W., Vallabh, C. K. P., To, A. C., Alam, M. J., & Zhao, X. (2026). Process-agnostic multi-metric nondestructive characterization of porosity in additive manufacturing via specialized single-camera two-wavelength pyrometry. Additive Manufacturing, 127, 105282. https://doi.org/10.1016/j.addma.2026.105282

Tempelman, J. R., Wachtor, A. J., Flynn, E. B., Depond, P. J., Forien, J.-B., Guss, G. M., Calta, N. P., & Matthews, M. J. (2022). Sensor fusion of pyrometry and acoustic measurements for localized keyhole pore identification in laser powder bed fusion. Journal of Materials Processing Technology, 308, 117656. https://doi.org/10.1016/j.jmatprotec.2022.117656

Li, Z., Zheng, H., Zhang, Z., Wang, J., Wen, G., Grasso, M., & Colosimo, B. M. (2026). On the use of acoustic emission methods for in-situ monitoring in metal additive manufacturing: a review study. Additive Manufacturing Frontiers, 5(2), 200347. https://doi.org/10.1016/j.amf.2026.200347

Brooke, R., Zhang, D., Qiu, D., Gibson, M. A., & Easton, M. (2025). Compositional criteria to predict columnar to equiaxed transitions in metal additive manufacturing. Nature Communications, 16, 5710. https://doi.org/10.1038/s41467-025-60162-0

Prabhune, B., Panicker, N., Jordan, B., Lim, B., Snyder, K., Delchini, M., See, N. D., Nag, S., Plotnikov, Y., & Lee, Y. (2025). Multi-physics melt pool modeling and process optimization for laser direct energy deposition of Nb-based refractory C103. Journal of Manufacturing Processes, 154, 444-461. https://doi.org/10.1016/j.jmapro.2025.09.082

Jeong, W., Cho, S., & Ryu, H. J. (2025). Optimization and characterization of C103 Nb alloy fabricated by laser wire directed energy deposition. International Journal of Refractory Metals and Hard Materials, 133, 107318. https://doi.org/10.1016/j.ijrmhm.2025.107318

Xu, Z. Y., Mao, J. Y., Ren, R. X., Wu, Z. Q., Lei, Y. J., Han, L. L., Chen, C., Fan, W., & Wang, Y. (2026). Research progress and challenges of intelligent in situ monitoring for laser additive manufacturing processes based on machine learning. Advances in Mechanical Engineering. https://doi.org/10.1177/16878132261442327

Pak, P., & Barati Farimani, A. (2025). AdditiveLLM: Large Language Models Predict Defects in Metals Additive Manufacturing. arXiv:2501.17784. https://arxiv.org/abs/2501.17784

Capítulo 7. Postprocesamiento y garantía de vida a la fatiga de alto ciclo de aleaciones aeroespaciales impresas en 3D

Introducción

Pensemos en la impresora 3D que usamos en casa. Crea objetos fundiendo filamentos de plástico y acumulándolos capa por capa. Los componentes de los motores cohete se fabrican de manera similar. Sin embargo, el material no es plástico, sino polvo metálico. Un láser funde instantáneamente el polvo metálico y lo solidifica capa por capa. De este modo, es posible fabricar formas complejas de una sola pieza.

Este método ha transformado radicalmente la fabricación de piezas de cohetes. En el pasado, se fabricaban varias piezas por separado y se unían mediante soldadura. En la actualidad, incluso las piezas con conductos de refrigeración internos se fabrican en una sola pieza. La cantidad de componentes disminuye y el peso se reduce. Al desaparecer el proceso de unión de múltiples piezas, el tiempo de producción también se acorta.

Sin embargo, existe un problema. El metal recién salido de la impresora no se puede instalar directamente en un cohete. La superficie exterior es rugosa y en su interior quedan pequeños poros. Además, contiene tensiones invisibles atrapadas dentro. Si se hace girar a alta velocidad en este estado, existe el riesgo de que la pieza se rompa.

Por ello, el acabado posterior a la impresión es fundamental. A este tratamiento se le denomina postprocesamiento. Se mecaniza la superficie para pulirla y dejarla suave. Se eliminan los poros internos mediante calor y presión. Con el tratamiento térmico, la microestructura se homogeniza. Este capítulo aborda paso a paso los procesos de postprocesamiento, desde el pulido superficial hasta la eliminación de defectos internos.

Imaginemos la escena de cortar carne dura en la cocina. El cuchillo atraviesa fácilmente el tofu blando. Pero en la carne dura, el cuchillo apenas entra. Las superaleaciones para cohetes se parecen a la carne dura. Son muy difíciles de mecanizar. A esta dificultad para el corte se la denomina baja maquinabilidad.

Hay una preocupación aún mayor: la fatiga. La fatiga es el fenómeno por el cual un material se fractura debido a la aplicación repetida de fuerzas pequeñas. Si doblamos y desdoblamos un clip para papel varias veces, se rompe. Se rompe aunque no hayamos aplicado una gran fuerza de golpe. Los álabes giratorios de la turbobomba de un cohete están expuestos a este peligro.

En este capítulo examinaremos cómo la inteligencia artificial ayuda en dos tareas. La primera es encontrar las condiciones óptimas para mecanizar metales de difícil corte. La segunda es predecir cuántos ciclos de carga repetida puede soportar una pieza. Ambas tareas requieren mucho tiempo si solo se recurre a la experiencia humana. La inteligencia artificial aprende de los datos y acota rápidamente las respuestas.

Sin embargo, la inteligencia artificial no lo decide todo. Los componentes de los cohetes son elementos de los que dependen vidas humanas y enormes inversiones económicas. Las predicciones de la inteligencia artificial deben reconfirmarse mediante experimentos e inspecciones. Este capítulo explica cómo la inteligencia artificial y los experimentos colaboran y se validan mutuamente.

Al terminar de leer este capítulo, dos aspectos quedarán claros. Uno es por qué los metales impresos en 3D necesitan un acabado posterior. El otro es cómo la inteligencia artificial contribuye a ese acabado y a su inspección. No utilizaremos fórmulas complejas. Lo explicaremos paso a paso con ejemplos de la vida cotidiana. Si avanza con calma, comprenderá sin dificultad el panorama general de este capítulo.

7.1 La transición hacia piezas de cohetes impresas en 3D y nuevos desafíos

Comencemos explicando lo que aborda esta sección. El núcleo de un motor cohete es un entorno extremadamente severo. Por un lado fluye combustible criogénico a más de 200 grados bajo cero. Por el otro pasan gases calientes a cientos de grados. La presión es alta y las vibraciones son intensas. En esta sección examinaremos por qué los metales utilizados en estos entornos se fabrican mediante impresión 3D y qué debilidades surgen en dicho proceso.

Los componentes de los motores cohete deben ser ligeros y resistentes. Si son pesados, la carga útil que se puede transportar disminuye. Si su resistencia es baja, no pueden soportar altas presiones. Estos dos requisitos entran en conflicto mutuo, ya que al aligerar una pieza suele volverse más frágil. La fabricación aditiva (Additive Manufacturing, AM) abrió el camino para resolver este problema.

La cámara de combustión con refrigeración regenerativa es un buen ejemplo. Las paredes de la cámara de combustión están en contacto directo con llamas ardientes. Si se dejaran tal cual, las paredes se fundirían. Por eso se crean finos conductos de refrigeración dentro de la pared. El combustible frío pasa por esos conductos y enfría la pared. Estructuras con conductos ocultos en el interior de las paredes eran difíciles de fabricar con métodos tradicionales. La fabricación aditiva produce estas complejas estructuras internas de una sola vez.

La tecnología representativa para fabricar estas piezas es la fusión por lecho de polvo láser (Laser Powder Bed Fusion, LPBF). Un láser pasa sobre una fina capa de polvo metálico extendido. La zona donde incide el láser se funde instantáneamente y se solidifica de inmediato. Sobre ella se vuelve a extender polvo y se aplica nuevamente el láser. Este proceso se repite miles de veces para construir la pieza capa a capa.

El metal más utilizado en componentes de cohetes es la superaleación (superalloy) con base de níquel. Una superaleación es un metal especial que no pierde su resistencia incluso a altas temperaturas. Su material más representativo es el Inconel 718 (Inconel 718). También se utilizan aleaciones de la serie Haynes (Haynes). Estos metales resisten durante mucho tiempo incluso en presencia de gases calientes.

El problema radica en las condiciones internas del metal fundido por el láser. El lugar por donde pasa el láser se enfría con extraordinaria rapidez, a una velocidad de cientos de miles de grados por segundo. Al enfriarse tan bruscamente, la estructura se solidifica de manera desigual. Los cristales crecen de forma alargada en la dirección de deposición. A estos cristales alargados se les llama granos columnares (columnar grain).

También surge el fenómeno por el cual las propiedades varían según la dirección. A esto se le conoce como anisotropía (anisotropy). Es fácil de entender si pensamos en la madera. La madera se divide con facilidad si se corta a lo largo de la veta, pero cortarla a través de la veta requiere más esfuerzo. De igual modo, el metal fabricado por adición tiene resistencias diferentes en cada dirección. Esta propiedad complica el diseño y el mecanizado de las piezas.

El enfriamiento rápido deja también otras huellas. Si vertemos agua caliente en un vaso de vidrio congelado en invierno, se agrieta. Esto ocurre porque el interior y el exterior se expanden a ritmos diferentes. En el metal fabricado por adición ocurre algo parecido. La velocidad de solidificación varía

según la zona. Por lo tanto, se acumulan fuerzas atrapadas dentro de la pieza. A esta fuerza atrapada se le denomina tensión residual (residual stress).

A esto se suma la segregación de componentes. Cuando el metal fundido se solidifica, ciertos elementos se concentran en zonas específicas. En el Inconel 718, el elemento niobio (Nb) tiende a concentrarse. En esos puntos se forman partículas duras. A estas partículas se les llama fase de Laves (Laves phase). La fase de Laves es mucho más dura y frágil que el metal base de la matriz.

Debido a estas circunstancias internas, surgen dos grandes desafíos. El primero es que resulta difícil de mecanizar. Las partículas duras dañan las herramientas de corte. El segundo es que los poros internos se convierten en semillas de grietas por fatiga. En piezas que giran a alta velocidad, estas semillas son peligrosas. Por ello, en las secciones siguientes abordaremos estos dos desafíos uno a uno.

Hay una razón por la que la fabricación aditiva se ha consolidado rápidamente en la industria de los cohetes: reduce drásticamente el tiempo necesario para fabricar una pieza. En el pasado, solo la creación de los moldes de fundición tomaba varios meses. La fabricación aditiva construye las piezas directamente sin necesidad de moldes. Si se modifica el diseño, se puede imprimir una nueva pieza al día siguiente. Poder repetir este ciclo de diseño y fabricación en intervalos tan cortos acelera el ritmo de desarrollo.

Su capacidad para reducir el peso también es notable. Al fusionar múltiples piezas en una sola, las uniones desaparecen. Las uniones añaden peso y suelen convertirse en puntos débiles. Al eliminarlas, la estructura se vuelve más ligera y resistente. Por esta razón, las empresas de cohetes invierten fuertemente en esta tecnología. La Administración Nacional de Aeronáutica y el Espacio (NASA) de Estados Unidos también ha probado cámaras de combustión aditivas durante mucho tiempo.

En el postratamiento, el tratamiento térmico es imprescindible. El tratamiento térmico consiste en calentar y enfriar el metal para transformar su microestructura. Primero, se calienta a alta temperatura para disolver las partículas duras en la matriz. A esta etapa se le llama tratamiento de solubilización. Luego, se mantiene a una temperatura adecuada durante un tiempo prolongado para que las partículas de refuerzo crezcan de manera uniforme. A esta etapa se le denomina envejecimiento. Solo tras superar estas dos etapas la superaleación alcanza su resistencia óptima.

En resumen, la fabricación aditiva trajo consigo grandes ventajas junto con nuevos desafíos. Las ventajas son la capacidad de crear formas complejas y la rapidez de producción. Los desafíos son las superficies rugosas y los defectos internos. La clave para resolver estos retos reside en el postprocesamiento y la inspección. Y la inteligencia artificial apoya de cerca todo ese proceso de postprocesamiento e inspección.

7.2 Por qué las superaleaciones fabricadas por adición son difíciles de mecanizar

Comencemos resumiendo los temas que se tratan en esta sección. La baja maquinabilidad se refiere a la dificultad para mecanizar un material. Las superaleaciones fabricadas por adición hacen que metales que ya de por sí eran difíciles de cortar sean aún más difíciles de mecanizar. Examinaremos las tres razones que lo explican: la direccionalidad de la estructura cristalina, las tensiones atrapadas y el endurecimiento por deformación, y el proceso de desgaste de las herramientas.

7.2.1 Estructura cristalina que varía según la dirección

En este apartado se aborda el impacto que la direccionalidad de la estructura cristalina tiene sobre el corte. Como se mencionó antes, los granos columnares crecen alargados en la dirección de deposición. Veremos en detalle por qué esta microestructura dificulta el mecanizado, y también examinaremos el impacto que la dura fase de Laves produce sobre la herramienta.

Los metales están formados por la agregación de pequeños cristales. En cada cristal, los átomos están alineados según un orden regular. La orientación de estos cristales determina las propiedades del material. En los metales aditivos, los cristales están orientados preferentemente en una dirección. Dado que los cristales crecen alargados a lo largo de la dirección de deposición, las propiedades varían según la dirección.

Cuando la herramienta de corte atraviesa esta microestructura, la resistencia varía. El ángulo en el que coinciden la orientación del cristal y el filo de la cuchilla cambia de un punto a otro. Algunas zonas se mecanizan suavemente, mientras que otras ofrecen una gran resistencia. Por ello, la fuerza de corte oscila de manera periódica. Al ser la fuerza irregular, el mecanizado se vuelve inestable y el acabado superficial resulta desparejo.

Estas oscilaciones provocan impactos repetidos en la herramienta. Al acumularse los impactos, el filo de corte se deteriora poco a poco y las vibraciones aumentan. A esta vibración se le llama trepidación (chatter). Si el chatter es severo, deja patrones ondulados en la superficie. En componentes de precisión, estas marcas representan un defecto grave.

A esto se añade que las partículas duras agravan el problema. Cuando el Inconel 718 se solidifica, el niobio se segrega hacia los límites de grano. En esos lugares se forma la fase de Laves en forma de placas o bloques. Estas partículas son mucho más duras que el metal de la matriz y poseen al mismo tiempo una alta fragilidad.

¿Qué ocurre cuando la herramienta de corte choca contra estas partículas duras? Es como intentar cortar con un cuchillo tierra blanda mezclada con grava. El filo sufre microdesconchados. Si estas microfracturas continúan acumulándose, el filo de la herramienta termina desgastándose y desafilándose por completo.

Para resolver este problema, primero se debe homogeneizar la microestructura. El tratamiento térmico, que trataremos más adelante, cumple esa función. El tratamiento térmico disuelve las partículas duras y las reincorpora a la matriz. También transforma los granos columnares alargados en formas más redondeadas. De este modo, se reduce la variación de propiedades según la dirección.

Esta diferencia en la microestructura incide directamente en el rendimiento de la pieza. Si la resistencia varía según la dirección, el diseño se complica. Es necesario orientar la pieza evitando las direcciones más débiles. Al homogeneizar la estructura, estas preocupaciones disminuyen. Por tanto, el postprocesamiento no solo facilita el mecanizado, sino que también mejora el rendimiento.

La inteligencia artificial incorpora esta información microestructural en sus predicciones. La orientación de los cristales y la cantidad de partículas duras presentes sirven como variables de entrada. El modelo aprende cómo cambiará la resistencia al corte en función de estos valores. Para una persona resulta difícil examinar visualmente cada fotografía de la microestructura, pero la inteligencia artificial analiza rápidamente numerosas imágenes para encontrar patrones.

La orientación de los cristales se mide con equipos especiales. Al emitir un haz de electrones sobre la superficie de la muestra, surge un patrón diferente para cada cristal. Mediante este patrón se conoce hacia dónde está orientado cada cristal. Esta información de orientación se convierte en una excelente

entrada para la inteligencia artificial. A los humanos les resulta difícil cuantificar visualmente las orientaciones, pero la máquina lee simultáneamente la orientación de innumerables cristales.

El ángulo de deposición también modifica las propiedades. La microestructura varía según si la pieza se construye en posición vertical o acostada. Aunque el plano sea idéntico, la resistencia cambia si la orientación de colocación es diferente. Por ello, la dirección de deposición se define desde la etapa de diseño. Este ángulo también influye en la maquinabilidad y en la fatiga. La inteligencia artificial introduce el ángulo como variable y aprende su impacto.

Sin embargo, la microestructura varía ligeramente entre una pieza y otra. Incluso si se fabrican con los mismos ajustes, existen variaciones según la zona. Pequeñas fluctuaciones del láser y diferencias en el tamaño del polvo generan estas desviaciones. En consecuencia, también persisten incertidumbres en las predicciones del modelo. En las secciones siguientes veremos en detalle cómo se gestiona esta variabilidad.

7.2.2 Tensiones atrapadas y endurecimiento progresivo durante el mecanizado

En este apartado se tratan las tensiones atrapadas y el endurecimiento por deformación. Como se mencionó anteriormente, en los metales aditivos quedan fuerzas atrapadas en su interior. Examinaremos qué problemas ocasionan estas fuerzas acumuladas durante el mecanizado, así como la propiedad por la cual las superaleaciones se vuelven más duras a medida que se cortan.

Repasemos primero la tensión residual. Dentro de la pieza existen fuerzas atrapadas en múltiples direcciones. Este conjunto de fuerzas se representa mediante una entidad matemática llamada tensor. Puede entenderse un tensor como una tabla que contiene simultáneamente valores en diversas direcciones. Aunque invisibles al ojo, estas fuerzas se empujan y atraen mutuamente en el interior de la pieza.

Estas fuerzas atrapadas manifiestan sus problemas en el instante del mecanizado. Al cortar y retirar la superficie, las fuerzas de esa zona se liberan. Entonces, la parte restante pierde el equilibrio. Componentes como álabes delgados o anillos se deforman y curvan. Al intentar ajustarse a dimensiones de precisión, la forma termina distorsionándose. Por esta razón, debe predefinirse rigurosamente la secuencia y el orden de mecanizado de cada superficie.

El siguiente factor es el endurecimiento por deformación (work hardening). El endurecimiento por deformación es la propiedad por la cual un metal se vuelve más duro a medida que se deforma. Si doblamos un alambre de hierro varias veces, esa zona se vuelve rígida. Las superaleaciones presentan esta propiedad de forma muy pronunciada. La fina capa situada justo delante del filo se endurece instantáneamente.

Entonces, la herramienta debe mecanizar de nuevo la capa recién endurecida. Requiere más fuerza. También genera más calor. El Inconel 718 es un metal que disipa mal el calor. Su conductividad térmica apenas alcanza un tercio de la del acero común. Por ello, el calor atrapado se acumula entre el filo de corte y la pieza, elevando aún más la temperatura.

Las fases de refuerzo también dificultan el trabajo de la herramienta. En el Inconel 718 se encuentran dispersas partículas de refuerzo muy pequeñas. La principal partícula de refuerzo de esta aleación es la fase gamma doble prima (γ″, Ni3Nb), que contiene niobio. Estas partículas mantienen firme y duro el metal. Incluso a altas temperaturas, estas partículas no se disuelven fácilmente. Por eso, el metal no se ablanda en la zona caliente de corte y la herramienta sigue soportando una gran fuerza.

Todos estos factores se combinan para provocar daños superficiales. Se produce una alteración en una capa delgada situada debajo de la superficie mecanizada. La microestructura se fragmenta finamente y

su color se vuelve blanco. A esta capa se le llama capa blanca (white layer). En la capa blanca es fácil que queden tanto tensiones residuales de tracción como microfisuras. Esta capa se convierte en el punto de inicio de grietas y reduce drásticamente la vida a fatiga de la pieza.

¿Por qué es peligrosa la capa blanca? En esta capa existen tensiones residuales de tracción invisibles a simple vista. La tensión de tracción es una fuerza que abre las grietas. Si a esto se suman las microfisuras, se convierten en semillas de falla. Al someterse a cargas repetidas, las grietas crecen a partir de estas semillas. Por lo tanto, en el acabado final no debe quedar ninguna capa blanca.

También existen métodos para mitigar este problema. Tras el mecanizado, se añade un postratamiento que bombardea la superficie con pequeñas partículas. Este tratamiento deja tensiones de compresión en la superficie. La tensión de compresión es una fuerza que presiona y cierra las grietas. Actúa en dirección opuesta a la tensión de tracción. Al inducir así tensiones de compresión en la superficie, se prolonga la vida a fatiga de la pieza.

La inteligencia artificial también aprende los efectos combinados de estos diversos postratamientos. Se introducen como entradas las condiciones de mecanizado, el tratamiento de granallado y el tratamiento térmico. A partir de ello, predice qué combinación produce buenos resultados. Esto reduce notablemente la necesidad de que los operarios realicen ensayos uno por uno. Aun así, la verificación final se realiza mediante pruebas reales.

La inteligencia artificial se utiliza para detectar este riesgo de forma anticipada. Si se introducen las condiciones de corte y los valores de tensión residual, predice el estado de la superficie. Aprende en qué condiciones se forma la capa blanca. De este modo, se pueden evitar las condiciones peligrosas. En piezas fabricadas por adición de titanio, ya han surgido modelos de aprendizaje profundo que predicen la rugosidad superficial y la dureza a partir de las condiciones del proceso. Este libro propone aplicar el mismo enfoque a las superficies mecanizadas de superaleaciones.

Sin embargo, las predicciones solo son fiables dentro del rango de datos con el que se ha entrenado. En condiciones nunca antes vistas, el margen de error aumenta. Por eso, las piezas reales siempre deben verificarse mediante inspección superficial. Esto se debe a que la capa blanca es tan delgada que resulta fácil pasarla por alto a simple vista. La inteligencia artificial filtra previamente las condiciones de riesgo para reducir las opciones candidatas.

7.2.3 Mecanismos de desgaste de la herramienta y daño superficial

En este apartado se abordan los diversos mecanismos de desgaste de la herramienta. El desgaste de la herramienta (tool wear) repercute de forma directa en la calidad de la producción. Cuando la herramienta se desgasta, la superficie mecanizada se vuelve más rugosa. Una superficie rugosa sirve de punto de inicio para grietas por fatiga. Por ello, conocer de antemano el grado de desgaste de la herramienta resulta crucial en el control de calidad.

Existen varios tipos de desgaste de la herramienta. El primero es el desgaste por entalla (notch wear). En el límite de la profundidad de corte se forma una ranura profunda en la herramienta. Esta zona coincide con el lugar donde rozan simultáneamente el oxígeno del aire y la capa de metal endurecida. La fricción conjunta del oxígeno y la capa metálica endurecida hace que esta ranura se vuelva cada vez más profunda.

El segundo es el desprendimiento por adhesión. El níquel tiende a adherirse fácilmente al material de la herramienta. Las virutas de metal cortadas se pegan sobre el filo. A esta acumulación se le denomina filo recrecido (built-up edge). Cuando esta masa se desprende, arranca consigo el recubrimiento de la superficie de la herramienta. Al arrancarse el recubrimiento, la superficie de la herramienta se deteriora rápidamente.

El tercero es el desgaste por difusión. La zona de corte alcanza temperaturas sumamente elevadas. A altas temperaturas, los átomos se difunden entre sí. Los átomos de la herramienta migran hacia las virutas metálicas. Como consecuencia, en la superficie de la herramienta queda una cavidad hundida llamada cráter de desgaste.

El cuarto corresponde a la capa blanca superficial y las microfisuras. La capa blanca explicada anteriormente vuelve a presentarse aquí. La combinación de una deformación severa y un enfriamiento rápido genera una delgada capa blanca. En esta capa es fácil que queden grietas ocultas. Aunque por fuera parezca lisa, en su interior se ocultan defectos como microfisuras.

Todos estos tipos de desgaste están interconectados. Si uno se intensifica, los demás también aumentan. Por eso no se debe analizar una sola causa de forma aislada. Es necesario examinar múltiples señales en conjunto. La fuerza de corte, las vibraciones, el sonido y la temperatura son señales indicadoras del desgaste.

La inteligencia artificial procesa estas señales de manera integrada. Los sensores recopilan señales en tiempo real durante el mecanizado. El modelo utiliza estas señales para estimar el desgaste actual de la herramienta. Investigaciones recientes han aplicado el aprendizaje profundo para analizar imágenes de herramientas desgastadas y predecir directamente su grado de desgaste. Las máquinas sustituyen así la tarea en la que los operarios debían detener el equipo e inspeccionar con una lupa.

El mismo enfoque funciona también con otros metales. Un estudio predijo el desgaste de la herramienta al mecanizar una aleación de níquel llamada Hastelloy C276 mediante un modelo basado en datos. Esto es posible gracias a que diversas superaleaciones presentan patrones de desgaste similares. Estos resultados sirven también como referencia para las superaleaciones empleadas en cohetes.

Predecir el desgaste de la herramienta permite además el mecanizado no tripulado. La máquina supervisa por sí misma el estado de la herramienta. Si detecta un riesgo, reduce la velocidad o cambia la herramienta. No es necesario que una persona esté vigilando continuamente. En fábricas que mecanizan piezas durante la noche sin presencia humana, los modelos predictivos garantizan esa seguridad operativa.

Estos cuatro tipos de desgaste dependen en gran medida de la temperatura. Cuanto más caliente esté la zona de corte, más rápido avanza el desgaste. Por eso la refrigeración es fundamental. Se aplica fluido de corte para disipar el calor. En ocasiones, el fluido de corte se inyecta a alta presión directamente sobre el filo. Este método de refrigeración también se convierte en una variable de entrada para la inteligencia artificial.

En resumen, la predicción del desgaste de la herramienta funciona como los ojos del control de calidad. Permite cambiar la herramienta antes de que el desgaste sea excesivo. Hace posible corregir las condiciones antes de que se produzcan piezas defectuosas. No obstante, las señales de los sensores contienen ruido. En las vibraciones se mezcla también la oscilación propia de la máquina. Por ello, el modelo debe entrenarse adecuadamente para tolerar el ruido. El uso conjunto de múltiples sensores reduce el impacto del ruido.

7.3 Cálculo de fuerzas para un mecanizado eficiente y optimización del postratamiento

En la sección anterior se examinó por qué resulta tan difícil mecanizar estos materiales. En esta sección se analiza cómo superar esas dificultades. Se abordan tres aspectos de forma sucesiva. Primero, se estudia cómo calcular las fuerzas que intervienen en el corte. A continuación, se examina cómo varía la maquinabilidad según el postratamiento. Por último, se analiza cómo la inteligencia artificial encuentra las condiciones óptimas.

7.3.1 Fuerzas que intervienen en el corte y modelos para gestionarlas

En este apartado se examinan las fuerzas que intervienen en el corte y los modelos para gestionarlas. La fuerza de corte (cutting force) es la fuerza con la que la herramienta empuja el metal. Conocer esta fuerza de antemano permite seleccionar adecuadamente las herramientas y las condiciones operativas. Se analizan aquí la idea fundamental para calcular la fuerza y el problema de las vibraciones.

La fuerza de corte no actúa en una sola dirección. Comprende simultáneamente una fuerza de empuje frontal, una fuerza de empuje lateral y una fuerza de empuje hacia abajo. Por ello, la fuerza de corte se representa como un vector con tres componentes direccionales. Un vector es una magnitud que posee tanto magnitud como dirección. En cambio, las tensiones que actúan en el interior del material se representan mediante un tensor. Un tensor es una magnitud que engloba valores en múltiples direcciones. El vector de fuerza de corte y el tensor de tensiones son magnitudes diferentes. La magnitud de esta fuerza depende del espesor de corte y de la resistencia del material.

Durante el corte ocurren tres fenómenos al mismo tiempo: el metal se deforma, se genera calor y se produce fricción. Un buen modelo debe abordar estos tres aspectos de forma integrada. La resistencia del material cuando se somete a compresión rápida varía con la temperatura. Al calentarse, el metal se ablanda ligeramente. Esta relación se introduce en el modelo a través de tablas en lugar de fórmulas matemáticas directas.

Este modelo utiliza constantes distintas para cada material. Las constantes se determinan experimentalmente. Incluso tratándose del mismo Inconel 718, las constantes varían entre el material fabricado por adición y el forjado. Esto se debe a que el material fabricado por adición posee una microestructura diferente. Por tanto, se requieren ensayos específicos para materiales de fabricación aditiva. El modelo solo alcanza precisión cuando se acumulan suficientes datos experimentales dedicados a estos materiales aditivos.

La inteligencia artificial también asiste en los experimentos para determinar estas constantes. Existe un método para deducir inversamente las constantes a partir de las señales de corte. Se miden la fuerza y la temperatura, y se seleccionan las constantes que se ajustan a esos valores. La máquina realiza con rapidez lo que antes los operarios ajustaban manualmente. Así, el modelo se puede calibrar con prontitud para nuevos materiales.

El problema de las vibraciones también es importante. Anteriormente se mencionó el chatter (vibración autoexcitada). El chatter es una vibración en la que la herramienta y la pieza se perturban mutuamente. Si se produce una vibración, en la siguiente rotación se vuelve a cortar sobre esa marca irregular. Con ello, la vibración aumenta todavía más. Una pequeña oscilación se amplifica por sí misma progresivamente y entorpece el mecanizado.

Para evitar estas vibraciones, es necesario conocer la profundidad de corte segura. Si se corta a demasiada profundidad, la vibración se intensifica; si se corta a poca profundidad, la vibración disminuye. Este límite varía según la velocidad de rotación. Al gráfico que traza este límite se le

denomina diagrama de lóbulos de estabilidad. A partir de este diagrama, se eligen condiciones de corte libres de vibraciones.

La inteligencia artificial agiliza estos cálculos. Utilizar exclusivamente modelos físicos requiere mucho tiempo de cómputo. Cada vez que se modifican ligeramente las condiciones, es necesario resolver las ecuaciones desde el principio. La inteligencia artificial aprende previamente los resultados de los modelos físicos. Después, proporciona instantáneamente la respuesta para nuevas condiciones. A estos modelos que reemplazan cálculos computacionalmente pesados se les llama modelos sustitutos (surrogate models). Sustituto significa actuar en representación o como delegado. Los cálculos precisos se realizan previamente solo unas pocas veces. El modelo sustituto, habiendo aprendido esos resultados, se encarga de las evaluaciones posteriores.

Este método reduce drásticamente el tiempo en los entornos de producción. En el pasado, se realizaban múltiples ensayos de corte modificando las condiciones una y otra vez. En la actualidad, el modelo filtra primero las condiciones de riesgo. Solo se prueban en la práctica las pocas opciones restantes. No obstante, si el modelo físico es incorrecto, la inteligencia artificial también fallará. Por ello, resulta más seguro combinar la física con los datos.

El método que combina la física y los datos se denomina aprendizaje automático informado por la física (PIML). Este método incorpora las leyes de la física dentro de la inteligencia artificial. Aunque los datos sean escasos, las leyes físicas guían y acotan la solución. Sirve como un mecanismo de seguridad que previene predicciones absurdas. Esta idea reaparece continuamente en los capítulos siguientes de este libro. Resulta de gran utilidad en campos como la ingeniería de cohetes, donde los datos son escasos.

7.3.2 Comparación de la maquinabilidad entre material forjado, tal como fue fabricado (as-built) y tratado por HIP

En este apartado se comparan tres estados del metal. Incluso tratándose del mismo Inconel 718, sus propiedades difieren según el método de fabricación y el postratamiento. Se examinan conjuntamente el material forjado, el estado tal como fue fabricado tras la adición y el estado sometido a prensado isostático en caliente.

Veamos primero el material forjado (wrought). El material forjado es aquel que se produce prensando y compactando el metal. Comparte el mismo principio que la técnica de forja tradicional en la que se golpea el hierro. Su microestructura es fina y homogénea. Los cristales son pequeños y redondeados. Por tanto, la fluctuación de la fuerza de corte se sitúa entre las más bajas de los tres materiales. El desgaste de la herramienta es correspondientemente menor, lo que facilita enormemente su predicción.

A continuación, se encuentra el estado inmediatamente posterior a la adición. A este estado se le denomina también 'tal como fue fabricado' (as-built). Es el material tal cual sale de la impresora sin ningún postratamiento. Presenta elevadas tensiones residuales y una concentración de la dura fase de Laves. Los granos columnares también son gruesos y largos. Por consiguiente, la resistencia al corte se incrementa notablemente en comparación con el material forjado. El riesgo de fractura de la herramienta se encuentra asimismo entre los más altos de los tres materiales.

El tercero es el estado sometido a prensado isostático en caliente (Hot Isostatic Pressing, HIP). El HIP es un postratamiento que aplica simultáneamente alta temperatura y alta presión. Por lo general, se realiza a temperaturas superiores a los 1000 grados Celsius. La presión también supera los 100 megapascales. En estas condiciones, los poros en el interior del metal se comprimen y se cierran.

Es fácil de entender si se piensa en la masa de pan. Dentro de la masa hay pequeñas burbujas de aire. Si se amasa y presiona durante un tiempo prolongado, estas burbujas se expulsan y la masa se vuelve más densa. El HIP funciona de forma similar. La alta presión comprime y une los poros en el interior del

metal. La combinación de alta presión y calor moviliza los átomos, haciendo que las paredes de los poros se unan entre sí.

El HIP también transforma la microestructura. No obstante, el grado de transformación depende de la temperatura y del tiempo de mantenimiento. Si la temperatura es lo suficientemente alta, la dura fase de Laves se disuelve dentro de la matriz. Si la temperatura es baja o el tiempo es corto, parte de la fase de Laves permanece. Los granos columnares alargados también pueden conservarse según las condiciones. Al homogeneizarse la microestructura, se reducen las diferencias según la dirección y las fluctuaciones en la fuerza de corte. Por ello, el HIP se emplea combinando y ajustando sus condiciones con tratamientos térmicos de solubilización, envejecimiento y homogeneización.

Sin embargo, el HIP también tiene sus limitaciones. No puede eliminar la totalidad de los poros. Los poros conectados a la superficie y los poros atrapados en el interior se comportan de manera diferente. Las burbujas de argón atrapadas durante el proceso de fabricación aditiva vuelven a causar problemas incluso después de haberse cerrado mediante HIP. Esto se debe a que el argón es un gas insoluble en el metal. Si posteriormente se somete de nuevo a altas temperaturas, las burbujas que se habían cerrado se expanden y vuelven a abrirse. A este fenómeno se le conoce como porosidad inducida térmicamente (thermally induced porosity). Por esta razón, es fundamental reducir las burbujas de gas desde el inicio.

Tras el HIP también surgen nuevos problemas. Al volverse la microestructura más blanda, las virutas metálicas no se rompen con facilidad. Las virutas se forman de manera continua y larga, enredándose alrededor de la herramienta. Si este enredo es severo, raya la superficie mecanizada. Por ello, se debe seleccionar una geometría de herramienta adecuada para favorecer la rotura de viruta.

La inteligencia artificial organiza esta comparación en forma de datos. Recopila los datos de corte de cada estado y aprende las diferencias. Con ello determina a qué estado se aproxima una pieza nueva. También propone qué postratamiento debe aplicarse. Estudios recientes examinan ampliamente métodos para optimizar conjuntamente el HIP y el tratamiento térmico.

El HIP es un postratamiento costoso y que requiere mucho tiempo. Exige calentar y presurizar durante periodos prolongados dentro de un equipo de gran tamaño. Por eso resulta difícil aplicarlo a todas las piezas. Es necesario determinar para qué piezas el HIP es indispensable. La inteligencia artificial también ayuda en este juicio. Al analizar la información sobre los defectos, estima de antemano la efectividad del HIP.

Las investigaciones públicas también sirven como una buena guía. La NASA analizó mediante simulaciones cómo el HIP reduce los defectos del Inconel 718 fabricado por manufactura aditiva. Un estudio midió la resistencia a la fatiga tras someter a HIP piezas con una elevada porosidad inicial. La mayoría de los poros se cerraron, pero las burbujas de argón permanecieron. La acumulación de estos resultados sienta las bases para el diseño del postratamiento.

Hay una razón para tomar los materiales forjados como referencia. El material forjado es un material validado durante mucho tiempo. Se conocen bien sus propiedades y se dispone de abundantes datos. El objetivo consiste en elevar el material fabricado por manufactura aditiva al nivel del material forjado. El postratamiento es el camino para aproximarse a esa meta. La inteligencia artificial ayuda a encontrar ese camino con mayor rapidez.

7.3.3 Búsqueda de buenas condiciones de corte con pocos experimentos

En este apartado se aborda el método para encontrar buenas condiciones con un número reducido de experimentos. Las condiciones de corte son variadas. La velocidad, el avance, la profundidad y el método de refrigeración constituyen variables. Si se combinan todas estas variables, el número de experimentos se dispara. Este problema se resuelve combinando el método de Taguchi y las redes neuronales artificiales.

El método de Taguchi (Taguchi method) es una técnica para comparar numerosas condiciones con pocos experimentos. No prueba todas las combinaciones posibles. Selecciona y experimenta únicamente con combinaciones representativas. Con solo unas pocas combinaciones bien elegidas, interpreta la tendencia general. Al reducirse el número de experimentos, es posible ahorrar tanto tiempo como costes.

Con este método se reduce drásticamente el número de experimentos. Por ejemplo, cientos de combinaciones pueden reducirse a dieciocho experimentos. En cada experimento se miden la rugosidad superficial y el desgaste de la herramienta. También se mide la energía consumida durante el mecanizado. Los valores recopilados de estas diversas combinaciones se convierten en datos de entrenamiento para la inteligencia artificial.

Las redes neuronales artificiales (Artificial Neural Network, ANN) aprenden patrones a partir de estos datos. Una red neuronal artificial es una estructura de cálculo inspirada en las conexiones neuronales del cerebro humano. En la entrada se introducen las condiciones de corte y el estado del tratamiento térmico. En la salida se obtienen la rugosidad superficial y el desgaste de la herramienta. El modelo encuentra por sí mismo la relación entre la entrada y la salida y la asimila como una regla.

Una vez finalizado el entrenamiento, también predice condiciones que no han sido experimentadas. Si se introducen una nueva velocidad y un nuevo avance, informa del resultado por anticipado. A continuación, se aplica una búsqueda para encontrar las condiciones óptimas. Se eligen condiciones que dejen una superficie lisa y minimicen el desgaste de la herramienta. Cuando ambos objetivos entran en conflicto, se busca un punto de equilibrio.

También se están acumulando casos reales. Un estudio abordó el microtorneado de Inconel 718 fabricado por manufactura aditiva. Analizó cómo variaba la maquinabilidad del material sometido a tratamiento térmico. Comparó velocidades de corte de entre 30 y 60 metros por minuto. La reunión de estos datos sirve como material de entrenamiento para el modelo.

Esta combinación resulta de gran utilidad en el terreno práctico. En primer lugar, se elabora una matriz experimental reducida. Luego, se recopilan datos mediante sensores y mediciones. Por último, la inteligencia artificial aprende la relación entre las condiciones y la calidad. De este modo, es posible determinar rápidamente las condiciones para piezas nuevas.

Cuando existen múltiples objetivos, es necesario encontrar un punto de equilibrio. Se busca lograr una superficie fina, un bajo desgaste de la herramienta y un tiempo reducido. Estos objetivos entran en conflicto mutuo. Si se mecaniza con rapidez, la superficie tiende a volverse rugosa con facilidad. Por lo tanto, si se intenta mejorar solo uno de ellos, otro empeora.

En estos casos, se emplea un método que aborda múltiples objetivos de manera simultánea. Se cede un poco en uno para ganar en otro. Se recopilan las combinaciones favorables obtenidas de esta manera. Entre ellas, se elige la combinación adecuada para la pieza. La inteligencia artificial encuentra estas combinaciones con rapidez.

El objetivo real es reducir la rugosidad superficial. Las piezas de cohetes deben tener una superficie lisa para ser resistentes a la fatiga. El uso conjunto de condiciones de corte bien seleccionadas y herramientas de calidad permite acercarse a esta meta. Una superficie fina reduce los puntos de inicio de grietas. Por ello, el acabado superficial se convierte en el primer paso para garantizar la resistencia a la fatiga.

Las limitaciones también son evidentes. El modelo solo resulta fiable dentro del rango en el que ha sido entrenado. En velocidades o avances no incluidos en el aprendizaje, las predicciones pierden precisión. Si los datos son escasos, el error aumenta. Por eso, en piezas críticas, las condiciones se confirman mediante pruebas de corte reales. Se verifica si la predicción fue acertada a través de la rugosidad superficial y el desgaste de la herramienta.

La elección de la herramienta también es una variable fundamental. Para las superaleaciones se utilizan herramientas provistas de recubrimientos especiales. El recubrimiento ayuda a resistir el calor y el desgaste. También se emplean herramientas fabricadas con materiales cerámicos. La inteligencia artificial aprende qué herramienta se adapta a cada condición. Así es posible equilibrar el coste de las herramientas con la calidad del mecanizado.

El reto para el futuro consiste en ampliar los datos. Si se reúnen datos experimentales de diversas instituciones, el modelo se vuelve más robusto. No obstante, los métodos de medición varían ligeramente entre instituciones. La forma de registrar las condiciones de corte también difiere. Por ello, se requieren normas para unificar los datos. Cuando se establezcan estos estándares, aumentará la fiabilidad de las predicciones.

7.4 Fallo por fatiga de muy alto número de ciclos en álabes de turbobombas

Hasta ahora se ha abordado cómo mecanizar adecuadamente. A partir de esta sección se analiza cuánto tiempo pueden resistir las piezas. Se examina por qué los álabes rotatorios de la turbobomba de un cohete son especialmente vulnerables y qué huellas dejan cuando se fracturan.

7.4.1 Entorno severo con coexistencia de temperaturas criogénicas y altas temperaturas

En este apartado se describe el entorno severo al que están expuestos los álabes de las turbobombas. Una turbobomba (turbopump) es una bomba que impulsa combustible y oxidante a alta presión. El rotor de álabes que gira a gran velocidad en su interior se denomina impulsor (impeller). Veremos por qué las cargas que soporta esta pieza son tan particulares.

El impulsor gira a una velocidad extremadamente alta. Rota entre 30.000 y 80.000 veces por minuto. Es una velocidad más de diez veces superior a la del motor de un automóvil. Durante el giro, la fuerza centrífuga tira de los álabes hacia el exterior. Esta fuerza centrífuga constituye una carga estacionaria cuya magnitud apenas cambia si la rotación es constante.

El entorno térmico también es riguroso. El propio impulsor está en contacto directo con propelentes criogénicos. El hidrógeno líquido es sumamente frío, a más de 250 grados bajo cero. El oxígeno líquido también se encuentra cerca de los 180 grados bajo cero. El gas de trabajo caliente no está en el impulsor, sino en la turbina montada sobre el mismo eje. Este calor se transmite paulatinamente hacia el lado de la bomba a través del eje, los rodamientos y los sellos. Así, los dos extremos de un mismo eje afrontan simultáneamente una zona extremadamente fría y otra muy caliente.

La gran diferencia de temperatura genera contracción y dilatación. El lado frío se contrae y el lado caliente se expande. Esta diferencia produce tensiones internas en el componente. A esto se suma la fuerza centrífuga estacionaria como base. Sobre ella se superponen cargas vibratorias cíclicas que sacuden la pieza.

Las cargas cíclicas proceden de múltiples fuentes. Se genera una fuerza cada vez que un álabe pasa cerca de un componente estacionario. El ligero desequilibrio del eje de rotación y las pulsaciones de presión también aportan fuerzas adicionales. El arranque y la parada aplican a su vez fuertes cargas transitorias. La fuerza de paso de los álabes se repite miles de veces por segundo. De este modo, el número de ciclos se acumula a un ritmo sumamente rápido.

La fatiga con un número tan elevado de ciclos se denomina fatiga de muy alto número de ciclos (Very High Cycle Fatigue, VHCF). Los diseños convencionales suelen tomar diez millones de ciclos como referencia. Para calcular cuántos ciclos de carga repetida soporta una pieza real, se debe multiplicar la frecuencia de excitación por el tiempo de funcionamiento acumulado. Los estudios con probetas abarcan regímenes de 100 millones e incluso 1.000 millones de ciclos. En este régimen, los patrones de fatiga cambian, por lo que deben analizarse por separado.

Los ensayos de fatiga convencionales difícilmente pueden abarcar este régimen. Experimentar 1.000 millones de ciclos a una velocidad estándar tomaría demasiado tiempo. Por eso se utilizan máquinas especiales de ensayo de alta frecuencia. Estas máquinas hacen vibrar la probeta aproximadamente 20.000 veces por segundo. Incluso con una vibración tan rápida, finalizar un solo ensayo requiere varios días.

Dado que los ensayos son prolongados, los datos resultan escasos y valiosos. Cuando los datos escasean, las predicciones pierden fiabilidad. Por ello, se requieren métodos con los que la inteligencia artificial aproveche al máximo los datos disponibles. Debe generar predicciones fiables incluso con pocos datos. Esta cuestión del uso eficiente de los datos se abordará detalladamente en secciones posteriores.

La selección del material también está estrechamente ligada a este entorno. Ciertos metales tienden a volverse frágiles o quebradizos en condiciones de frío extremo. El Inconel 718 mantiene propiedades relativamente estables incluso a temperaturas criogénicas. Por ello, se utiliza ampliamente en componentes criogénicos. Las aleaciones de la serie Haynes muestran un excelente rendimiento a temperaturas más altas. El material debe elegirse en función de la ubicación en la que opere la pieza.

Incluso dentro de una misma pieza, las condiciones varían según la zona. La raíz del álabe soporta grandes esfuerzos. La punta del álabe experimenta fuertes vibraciones. Por lo tanto, es necesario identificar primero las zonas más débiles a las que les resulta difícil soportar las fuerzas. Esta zona se denomina sección crítica. La garantía de resistencia a la fatiga se basa en dicha sección crítica.

Para localizar la sección crítica también se requieren cálculos. Las fuerzas que actúan sobre la pieza se resuelven mediante ordenador. Este cálculo se conoce como análisis por elementos finitos. La pieza se divide en pequeños elementos y se calculan las fuerzas una a una. La zona donde se concentran grandes tensiones es la sección crítica. La inteligencia artificial a menudo sustituye estos cálculos para realizarlos con rapidez.

La vibración del impulsor sigue un patrón particular. El número de veces que un álabe pasa junto a un componente fijo está predeterminado. Al multiplicar este número por la velocidad de rotación se obtiene la frecuencia de vibración. A ciertas velocidades de rotación, esta vibración coincide con la frecuencia natural de oscilación de la pieza. En tal caso, la oscilación aumenta de forma considerable. Esta coincidencia se denomina resonancia. La resonancia deteriora rápidamente el componente. Por ello, el diseño se realiza evitando las velocidades de rotación peligrosas.

Garantizar la resistencia a la fatiga del impulsor es clave para la seguridad del cohete. Si se rompe un solo álabe, la turbobomba entera queda destruida. Como consecuencia, el motor se apaga o explota. Por ello, esta pieza se somete a inspecciones más rigurosas que cualquier otra. Es indispensable controlar tanto la calidad superficial como los defectos internos.

7.4.2 Fallo originado en el interior y patrón de ojo de pez

En este apartado se analiza cómo la fatiga de muy alto número de ciclos fractura los componentes. Anteriormente se mencionó que las reglas en este régimen son diferentes. Veremos en detalle de qué manera difieren. También se examinarán los patrones singulares que quedan en la superficie de fractura.

La fatiga convencional suele originarse en la superficie. La superficie es una zona propensa a sufrir imperfecciones. Al someterse a cargas repetidas, aparecen bandas de deslizamiento en la superficie. Las grietas se inician a partir de esas marcas. Por ello, un acabado superficial fino prolonga la vida útil.

La fatiga de muy alto número de ciclos es diferente. Las grietas no se originan en la superficie, sino en el interior. Cuando el número de ciclos es sumamente elevado, los defectos internos no resisten aunque la fuerza aplicada sea pequeña. Pequeños poros o inclusiones duras en el interior actúan como núcleos de iniciación. Las diminutas burbujas de argón que ni siquiera el HIP pudo eliminar constituyen un ejemplo peligroso.

Alrededor del núcleo de la grieta se forma una zona característica. En torno a este núcleo se acumulan cientos de millones de microdeformaciones. La grieta iniciada en el interior crece en un entorno aislado del aire. Ambas caras de la estrecha fisura friccionan entre sí, triturando la microestructura de forma muy fina. Esta zona finamente triturada se denomina área granular fina (Fine Granular Area, FGA). Bajo el microscopio, esta área se observa más oscura que su entorno. Esta región oscura se sitúa justo en el centro del ojo de pez.

La formación de esta región oscura requiere mucho tiempo. La mayor parte de la vida total a la fatiga se consume en esta etapa. El período previo a que la grieta crezca hasta ser visible resulta prolongado. Aunque el componente parezca intacto por fuera, el daño ya está progresando en su interior.

Este aspecto reviste gran importancia para determinar los intervalos de inspección. El daño tarda en manifestarse en el exterior. Por ello, no se debe confiar únicamente en la apariencia superficial. Es necesario examinar el interior con una periodicidad establecida. En los cohetes reutilizables, la gestión de este ciclo resulta aún más crucial. Antes de reutilizar una pieza ya utilizada, se debe verificar su interior.

A medida que la grieta se propaga a partir del núcleo, deja un patrón circular. Este patrón se asemeja al ojo de un pez. Por esta razón, se denomina ojo de pez (fish-eye). El patrón de ojo de pez es una prueba de que la grieta se originó en el interior. Al observar la superficie de fractura, este patrón revela la causa del fallo.

Al final, la grieta alcanza la superficie. En ese instante, la pieza se rompe de manera repentina. La grieta que crecía lentamente se propaga de forma súbita. Esta clase de fractura repentina, difícil de detectar con antelación, es la más peligrosa de todas.

Este mecanismo de fallo tiene un gran significado para las inspecciones. Limitarse a observar la superficie hace que se pasen por alto los peligros. Es imprescindible inspeccionar el interior. Por ello, se utilizan conjuntamente la tomografía computarizada por rayos X (X-ray CT) y los ensayos por ultrasonidos. Estas técnicas permiten observar el interior sin dañar la pieza. Tales procedimientos se conocen como ensayos no destructivos (non-destructive testing).

La inteligencia artificial se une a estas inspecciones. Toma como entrada el tamaño y la ubicación de los defectos internos detectados mediante TC. Con estos valores, predice la vida a la fatiga. Establece un orden de prioridad sobre qué defectos resultan peligrosos. Investigaciones recientes han propuesto métodos para detectar de forma no destructiva los defectos críticos en piezas fabricadas por

manufactura aditiva. La máquina identifica con antelación los defectos internos que las personas podrían pasar por alto con facilidad.

Eliminar por completo los defectos internos resulta difícil. Por tanto, se diseña de modo que el componente sea seguro incluso con la presencia de defectos. Este concepto se denomina diseño con tolerancia a defectos. Se define de antemano el tamaño máximo de defecto que puede tolerarse. Las piezas con defectos mayores a dicho límite se descartan. La inteligencia artificial genera los datos necesarios para fijar estos criterios.

Este método también se conecta con los criterios de inspección. Los equipos de inspección no pueden detectar defectos extremadamente pequeños. Existe un límite definido en el tamaño de los defectos que se pueden detectar. El diseño garantiza que los defectos que no superan este límite sean seguros. De este modo, los defectos no detectados en la inspección tampoco resultan peligrosos. Así, la inspección y el diseño cubren los vacíos mutuos y protegen la seguridad.

7.5 Garantía de fiabilidad ante la fatiga que predice conjuntamente la variabilidad

La vida a fatiga no se reduce a un número exacto. Incluso componentes bajo las mismas condiciones presentan vidas útiles diferentes. Por ello, la predicción debe incluir el rango de variabilidad. Esta sección explica la inteligencia artificial que aborda dicha variabilidad. Se analiza cómo convertir los defectos en números, los modelos que incorporan la variabilidad y cómo conectarlos con los criterios de seguridad.

7.5.1 Conversión del estado del material en números

En este apartado se aborda cómo convertir el estado del material en números. La inteligencia artificial se nutre y aprende de los números. Por ello, es necesario convertir los defectos y las superficies en números. A estos números se les denomina descriptores (descriptor). Aquí examinamos uno a uno qué números son importantes para la vida a fatiga.

En primer lugar, la forma del defecto se traduce en números. Al escanear el interior de la pieza mediante tomografía computarizada (CT), se obtiene la forma tridimensional del poro. Esta forma se organiza en un tensor. Como se mencionó anteriormente, un tensor es una caja que contiene valores en múltiples direcciones. En él se registran en qué dirección se alarga el poro y qué tan plano es.

La tomografía computarizada (CT) por rayos X comparte el mismo principio que la tomografía de los hospitales. Emite rayos X desde múltiples direcciones para observar el interior. Con esa información recopilada, genera una imagen tridimensional. Permite observar los poros internos sin cortar la pieza. Se miden uno a uno el tamaño y la forma de los poros. La información de tamaño y forma recolectada de este modo sirve como dato fundamental para la predicción de la fatiga.

El tamaño del defecto también es importante. Un método excelente para medir el tamaño es el método de la raíz cuadrada del área de Murakami (Murakami square root of area). Este método trata el defecto como si fuera una pequeña grieta. Proyecta el defecto sobre un plano perpendicular a la dirección de la carga y mide su área. Se toma la raíz cuadrada de esa área como el tamaño del defecto. Esto se debe a que dicha raíz cuadrada se relaciona estrechamente con la magnitud de la fuerza que abre la grieta. Por tanto, un poro tridimensional complejo se simplifica y se maneja mediante este único valor. Cuanto mayor sea este valor, peor será su impacto sobre la fatiga.

La ubicación del defecto también se tiene en cuenta. Aunque tengan el mismo tamaño, los defectos cercanos a la superficie son más peligrosos. En los defectos superficiales hay menos material alrededor para sostenerlos, por lo que la fuerza se concentra más. Los efectos ambientales derivados del contacto

con el aire también aceleran el crecimiento de las grietas. Los defectos profundos en el interior están rodeados de material por todas partes y resultan menos peligrosos. Por eso, se aplican factores de corrección distintos según la posición.

La rugosidad de la superficie también se convierte en números. Una superficie mecanizada presenta valles y picos microscópicos. Cuanto más profundo y afilado sea el valle, más se concentrará la fuerza. A este grado de concentración se le llama factor de concentración de tensiones (stress concentration factor). Es como cuando se hace un pequeño corte en un papel para rasgarlo fácilmente. Una pequeña incisión concentra en un solo punto la fuerza dispersa, debilitando esa zona.

Estos números se combinan y se vinculan con la vida a fatiga. Existe un gráfico que representa la relación entre la magnitud de la fuerza y el número de ciclos repetidos. Este gráfico se denomina curva S-N. La información de los defectos y la rugosidad desplaza esta curva hacia arriba o hacia abajo. Si los defectos son grandes y la superficie es rugosa, la curva desciende. Que la curva descienda significa que, bajo la misma fuerza, la pieza se romperá más rápido.

Estudios recientes han logrado resultados notables al introducir estos descriptores en la inteligencia artificial. Una investigación predijo la vida a fatiga de metales fabricados aditivamente a partir de la información de defectos medida mediante tomografía computarizada (CT). También aumentan los estudios que aplican el método de Murakami a diversos casos de fabricación aditiva. Se confirma una y otra vez que el tamaño, la ubicación y la forma de los defectos influyen enormemente en la vida útil.

Entre los descriptores también se incluye la dureza del material. La dureza indica qué tan duro es el material. En la fórmula de Murakami, una mayor dureza se asocia con una mayor capacidad para tolerar defectos. Sin embargo, la dureza por sí sola no explica completamente la tolerancia a los defectos. Los materiales sumamente duros pueden volverse incluso más sensibles a los defectos. Por tanto, es necesario evaluar conjuntamente la dureza, el tamaño del defecto, la ubicación y la relación de carga.

Las condiciones de carga también se introducen como números. Es fundamental saber cuán grande es la fuerza y en qué dirección actúa. También cambia si la fuerza proviene de una sola dirección o si se combina en múltiples direcciones. Asimismo, se incluye la temperatura, ya que las propiedades del material varían entre ambientes cálidos y fríos.

Investigaciones recientes han obtenido excelentes resultados con estos descriptores. Un estudio predijo la vida a fatiga por alto número de ciclos del Inconel 718 fabricado por manufactura aditiva. Utilizó como entradas el tamaño del defecto y la magnitud de la carga. El problema de la escasez de datos se resolvió aumentando los datos mediante inteligencia artificial. Como resultado, el coeficiente de determinación de la predicción alcanzó 0,975. El error medio fue de apenas un 1,13 por ciento. El coeficiente de determinación es un valor que muestra qué tan bien coincide la predicción con la realidad. Cuanto más cercano a 1 sea este valor, mejor se ajusta la predicción a la realidad. No obstante, esta cifra proviene de datos limitados de un solo material. Al ser un resultado obtenido con datos aumentados, resulta difícil aplicarlo directamente a otros procesos y materiales.

Sin embargo, definir los descriptores no es una tarea sencilla. Se requiere juzgar qué números incluir y cuáles excluir. Si se omite un número importante, la predicción se distorsiona. Si se introducen números innecesarios, el aprendizaje se vuelve confuso. Por ello, los expertos en materiales y la inteligencia artificial deben seleccionarlos conjuntamente.

7.5.2 Red neuronal bayesiana (BNN) que genera un rango de respuestas

En este apartado se aborda la inteligencia artificial que incorpora la variabilidad. Las redes neuronales convencionales ofrecen una única respuesta. En la predicción de la fatiga, esto no es suficiente, ya que la vida útil presenta una amplia dispersión incluso bajo las mismas condiciones. Por ello, se necesita un modelo que proporcione conjuntamente un rango de respuestas.

A este tipo de modelo se le denomina red neuronal bayesiana (Bayesian Neural Network, BNN). En una red neuronal convencional, los valores internos se fijan en un único número. La red neuronal bayesiana trata los valores internos como distribuciones. En lugar de un solo valor, posee un rango de valores. Por tanto, la salida de la predicción no es un valor único, sino que se expresa como un intervalo.

Aquí aparece con frecuencia el término gaussiano. Gaussiano se refiere a una distribución en forma de campana. Es una forma alta en el centro y baja hacia ambos lados. Resulta fácil imaginar la distribución de las calificaciones de un examen: hay muchas personas cerca del promedio y pocas hacia los extremos. Sin embargo, la vida a fatiga en sí a menudo no sigue una forma de campana; se dispersa siguiendo otras formas, como la log-normal o la de Weibull. Por eso, se aplica la suposición de forma de campana al valor logarítmico de la vida útil. Esta suposición gaussiana es también una conveniencia para simplificar los cálculos.

En esta distribución son importantes dos valores. Uno es la media. La media es el valor representativo. El otro es el ancho de dispersión. Un ancho amplio significa que la variabilidad es grande. La red neuronal bayesiana entrega ambos valores conjuntamente.

Usemos el pronóstico del tiempo como analogía de por qué la predicción tiene un margen. El pronóstico dice que la probabilidad de lluvia para mañana es del 70 por ciento. No afirma de manera tajante que sea del 100 por ciento. Esto se debe a que el futuro contiene incertidumbre. La red neuronal bayesiana también proporciona un rango de probabilidades en lugar de una sola respuesta.

El método de entrenamiento de este modelo también es especial. Como los valores internos son distribuciones, los cálculos resultan complejos. El cálculo exacto es prácticamente imposible. Por ello, se utiliza un método que lo sustituye por una distribución aproximada. Este proceso de reducir la carga computacional sustituyéndola por una distribución cercana constituye el núcleo del entrenamiento.

Una vez finalizado el entrenamiento, se extraen predicciones múltiples veces. Dado que los valores internos son distribuciones, cada vez que se extrae una muestra se obtiene una respuesta ligeramente distinta. Al reunir estas respuestas se calculan la media y la dispersión. Si las respuestas son similares entre sí, el intervalo es estrecho. Si se dispersan, el intervalo es amplio. La magnitud de este intervalo muestra qué tan confiable es la predicción.

Este método se asemeja al juicio humano. Incluso un ingeniero experimentado distingue entre la certeza y la suposición. Habla con confianza sobre piezas que conoce bien, pero con prudencia sobre componentes que ve por primera vez. La red neuronal bayesiana transmite de manera similar su nivel de certeza. De este modo, para las personas resulta más fácil confiar en los resultados y utilizarlos evaluando dicho grado de certeza.

Este enfoque tiene una ventaja adicional. Cuando el modelo encuentra condiciones nunca antes vistas, amplía el intervalo. Envía por sí mismo la señal de que no tiene suficiente información. El ser humano observa esta señal y procede con cautela. Las predicciones con un intervalo amplio se verifican con mayor detenimiento antes de usarse. Que el modelo comunique sus propios límites contribuye enormemente a la seguridad.

Este enfoque se adapta perfectamente a la predicción de la fatiga. La fatiga es un fenómeno con una gran variabilidad. Un intervalo resulta más honesto que un solo número. Múltiples investigaciones recientes siguen esta dirección. Han surgido estudios que predicen la vida a fatiga de metales aeroespaciales junto con su variabilidad. También existen investigaciones que incorporan leyes físicas para aumentar la fiabilidad.

¿Por qué es beneficioso incorporar leyes físicas? La fatiga cuenta con teorías acumuladas a lo largo de mucho tiempo. Son ecuaciones que describen la relación entre la fuerza y la vida útil. Al integrar estas fórmulas en la inteligencia artificial, las respuestas no se desvían drásticamente incluso si los datos son escasos. También han aparecido investigaciones que abordan conjuntamente cargas multiaxiales. Esta tendencia hace que la garantía de vida útil de los componentes de cohetes sea aún más sólida.

Este método también conlleva costes. Al requerir múltiples muestreos para obtener respuestas, el cálculo aumenta. El entrenamiento también resulta más complejo que el de una red neuronal convencional. Aun así, para los componentes de cohetes este coste merece la pena. Conocer la variabilidad es lo que protege la seguridad. Si ocurre un accidente por confiar ciegamente en un solo número, las pérdidas son incomparablemente mayores.

7.5.3 Descomposición de la variabilidad y su conexión con los criterios de seguridad

En este apartado se aborda cómo desglosar y analizar la variabilidad. Antes se mencionó que las predicciones tienen un margen. Este margen no es de un solo tipo; en él se mezclan dos tipos de variabilidad. Al separar ambas, queda claro qué acciones deben tomarse.

La primera es la incertidumbre epistémica (epistemic uncertainty). Esta es la variabilidad que surge porque el modelo carece de conocimiento suficiente. Se manifiesta con fuerza en condiciones donde los datos son escasos. Ante situaciones que nunca ha aprendido, el modelo muestra falta de confianza. Esta variabilidad se reduce progresivamente a medida que se incorporan más datos faltantes.

La segunda es la incertidumbre aleatoria (aleatoric uncertainty). Esta se origina en la variabilidad intrínseca del material y de las cargas. Incluso piezas bajo idénticas condiciones presentan defectos internos dispares. En las mediciones también se cuela el ruido. Al estar presente en el fenómeno mismo, esta variabilidad apenas disminuye aunque se aumenten los datos.

¿Por qué es importante distinguir entre ambas? Porque las soluciones son diferentes. Si la variabilidad epistémica es alta, deben realizarse más experimentos. Si la variabilidad aleatoria es alta, se debe añadir un mayor margen de seguridad. Si se confunden ambas, es fácil prescribir una solución errónea. Al separar los dos tipos de variabilidad, queda claro qué paso debe darse a continuación.

Ahora se conectan las predicciones con los criterios de seguridad. Los componentes de cohetes deben superar criterios de seguridad establecidos. En estos criterios se integran dos factores: uno es qué porcentaje de piezas sobrevive, y el otro es con qué grado de confianza se puede respaldar dicho valor. Por ejemplo, se establece un límite inferior con una supervivencia del 90 por ciento y una confianza del 95 por ciento. A este valor, definido conjuntamente por la tasa de supervivencia y el nivel de confianza, se le denomina admisibles del material (material allowables). Los valores base A y base B utilizados en materiales aeroespaciales son ejemplos de ello. Este valor no se puede garantizar simplemente restando la dispersión del promedio. Se calcula estadísticamente incorporando tanto el número de datos de prueba como su dispersión.

A este límite inferior se le denomina valor admisible de diseño. El componente debe resistir más tiempo que este valor admisible. A esto se le puede multiplicar adicionalmente un factor de seguridad. El factor de seguridad es un margen destinado a prevenir riesgos imprevistos. De este modo, se establecen

múltiples capas de margen para evitar accidentes. Si el margen es excesivo, la pieza se vuelve demasiado pesada, por lo que se busca un equilibrio.

De esta manera, el juicio se inclina hacia el lado de la seguridad. En condiciones donde el modelo no tiene certeza, estima una vida útil más baja. En condiciones donde tiene confianza, otorga un poco más de margen. Esta actitud es la adecuada para los componentes de cohetes. Es preferible ser algo conservador con tal de prevenir accidentes.

La certificación no concluye únicamente con la inteligencia artificial. La garantía contra la fatiga requiere de múltiples herramientas coordinadas: la predicción por inteligencia artificial, los ensayos normalizados de fatiga, los ensayos no destructivos y los factores de seguridad son todos necesarios. Los métodos de ensayo estándar también están predeterminados. Existen normativas internacionales para el modo de aplicación de fuerzas y el análisis de los datos.

El papel de la inteligencia artificial es claro: localiza rápidamente los candidatos peligrosos, señala qué áreas requieren más ensayos y ayuda a obtener abundante información con pocos experimentos. Los seres humanos aprovechan esa información para diseñar eficazmente los planes de ensayo, reduciendo en gran medida el tiempo y los costes.

La decisión final corresponde a los seres humanos y a los experimentos. La predicción de la inteligencia artificial es el punto de partida, no el final. Los datos de los ensayos reales deben respaldar las predicciones. La documentación de calidad debe registrar todo el proceso. Solo cuando se cumplen todos estos requisitos, el componente sube al cohete.

También quedan desafíos pendientes. Los datos de fatiga siguen siendo insuficientes, y los correspondientes al régimen de fatiga de muy alto número de ciclos son aún más escasos debido a que los ensayos requieren mucho tiempo. Por ello, se investigan continuamente métodos para aprender a partir de pocos datos. También existen métodos para transferir y utilizar datos de otros materiales. A medida que estos métodos avancen, las predicciones se volverán más robustas.

La calidad de los datos también es fundamental. La forma en que se midieron los defectos varía según cada conjunto de datos, al igual que el modo en que se realizaron los ensayos. Si los registros son deficientes, el modelo aprende de forma incorrecta. Por ello, es necesaria una cultura de registro meticuloso de los datos. Solo disponiendo de datos excelentes y bien documentados se pueden obtener modelos sobresalientes.

Las normas y la certificación también deben evolucionar a la par. Las reglas de certificación actuales se crearon con base en materiales convencionales. Los materiales de fabricación aditiva tienen propiedades diferentes. Por eso se necesitan nuevas reglas. También se está debatiendo cómo incorporar las predicciones de la inteligencia artificial en la certificación. Una vez que este debate se resuelva, las piezas fabricadas por adición se utilizarán de forma más amplia.

Resumen

En este capítulo se han abordado dos grandes problemas. Uno es mecanizar adecuadamente metales difíciles de cortar. El otro es saber cuánto tiempo resistirán las piezas. Ambos problemas surgen en las piezas de cohetes fabricadas mediante impresión 3D. Y en ambos problemas, la inteligencia artificial ofrece una gran ayuda.

En realidad, ambos problemas están conectados como uno solo. Solo con un buen mecanizado se obtiene una superficie lisa. Una superficie lisa es resistente a la fatiga. Es necesario eliminar los poros internos para reducir las semillas de las grietas. Al final, un buen procesamiento se traduce en una larga

vida útil. Por eso, el mecanizado y la fatiga siempre deben considerarse juntos. Este capítulo ha reunido y tratado conjuntamente estos dos problemas de mecanizado y fatiga.

En la primera parte, examinamos por qué las superaleaciones de fabricación aditiva son difíciles de mecanizar. La microestructura enfriada rápidamente tiene propiedades diferentes según la dirección. En su interior quedan atrapadas tensiones y se concentran partículas duras. Estas propiedades desgastan rápidamente las herramientas y dejan capas blancas y microgrietas en la superficie.

También vimos la manera de superar estas dificultades. Se calculan con anticipación la fuerza de corte y las vibraciones para seleccionar condiciones seguras. Mediante el prensado isostático en caliente, se reducen los poros internos y se homogeniza la microestructura. Con el método Taguchi y las redes neuronales artificiales, se encuentran condiciones óptimas con pocos experimentos. En este proceso, la inteligencia artificial reduce rápidamente los candidatos.

En la última parte, abordamos el problema de la fatiga. Los álabes de las turbobombas de cohetes están expuestos a fatiga de ciclo muy alto. En este régimen, las grietas se originan en el interior y no en la superficie. En la superficie de fractura queda un patrón circular conocido como ojo de pez. Por lo tanto, se requieren ensayos no destructivos para observar el interior sin romper la pieza.

La predicción de fatiga debe incluir también la incertidumbre. Esto se debe a que las piezas bajo las mismas condiciones tienen vidas útiles variables. Las redes neuronales bayesianas proporcionan un rango de respuestas en lugar de una sola respuesta fija. Al dividir este rango en dos tipos, se hace evidente la siguiente tarea por realizar. Se puede decidir si aumentar los datos o incrementar el margen de seguridad.

La clave es la colaboración entre la inteligencia artificial y los seres humanos. La inteligencia artificial interpreta los datos con rapidez y señala los riesgos. Reduce de forma inteligente los planes de experimentación. Sin embargo, la certificación final corresponde a los experimentos y las inspecciones. Estas dos fuerzas protegen conjuntamente las piezas de cohetes. Esta cooperación entre la inteligencia artificial y la experimentación constituye la base que sustenta un vuelo seguro.

Resumamos el contenido de este capítulo en una sola línea. Es necesario mecanizar bien, pulir bien e inspeccionar bien. Y es preciso predecir los resultados junto con su incertidumbre. La inteligencia artificial ayuda con rapidez en todas estas etapas. Con esa ayuda, los humanos fabrican piezas más seguras. El futuro de las piezas de cohetes impresas en 3D depende, en última instancia, de esta cooperación.

Fuentes y referencias relacionadas

Ojo, O. O., & Taban, E. (2023). Post-processing treatments-microstructure-performance interrelationship of metal additive manufactured aerospace alloys: A review. Materials Science and Technology, 39(1), 1-41. https://doi.org/10.1080/02670836.2022.2130530

Song, Z., Peng, J., Zhu, L., Deng, C., Zhao, Y., Guo, Q., & Zhu, A. (2025). High-Cycle Fatigue Life Prediction of Additive Manufacturing Inconel 718 Alloy via Machine Learning. Materials, 18(11), 2604. https://doi.org/10.3390/ma18112604

Yu, C., Huang, Z., Zhang, Z., Shen, J., Wang, J., & Xu, Z. (2021). Influence of post-processing on very high cycle fatigue resistance of Inconel 718 obtained with laser powder bed fusion. International Journal of Fatigue, 153, 106510. https://doi.org/10.1016/j.ijfatigue.2021.106510

ASTM International. (2021). ASTM E466-21: Standard Practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials. ASTM International, West Conshohocken, PA. https://www.astm.org/e0466-21.html

ASTM International. (2015). ASTM E739-10(2015): Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life and Strain-Life Fatigue Data. ASTM International, West Conshohocken, PA. https://www.astm.org/e0739-10r15.html

Yi, M., Tang, W., Zhu, Y., Liang, C., Tang, Z., & Yin, Y. (2023). A holistic review on fatigue properties of additively manufactured metals. arXiv preprint. (Preimpresión sin revisión por pares) https://arxiv.org/abs/2311.07046

Altıparmak, S. C., Lombardi, M., Bondioli, F., Fino, P., Biamino, S., & Ugues, D. (2025). Hot isostatic pressing for powder-based additive manufacturing of metals: State-of-the-art review and future perspectives. Journal of Materials Research and Technology, 39, 4794-4822. https://doi.org/10.1016/j.jmrt.2025.10.135

Murakami, Y. (2002). Metal Fatigue: Effects of Small Defects and Nonmetallic Inclusions. Elsevier Science, Oxford, UK.

Basquin, O. H. (1910). The exponential law of endurance tests. Proceedings of the American Society for Testing and Materials, 10(2), 625-630.

Coffin, L. F. (1954). A study of the effects of cyclic thermal stresses on a ductile metal. Transactions of the American Society of Mechanical Engineers, 76(6), 931-950.

Manson, S. S. (1953). Behavior of materials under conditions of thermal stress. National Advisory Committee for Aeronautics (NACA) Technical Note, NACA-TN-2933.

Arola, D., & Ramulu, M. (1999). An examination of the effects of surface texture on the strength, fatigue life, and stress concentration of machined components. Journal of Engineering Materials and Technology, 121(3), 376-381.

Blavette, D., Cadel, E., Fraczkiewicz, A., & Menand, A. (1999). Three-dimensional atomic-scale imaging of segregation and precipitation in superalloys. Science, 286(5448), 2317-2319. https://doi.org/10.1126/science.286.5448.2317

Wang, Q., Yao, G., & Hou, M. (2026). Fatigue life prediction and uncertainty quantification of aerospace metals: A Bayesian physics-informed neural network model. Reliability Engineering & System Safety, 266, 111724. https://doi.org/10.1016/j.ress.2025.111724

Zhang, D., Dai, W., & Ding, Y. (2025). Bayesian optimization-based physics-informed neural network for adaptive prediction of multiaxial fatigue life in metallic materials. Fatigue & Fracture of Engineering Materials & Structures, 49(1), 28-51. https://doi.org/10.1111/ffe.70101

Chang, J., Basvoju, D., Vakanski, A., Charit, I., & Xian, M. (2025). Predictive modeling and uncertainty quantification of fatigue life in metal alloys using machine learning. arXiv preprint. (Preimpresión sin revisión por pares) https://arxiv.org/abs/2501.15057

Li, L., Chang, J., Vakanski, A., Wang, Y., Yao, T., & Xian, M. (2024). Uncertainty quantification in multivariable regression for material property prediction with Bayesian neural networks. Scientific Reports, 14, 10543. https://doi.org/10.1038/s41598-024-61189-x

Vishnu, S., Kuriachen, B., & Mathew, J. (2026). Microturning response of LPBF Inconel 718: Effect of homogenization annealing. Materials and Manufacturing Processes, 41(7), 964-973. https://doi.org/10.1080/10426914.2026.2660062

Truong, T. T., Airao, J., Fattahi, S., Azarhoushang, B., Karras, P., & Aghababaei, R. (2025). Image-based machine learning model for tool wear estimation in milling Inconel 718. Wear, 571, 205865. https://doi.org/10.1016/j.wear.2025.205865

Abdullah, M., Rao, A. C. U., Ramachandran, T., Biswal, S., Chaudhary, S. P., Singh, R., Mishra, A., & Mann, V. S. (2026). A data-driven approach for high-accuracy tool wear prediction in machining Hastelloy C276. Scientific Reports, 16, 19442. https://doi.org/10.1038/s41598-026-50824-4

Medibew, T. M., Zieliński, D., Agebo, S. W., & Deja, M. (2025). Recent research progress in the abrasive machining and finishing of additively manufactured metal parts. Materials, 18(6), 1249. https://doi.org/10.3390/ma18061249

Liu, Y., Zhao, L., Huang, S., Meng, X., Zhang, Z., Lu, H., Zhou, S., & Dai, G. (2026). Recent progress in surface post-treatment techniques for additive manufactured alloy metal parts: A comprehensive review. Additive Manufacturing Frontiers, 200339. https://doi.org/10.1016/j.amf.2026.200339

Hassan, M., Hussain, M. A., & Jahan, M. P. (2026). Post-processing of additively manufactured metallic parts - A review. The International Journal of Advanced Manufacturing Technology. https://doi.org/10.1007/s00170-026-18615-3

Sommer, D., Truetsch, B., Esen, C., & Hellmann, R. (2025). Effect of hot isostatic pressing and sequenced heat treatment on the mechanical properties of hybrid additive manufactured Inconel 718 components. Journal of Manufacturing and Materials Processing, 9(11), 378. https://doi.org/10.3390/jmmp9110378

Kaletsch, A., Qin, S., Herzog, S., & Broeckmann, C. (2021). Influence of high initial porosity introduced by laser powder bed fusion on the fatigue strength of Inconel 718 after post-processing with hot isostatic pressing. Additive Manufacturing, 47, 102331. https://doi.org/10.1016/j.addma.2021.102331

Fody, J. M., & Lang, C. G. (2021). Hot isostatic pressing for reduction of defects in additively manufactured Inconel-718: Preliminary simulation approach and results. NASA Technical Reports Server, Document ID 20210015451. https://ntrs.nasa.gov/citations/20210015451

Baig, S., Jam, A., Beretta, S., Shao, S., & Shamsaei, N. (2025). Non-destructive detection of critical defects in additive manufacturing. Scientific Reports, 15, 6740. https://doi.org/10.1038/s41598-025-91608-6

Novelino, A. L. B., de Oliveira, D., Araújo, J. A., & Ziberov, M. (2026). Defects characterization of wall components fabricated via WAAM and fatigue strength estimation: A theoretical prediction approach based on Murakami's square root of area method. Progress in Engineering Science, 3(1), 100238. https://doi.org/10.1016/j.pes.2026.100238

He, P., Huang, Y., Wang, W., Guo, Y., Yu, Z., Chen, W., Li, B., & Wang, X. (2025). Deep learning driven prediction method for roughness and hardness of laser in-situ forging additive manufacturing Ti-6Al-4V. Virtual and Physical Prototyping, 20(1), e2569541. https://doi.org/10.1080/17452759.2025.2569541

Capítulo 8. Modelado de flujos reactivos gaseosos y química de combustión mediante aprendizaje automático informado por la física (PIML/PINNs) y aprendizaje de operadores (FNO)

Introducción

Al encender la cocina de gas, brota una llama azul. El gas, que actúa como combustible, se encuentra con el oxígeno del aire y genera calor. A este proceso se le llama combustión. La combustión está presente en muchos rincones de nuestra vida cotidiana. Tanto los motores de los automóviles como las calderas funcionan mediante la combustión.

Los motores de cohete generan fuerza siguiendo el mismo principio. No obstante, la escala y las condiciones son enormemente diferentes. En el interior de la cámara de combustión de un cohete hace mucho más calor y la presión es más alta. El combustible y el oxígeno se mezclan y reaccionan a gran velocidad. En su interior, el gas fluye, se arremolina y produce sonido.

Resulta difícil calcular todos estos fenómenos al mismo tiempo. Las reacciones químicas ocurren en un abrir y cerrar de ojos. En cambio, el flujo que recorre toda la cámara de combustión es más lento. Cambios rápidos y lentos se mezclan en un mismo espacio. Para que una computadora siga ambos simultáneamente, la cantidad de cálculo aumenta de forma drástica.

Hasta ahora, las ecuaciones se resolvían dividiéndolas en partes muy pequeñas. A este método se le conoce como dinámica de fluidos computacional (CFD). Es preciso, pero su coste computacional es elevado. Para resolver en detalle una cámara de combustión grande de un cohete, una supercomputadora debe funcionar durante días. Quienes diseñan llevan mucho tiempo anhelando herramientas más rápidas.

La inteligencia artificial ofrece ayuda en este punto. Una red neuronal bien entrenada reduce los cálculos con rapidez. También puede rellenar información faltante cuando hay pocos sensores. Sin embargo, resulta peligroso que la inteligencia artificial proporcione respuestas arbitrarias en problemas de combustión. Esto se debe a que, en un cohete, un pequeño error puede derivar en un accidente grave.

Por ello, surgió el método de integrar directamente las leyes de la física en el aprendizaje. A este enfoque se le denomina aprendizaje automático informado por la física (PIML). La red neuronal no se limita a imitar datos. Se la obliga a cumplir también leyes como la conservación de la masa o de la energía. Su método representativo es la red neuronal informada por la física (PINN).

Este capítulo aborda diversos métodos de dicho aprendizaje informado por la física. En primer lugar, se explican con palabras sencillas las ecuaciones que rigen la combustión. A continuación, se examinan las dificultades de cálculo provocadas por las rápidas reacciones químicas. Luego, se analiza una nueva estructura denominada red de Kolmogórov-Arnold (KAN). Asimismo, se presenta el operador neuronal de Fourier (FNO), que transforma funciones en funciones. Por último, se concluye con el caso de la reconstrucción del interior de un motor de detonación rotativa.

Planteemos aquí una analogía. Basta con imaginar dos formas de trazar un mapa. Una consiste en medir todo el terreno caminando paso a paso. Es precisa, pero requiere mucho tiempo. La otra consiste en medir solo unos pocos puntos y rellenar el resto. Es rápida, pero si se rellena sin reglas, se obtiene un mapa disparatado. El aprendizaje informado por la física consiste en incorporar las reglas de la naturaleza a este proceso de relleno.

Al terminar de leer este capítulo, se obtiene una visión general. Se comprende cómo la inteligencia artificial está transformando el cálculo de la combustión. También se distingue qué es factible en la actualidad y qué sigue siendo difícil. Las nuevas herramientas son potentes, pero no omnipotentes. Solo cuando van acompañadas de la física y de la validación se puede confiar en ellas. Mantener este equilibrio es la idea central que recorre todo este capítulo.

Este libro maneja con cautela los nombres complejos y las cifras. Entre las investigaciones recientes, existen casos que solo tienen nombres llamativos. También se mezclan cifras de rendimiento cuya base no ha sido verificada. Por ello, este capítulo selecciona y explica únicamente estudios confirmados. En el texto no se incluyen cifras no verificadas. Las investigaciones en fase inicial se indican con claridad como estudios preliminares.

8.1 Cambio de paradigma en la ingeniería de combustión computacional e inteligencia artificial

Contenido de esta sección

Esta sección examina primero por qué es difícil calcular la combustión. Observa la gran complejidad de los fenómenos que ocurren dentro de la cámara de combustión de un cohete. A continuación, analiza las limitaciones de los métodos tradicionales de cálculo. Finalmente, explica por qué ha surgido la inteligencia artificial para superar dichas limitaciones.

Aquí no se utilizan fórmulas matemáticas complejas. En su lugar, se trasladan los conceptos a situaciones de la vida cotidiana. Si se comprende esta sección, todas las secciones posteriores resultarán más sencillas. Al fin y al cabo, todo este capítulo trata sobre diversos métodos para resolver este problema.

El interior de la cámara de combustión de un cohete es un entorno extremo. La temperatura asciende a miles de grados. La presión alcanza cientos de veces la presión atmosférica. El combustible y el oxidante reaccionan con violencia en un breve lapso. Los motores que utilizan metano y oxígeno líquido constituyen un ejemplo representativo.

Estas reacciones son mucho más complejas de lo que aparentan a simple vista. No existe una única vía por la cual el metano se combine con el oxígeno para transformarse en agua y dióxido de carbono. En el proceso intermedio surgen decenas de sustancias intermedias. Estas especies, denominadas radicales, poseen una alta reactividad. Por ello, cientos de reacciones elementales se entrelazan y avanzan de forma simultánea.

Comparemos los radicales con una carrera de relevos. El testigo pasa velozmente de mano en mano. Ningún corredor retiene el testigo durante mucho tiempo. Los radicales también se generan en un instante y se transfieren a la siguiente sustancia. Este continuo intercambio impulsa la combustión. Si se pierde de vista esa velocidad, no es posible modelar adecuadamente la forma de la llama.

A esto se añade además la turbulencia. La turbulencia es un flujo en el que el gas se arremolina y se mezcla caóticamente. Es semejante al agua de un río que choca contra las rocas y se rompe en espuma blanca. El gas dentro de la cámara de combustión de un cohete se mezcla de esta misma manera. La turbulencia hace que el combustible y el oxígeno se encuentren rápidamente. Sin embargo, debido a ese movimiento irregular, el cálculo se vuelve aún más difícil.

Para resolver esta reacción con precisión, es necesario dividir el espacio-tiempo en partes sumamente finas. La zona donde se sitúa la delgada llama tiene un grosor inferior al de un cabello. En su interior, la temperatura y la concentración varían drásticamente. Si se pasa por alto esta fina capa, los cálculos fallan. Por lo tanto, para capturar dicha capa se debe disponer una malla extraordinariamente densa.

Apreciemos este problema multiescala mediante cifras. La delgada capa donde se forma la llama mide unas decenas de micrómetros. Por el contrario, la longitud total de la cámara de combustión es de varias decenas de centímetros. La diferencia de tamaño alcanza varios miles de veces. En el tiempo se presenta una disparidad similar. Para abarcar este amplio rango a la vez, el número de elementos de la malla aumenta de forma explosiva.

Entre los métodos de cálculo tradicionales se encuentran la simulación numérica directa (DNS) y la simulación de grandes remolinos (LES). La simulación numérica directa resuelve hasta el más mínimo vórtice sin omitir ninguno. La simulación de grandes remolinos resuelve únicamente los vórtices grandes y aproxima los pequeños. Ambos métodos son precisos, pero requieren una enorme cantidad de cálculo. Resolver en detalle una cámara de combustión grande lleva mucho tiempo.

Traduzcamos la idea de un alto coste computacional al lenguaje cotidiano. Es semejante a producir una película de alta definición. Si cada escena se dibuja con minucioso detalle, se tardan meses. El cálculo de la combustión es igual. Cada vez que cambian las condiciones, hay que calcular todo de nuevo desde el principio. Quien diseña desea probar decenas de condiciones. No obstante, esperar varios días cada vez que se modifican los parámetros resulta poco práctico en la realidad.

En este punto, la inteligencia artificial abre un nuevo camino. Una vez bien entrenada, la red neuronal responde con rapidez a problemas similares. Se convierte en un modelo subrogado (surrogate model) ligero que sustituye a los cálculos pesados. Un modelo subrogado es un modelo de aproximación rápida que imita el cálculo original. Gracias a él, quien diseña puede examinar múltiples condiciones en poco tiempo.

Comparemos el modelo subrogado con una aplicación de mapas. Indica el tiempo necesario sin necesidad de recorrer el camino a pie. Compara innumerables rutas en un instante. El modelo subrogado funciona de la misma manera. Estima los resultados sin ejecutar todos los cálculos reales. Gracias a ello, se pueden examinar varias opciones de diseño en poco tiempo.

Sin embargo, la aplicación de mapas también puede equivocarse en caminos desconocidos. El modelo subrogado también falla si se sale del rango en el que fue entrenado. Por ello, no se confía a ciegas en las respuestas estimadas. Ante decisiones cruciales, se vuelve a verificar mediante cálculos de alta precisión. El modelo subrogado es una brújula que orienta la dirección, no un veredicto definitivo. Esta actitud se mantiene de principio a fin en todo este capítulo.

Expliquemos el estado supercrítico con más detalle. Resulta fácil de entender tomando el agua como ejemplo. Por lo general, el agua hierve y se convierte en vapor de agua. La frontera entre líquido y gas es clara. No obstante, si la presión y la temperatura se elevan a niveles extremos, dicho límite se difumina. El combustible dentro de la cámara de combustión de alta presión de un cohete se encuentra en este estado. Calcular con exactitud esta frontera difusa entre líquido y gas resulta muy complejo.

Existen también casos reales. DeepFlame es un software que integra redes neuronales para calcular con rapidez la combustión supercrítica. El estado supercrítico se refiere a un estado de alta presión en el que la frontera entre líquido y gas se desvanece. El equipo de investigación optimizó este código para resolver a gran escala la combustión turbulenta de metano y oxígeno. Según los datos públicos, la escala de cálculo se amplió hasta el nivel de decenas de miles de millones de celdas. Esto representa un alcance considerablemente mayor que antes. Este resultado corresponde a una investigación preliminar presentada en una conferencia de computación de alto rendimiento en 2025.

En este caso, la función que asume la inteligencia artificial es clara. El cálculo de las reacciones químicas es la etapa que más tiempo consume. La red neuronal sustituye este cálculo con rapidez. A cambio,

preserva fielmente las propiedades reales del fluido. Se trata de un intento por lograr tanto velocidad como precisión. Este enfoque acelera los cálculos de combustión en cohetes a gran escala.

Examinemos por qué es importante el cálculo a gran escala. Un motor cohete real cuenta con cientos de inyectores. El inyector es el componente que pulveriza y mezcla el combustible y el oxígeno. Las llamas de estos numerosos inyectores interactúan entre sí. Con cálculos a pequeña escala resulta difícil capturar esta interacción. Solo cuando se hace posible el cálculo a gran escala nos aproximamos al comportamiento del motor real.

La inteligencia artificial se introduce en la combustión principalmente de tres formas. La primera es como un modelo subrogado rápido que sustituye a los cálculos que requieren mucho tiempo. La segunda es la reconstrucción de todo el campo a partir de mediciones dispersas. La tercera consiste en diagnosticar y controlar el estado en tiempo real. Este capítulo se centra en las dos primeras formas, mientras que el control en tiempo real se aborda en capítulos posteriores.

Veamos un ejemplo de por qué es peligroso que la inteligencia artificial vulnere la física. Supongamos que cierto modelo predice una temperatura inferior a la real. Quien diseña concluye que la pared puede resistirlo. Sin embargo, si la temperatura real es mayor, la pared se funde. Un pequeño error conduce a un grave accidente. Por eso, en el ámbito de los cohetes, respetar la física está directamente vinculado con la seguridad.

Desde luego, las limitaciones también son evidentes. La inteligencia artificial se equivoca con frecuencia fuera del rango aprendido. En los cohetes siempre surgen condiciones inéditas. Si no se respetan las leyes de la física, las predicciones se vuelven peligrosas. Por ello, todo este capítulo trata sobre los métodos que integran la física y el aprendizaje. En 2026 se publicó también una revisión que recopila el aprendizaje informado por la física para la propulsión aeroespacial. Esta tendencia demuestra que este campo está creciendo rápidamente.

8.2 Ecuaciones rectoras de flujos reactivos compresibles multicomponente y aprendizaje informado por la física

Contenido de esta sección

Esta sección explica con palabras sencillas las ecuaciones fundamentales que rigen la combustión. Hablar de ecuaciones puede sonar complejo. No obstante, el significado es simple: es la regla que establece que nada desaparece ni surge de la nada.

A continuación, se analiza cómo incorporar estas reglas en una red neuronal. Se examina la forma en que las redes neuronales informadas por la física cumplen las leyes. Estos dos conceptos constituyen la estructura fundamental de este capítulo. Todos los métodos presentados más adelante se derivan de estos dos conceptos.

8.2.1 Ecuaciones rectoras de Navier-Stokes reactivas

Las ecuaciones de Navier-Stokes son las reglas que describen el movimiento de los fluidos. Fenómenos como el fluir del agua o el soplo del aire se resuelven con estas reglas. En la combustión de cohetes se añaden además las reacciones químicas y el calor. Por ello, la cantidad de ecuaciones que deben resolverse conjuntamente se incrementa.

Para un fluido convencional sin reacciones, basta con el equilibrio de tres magnitudes: masa, cantidad de movimiento y energía. Sin embargo, en un fluido reactivo como el de la combustión, se añade la conservación de las especies químicas. Por lo tanto, estas ecuaciones engloban cuatro leyes de conservación. La primera de ellas es la conservación de la masa. La cantidad de gas que entra debe

coincidir con la que sale. El gas no se crea ni se destruye por sí solo. Este equilibrio donde la masa cuadra constituye la base de todos los cálculos.

La siguiente es la conservación de la cantidad de movimiento. La cantidad de movimiento es el producto de la masa por la velocidad. Al recibir una fuerza, la cantidad de movimiento cambia. La diferencia de presión y la viscosidad generan esta fuerza. Esa es la razón por la cual el gas es expulsado a gran velocidad en el interior del cohete.

Pensemos en la conservación de la cantidad de movimiento a través de una bicicleta. Al pedalear, la bicicleta avanza. La fuerza del pedaleo produce un cambio en la velocidad. En el cohete, la diferencia de presión actúa como esa fuerza. La viscosidad es la fricción que retiene y ralentiza el flujo. Estas dos fuerzas —la diferencia de presión y la viscosidad— alteran la velocidad del gas.

Otra es la conservación de las especies químicas. Las especies químicas son las sustancias que participan en la reacción, como el oxígeno, el metano y el agua. Cuando se produce la reacción, algunas sustancias disminuyen y otras aumentan. La velocidad de reacción determina estas cantidades de incremento y disminución. En un fluido reactivo se plantea una ecuación de conservación independiente para cada sustancia. Por consiguiente, se añaden tantas ecuaciones como número de especies químicas haya. La última es la conservación de la energía. El calor desprendido por la reacción calienta el gas e impulsa el flujo.

La conservación de la energía se puede comparar con una hucha. Si se cuadra el dinero ingresado con el gastado, se obtiene el dinero restante. En la combustión, las reacciones químicas no crean energía nueva. La energía química contenida en el combustible simplemente se transforma en calor. Ese calor calienta el gas y sale realizando trabajo. También hay calor que se escapa por radiación. La regla que equilibra todas estas entradas y salidas es la conservación de la energía.

El término compresibilidad también es importante. La compresibilidad significa que la densidad del gas cambia significativamente según la presión. En la cámara de combustión de un cohete, la presión es alta y experimenta fuertes variaciones. Por lo tanto, no se puede fijar la densidad como una constante. Esta variación continua de la densidad hace que el cálculo sea aún más difícil.

Examinemos por qué estos cuatro equilibrios son tan importantes. Un cálculo con el equilibrio roto se convierte pronto en una mentira. Si la masa aumenta por sí sola, es como si surgiera un empuje inexistente. Si la energía se fuga, la predicción de la temperatura falla. En el diseño de cohetes, tales errores conducen directamente a accidentes. Por eso, la inteligencia artificial también debe respetar obligatoriamente este equilibrio.

La velocidad de reacción química se determina mediante la ley de Arrhenius. La ley de Arrhenius es una regla según la cual la reacción se acelera drásticamente a medida que sube la temperatura. Se parece a cómo el fuego de una cerilla se propaga en un instante una vez encendido. Esta naturaleza abrupta se convierte en la raíz de los problemas de cálculo que veremos más adelante. Los diseñadores de cohetes deben predecir con precisión estas reacciones para mantener la seguridad.

La difusión y la transferencia de calor también deben resolverse de forma conjunta. La difusión es el fenómeno por el cual una sustancia se propaga desde un lugar de alta concentración hacia uno de baja concentración. Es como cuando el aroma de la cocina se extiende por la habitación. La transferencia de calor es el fenómeno por el cual el calor pasa de un lugar caliente a uno frío. En la cámara de combustión, ambos procesos ocurren simultáneamente con las reacciones químicas. Por ello, las ecuaciones se entrelazan y se resuelven como un solo bloque.

8.2.2 Función de pérdida compuesta informada por la física

El núcleo de las redes neuronales informadas por la física reside en la función de pérdida. La función de pérdida es el criterio que mide cuánto se ha equivocado la red neuronal. Una red neuronal convencional solo mide la discrepancia con los datos. La red neuronal informada por la física añade a esto el error físico.

El principio de funcionamiento es el siguiente. La red neuronal recibe como entradas la posición y el tiempo. Luego, genera la presión, la temperatura, la velocidad y la concentración en ese punto. Es correcto que estos valores satisfagan las ecuaciones de conservación. Si estos valores no cumplen las ecuaciones, reciben una penalización proporcional.

Para medir este error físico, se requiere la derivación. Aquí se utiliza la diferenciación automática (automatic differentiation). La diferenciación automática es una técnica que calcula las derivadas siguiendo el proceso de cómputo de la red neuronal. Es distinta de las diferencias finitas, que aproximan mediante diferencias en una malla. Por ello, no presenta el error de truncamiento originado en las diferencias finitas. No obstante, aún permanecen el error de aproximación propio de la red neuronal y los errores derivados de la selección de puntos de muestreo. Lo que elimina la diferenciación automática es el error generado al aproximar las derivadas.

Analicemos por qué son necesarias las derivadas. Las ecuaciones de conservación tratan sobre razones de cambio. Observan qué tan rápido cambia la presión o cómo se transforma la velocidad. Esta razón de cambio es precisamente la derivada. La variación de los valores producidos por la red neuronal se obtiene mediante derivadas. Luego, se comprueba si la tasa de cambio así calculada satisface las ecuaciones.

El cálculo tradicional aproximaba las derivadas sobre una malla. Estimaba los cambios mediante las diferencias entre puntos de malla vecinos. Con este método, si la malla es gruesa, el error aumenta. La diferenciación automática no utiliza esta aproximación. Deriva directamente la expresión matemática de la red neuronal. Por lo tanto, no es necesario construir por separado una malla densa para derivar. Sin embargo, se deben seleccionar los puntos donde se medirá el error físico. A estos puntos se les llama puntos de colocación. Si los puntos de colocación se distribuyen de forma demasiado dispersa, la solución se vuelve inestable. Que esté libre de mallas no significa que sea precisa en cualquier lugar.

En la función de pérdida se incluyen múltiples términos. La discrepancia con los datos es un término. Los errores de las ecuaciones de masa, cantidad de movimiento, energía y especies químicas constituyen términos individuales. Las condiciones de frontera y los estados iniciales también se introducen como términos. El entrenamiento progresa ponderando estos múltiples términos y reduciéndolos de manera conjunta.

Hacer esto ofrece una gran ventaja. Aunque los datos sean escasos, las leyes físicas llenan los vacíos. En el laboratorio solo se suelen utilizar sensores de presión en paredes o sondas ópticas. Con estas mediciones dispersas no es posible conocer el flujo completo. Si se establecen las leyes físicas como reglas, se pueden cubrir los intervalos intermedios.

Esta propiedad destaca en los problemas inversos. El cálculo convencional introduce las causas para obtener resultados. Un problema inverso observa los resultados para encontrar las causas. Un ejemplo es deducir la fuente de calor interna observando únicamente la temperatura en la pared. Las redes neuronales informadas por la física son eficaces en este tipo de problemas inversos porque gestionan las mediciones y la física juntas dentro de una sola función de pérdida.

Esta capacidad es invaluable en las pruebas de cohetes. Es difícil colocar muchos sensores dentro de la cámara de combustión. Al ser un entorno caliente y reducido, los sensores no pueden resistir. Por ello,

solo se miden unos pocos puntos desde el exterior. Reconstruir el estado interno a partir de estos pocos puntos es una tarea fundamental. La red neuronal informada por la física tiende un puente entre las mediciones y el estado interno.

En esta función de pérdida subyace la dificultad de la ponderación. El problema radica en cómo determinar la importancia de cada término. Si se le da un peso excesivo a un término, los demás se ignoran. Si solo se persigue el error físico, se pierden los datos. Si solo se persiguen los datos, se infringe la física. La forma de hallar este equilibrio es un tema central de investigación.

Comparemos esta ponderación con sazonar la comida. Si hay demasiada sal queda salado, y si hay demasiado azúcar queda dulce. Si alguno se excede, el sabor se arruina. Lo mismo ocurre con los diversos términos de la función de pérdida. Si un término es excesivamente dominante, el entrenamiento se sesga hacia un lado. Por esta razón, también se investigan métodos para ajustar los pesos de forma automática.

El estudio que dio a conocer ampliamente este concepto por primera vez es el artículo de Raissi et al. de 2019. Posteriormente, se extendió ampliamente al campo de la combustión. En 2025, se publicó también una revisión exhaustiva que sistematiza el aprendizaje informado por la física para la combustión. Esta revisión clasifica los temas en reacciones químicas, flujos reactivos y otros problemas de combustión. Demuestra que el aprendizaje informado por la física se ha consolidado como una corriente principal.

También debemos examinar con honestidad las limitaciones. Las redes neuronales informadas por la física encuentran bien la solución adaptada a una condición específica. Sin embargo, cuando las condiciones cambian, a menudo se debe volver a entrenar. Todavía resulta difícil abarcar un amplio rango de condiciones de una sola vez. En presencia de ondas de choque intensas o frentes de llama abruptos, el entrenamiento también puede volverse inestable. Las investigaciones para subsanar estas debilidades continúan hoy en día.

8.3 Dinámica de reacciones rígidas y aprendizaje físico con restricciones de conservación

8.3.1 Rigidez numérica en reacciones químicas y el colapso del aprendizaje estándar

La rigidez surge cuando ciertos cambios son sumamente rápidos y otros son lentos. Esta situación es común en la química de la combustión. Las sustancias intermedias, como los radicales, se forman y desaparecen en una fracción de segundo. En cambio, la temperatura o los componentes principales cambian de manera relativamente lenta.

Veamos qué tan grande es esta diferencia temporal. Las reacciones rápidas progresan en escalas de nanosegundos (milmillonésimas de segundo). Los flujos lentos se mueven en escalas de milisegundos (milésimas de segundo). La diferencia entre ambas escalas supera el millón de veces. Esta enorme brecha se convierte en la raíz de la rigidez que desestabiliza los cálculos.

Surgen problemas si los cambios rápidos y lentos se resuelven con el mismo intervalo. Si el intervalo se ajusta a las reacciones rápidas, el cálculo se prolonga indefinidamente. Si se ajusta al flujo lento, se omiten las reacciones rápidas y el cálculo diverge. El análisis numérico tradicional también ha luchado durante mucho tiempo con este punto. Por ello, el análisis numérico tradicional ha abordado esto mediante métodos especiales de integración temporal.

Esta discrepancia temporal se asemeja a la fotografía. Imaginemos capturar el aleteo de un colibrí y el paso de una tortuga en una sola imagen. Si nos ajustamos al colibrí, debemos accionar el obturador incontables veces. Si nos ajustamos a la tortuga, el colibrí sale borroso. Es difícil capturar ambos en un mismo cuadro. Lo mismo ocurre con las reacciones rápidas y los flujos lentos de la combustión.

Las redes neuronales informadas por la física estándar suelen fallar aquí. Esta red neuronal intenta reducir todos los errores a la vez. Los grandes errores generados por las reacciones rápidas dominan el entrenamiento. En consecuencia, no logra aprender adecuadamente el equilibrio de las variables lentas. A este fenómeno en el que el entrenamiento se sesga hacia un lado se le denomina patología de gradientes.

Ilustremos esta patología. Supongamos que una voz estridente monopoliza una reunión. Las palabras de las personas tranquilas quedan sepultadas. La reunión se inclina solo hacia un lado. El gran error de las reacciones rápidas es esa voz estridente. Por ello, el equilibrio de las variables lentas queda relegado a un segundo plano en el entrenamiento.

Aquí se ocultan dos causas adicionales. Una es la propiedad de las redes neuronales de aprender los cambios bruscos con retraso. La red neuronal aprende primero las tendencias suaves y posterga las variaciones pronunciadas. A esta propiedad se le llama sesgo espectral. La variación abrupta en el momento de la ignición es precisamente esa parte pronunciada. La otra es el problema de no respetar la causalidad temporal. En la combustión, el instante previo es la causa del instante posterior. Sin embargo, el entrenamiento estándar intenta ajustar todos los instantes temporales al mismo tiempo. Al hacerlo, el orden de causa y efecto se distorsiona. En casos severos, colapsa en una solución trivial donde no ocurre ninguna reacción. Por ello, han surgido métodos para dividir el tiempo y aprender secuencialmente, o para aplicar ponderaciones distintas a cada especie química.

Otro problema es el colapso de las leyes de conservación. Al ponderar múltiples términos de forma simultánea, el equilibrio se rompe. La suma de las concentraciones de las especies químicas siempre debe ser constante. El número de átomos también debe ser igual antes y después de la reacción. Si estas reglas se rompen, surgen valores absurdos como concentraciones negativas.

Analicemos por qué las concentraciones negativas son un grave problema. En la realidad no existe una cantidad negativa de materia. El agua no puede ser de menos una taza. Si el cálculo produce tales valores, la etapa siguiente se desmorona. El cálculo de la velocidad de reacción se descontrola o se detiene. Por lo tanto, se debe evitar a toda costa que las concentraciones caigan en valores negativos.

Existen varios estudios iniciales que abordaron este problema. El modelo Stiff-PINN de Ji et al. en 2021 es un ejemplo representativo. Este estudio introdujo un tratamiento especial para manejar la química rígida. Las redes neuronales informadas por la física multiescala surgieron en una época similar. Estos estudios confirmaron la gran barrera que representa la rigidez para el entrenamiento.

Hay una razón por la cual esta barrera se eleva aún más en los cohetes. La combustión de metano y oxígeno cuenta con numerosas rutas de reacción. Las sustancias involucradas alcanzan decenas de especies. A medida que aumentan las sustancias, la rigidez también se incrementa. Por ello, la combustión de cohetes se sitúa en el extremo más difícil de los problemas rígidos. La necesidad de imponer la física con firmeza es, por tanto, apremiante.

8.3.2 Combinación de restricciones de conservación e integración temporal

La clave para resolver los problemas rígidos radica en imponer la física con mayor solidez. Existen principalmente dos formas de imponer la física. Una son las restricciones blandas (soft constraints), que penalizan la infracción de las reglas. Añadir un término de error físico a la función de pérdida pertenece a este enfoque. La otra son las restricciones duras (hard constraints), que se incrustan en la estructura para que las reglas no puedan infringirse en absoluto. Cuanto más severa sea la rigidez, más se combinan ambos métodos.

La primera herramienta consiste en conectar un integrador temporal probado con la red neuronal. El método de Runge-Kutta es una técnica estándar para resolver ecuaciones diferenciales dividiendo el

tiempo en pequeños pasos. No obstante, se debe elegir el tipo adecuado. Los métodos explícitos, que proyectan los valores directamente hacia adelante, son débiles ante problemas rígidos. Para resolverlos de forma estable, el paso de tiempo debe dividirse de manera excesivamente fina. Por ello, para los problemas rígidos se emplean métodos de integración estables de tipo implícito. Al incorporar en la red neuronal esta estructura de integración de estabilidad comprobada, el aprendizaje profundo hereda la estabilidad acumulada a lo largo del tiempo por el análisis numérico.

La segunda herramienta es la proyección que fuerza la conservación. No se utilizan directamente las concentraciones generadas por la red neuronal. En primer lugar, los valores se transforman para que sean siempre positivos. Luego, se reajustan para que su suma sea constante. Finalmente, se ajustan los valores para que se conserve el número de átomos. A través de este proceso, no se producen en absoluto valores que violen la física.

Comparemos la proyección con un guardia en la puerta. Sin importar el valor que produzca la red neuronal, se examina ante la puerta. Si no se ajusta a las reglas, se corrige al valor válido más cercano. Los números negativos se convierten en positivos. Si la suma se desvía, se reajusta. De este modo, los valores que salen al exterior siempre cumplen las reglas.

Este enfoque presenta diferencias con respecto al método blando. El método blando solo impone una penalización si se infringen las reglas. La penalización guía el entrenamiento, pero no puede impedir la infracción. El método duro bloquea estructuralmente la infracción misma. Incluso si el entrenamiento es imperfecto, la física se respeta. Por tanto, es un método idóneo para los cohetes, donde la seguridad es primordial.

Un ejemplo real es CRK-PINN. Este estudio es un caso en el que las ecuaciones de velocidad de reacción de combustión se resolvieron mediante redes neuronales informadas por la física. Los autores son Shihong Zhang, Qi Zhang y Baosen Wang. Este artículo fue revisado por pares y publicado en la revista Combustion and Flame. Este método utiliza restricciones blandas que introducen las leyes físicas como términos en la función de pérdida. Incluyó la ley de acción de masas y la ley de Arrhenius en la pérdida. La conservación de la entalpía, la conservación de elementos y la conservación de la suma de concentraciones se impusieron de la misma forma. A diferencia de la proyección dura anterior, este enfoque induce la física mediante penalizaciones.

El rendimiento también ha sido publicado. Los cálculos se aceleraron en el sistema de reacción de hidrógeno y aire. Se abordaron diez especies químicas y veintiuna reacciones. Fue entre seis y catorce veces más rápido en comparación con los cálculos estándar. Al integrarse con cálculos de flujo reactivo real, también resultó entre dos y cinco veces más rápido. Este es un ejemplo en el que se respetaron las leyes físicas al tiempo que se obtuvo velocidad de cálculo.

Examinemos qué significa acelerar varias veces. El cálculo de reacciones químicas es la parte que más tiempo consume en el análisis de combustión. Una gran proporción del tiempo de cómputo total se gasta aquí. Si esta parte se acelera diez veces, el tiempo total se reduce considerablemente. Un cálculo que tomaba varios días se reduce a un solo día. En esa misma medida, el ciclo de modificar el diseño y volver a probar se vuelve más rápido.

Conviene aclarar aquí que no se deben confundir las restricciones blandas con las duras. CRK-PINN es un método blando que induce la física mediante la pérdida. Difiere en naturaleza de la proyección dura, que fuerza la corrección de los valores. Ambos métodos pueden utilizarse juntos o por separado. Las restricciones blandas son flexibles, pero no impiden por completo las violaciones físicas. Las restricciones duras previenen las violaciones, pero vuelven la estructura más compleja. Determinar cuál usar y en qué medida es una decisión de diseño.

La lección que ofrece este enfoque es sencilla y clara: no descarta el conocimiento de análisis numérico ya validado. Métodos como el de Runge-Kutta se han perfeccionado durante décadas. La inteligencia artificial se apoya en este conocimiento para potenciar sus capacidades. Es una forma en la que el conocimiento físico tradicional y el nuevo aprendizaje no compiten, sino que se ayudan mutuamente.

Sin embargo, las limitaciones también son evidentes. Este método se ha validado dentro del sistema de reacciones analizado. La combustión de metano y oxígeno en un cohete es más compleja. La cantidad de sustancias y reacciones es mucho mayor. Para aplicarlo a un nuevo sistema de reacciones, es necesario volver a entrenar y validar. También existe un compromiso en el que, cuanto más estrictamente se impone la física, más pesado se vuelve el cálculo.

8.4 Modelado de flujos reactivos basado en redes de Kolmogorov-Arnold

8.4.1 Fundamentos de la asignación de funciones a las conexiones y funciones de activación adaptativas

Primero, recordemos la estructura de una red neuronal convencional. Una red neuronal se compone de nodos y líneas de conexión que los unen. En los nodos se coloca una función de activación fija. En las conexiones solo se asigna un número llamado peso. El aprendizaje consiste en ajustar estos valores de peso.

Expliquemos primero el término función de activación. La red neuronal recibe una entrada y la emite a través de cierta curva. Esta curva es la función de activación. Las redes neuronales convencionales fijan la forma de esta curva. Superponen varias curvas fijas para crear relaciones complejas. El aprendizaje mantiene la forma de la curva y solo modifica la intensidad de la superposición.

Las redes de Kolmogorov-Arnold invierten esta disposición. Colocan las funciones en las conexiones y no en los nodos. Esa misma función se modifica a través del aprendizaje. En lugar de usar una forma fija, modelan la curva de manera flexible. Esta idea proviene de un teorema matemático demostrado hace mucho tiempo.

Ese teorema es el teorema de representación de Kolmogorov-Arnold. Este teorema establece que una función compleja puede representarse como la suma de funciones unidimensionales simples. Esta estructura expresa la función de cada conexión mediante una curva flexible. Dicha curva puede cambiar de forma localmente. Por ello, es posible seleccionar y refinar minuciosamente solo las zonas donde los cambios son abruptos, como en una llama.

Traslademos este teorema al juego de bloques de construcción. Por compleja que sea una forma, es la combinación de unos pocos bloques básicos. Una gran estructura se divide en piezas pequeñas y se vuelve a conectar. Del mismo modo, una función compleja se puede descomponer en la suma de piezas unidimensionales simples. Las redes de Kolmogorov-Arnold adoptan esta descomposición como su estructura. Las matemáticas clásicas se han convertido en la base de una nueva red neuronal.

El material utilizado para crear estas curvas son las B-splines. La B-spline es un método para unir suavemente múltiples segmentos de curvas. Pertenece a la misma familia de herramientas que utilizan los diseñadores de automóviles para trazar curvas suaves. Cada segmento influye únicamente en su propio intervalo. Por lo tanto, modificar una zona no altera las demás.

Esta localidad ofrece ventajas para el aprendizaje. Permite dividir con densidad solo el estrecho intervalo donde se sitúa la llama, mientras que el resto de las zonas amplias se deja con menor densidad. Es una forma de concentrar los recursos donde se necesitan. Esto se denomina refinamiento de malla. De este modo, se ahorra cómputo al tiempo que se obtiene la precisión requerida.

Se comprende por qué esta propiedad resulta ventajosa. Permite observar con mayor claridad qué entrada influye en qué salida. Es posible representar con precisión únicamente los intervalos necesarios. También resulta sencillo examinar y analizar la función visualmente. Es adecuada para problemas con variaciones locales severas, como la combustión.

Ilustremos la diferencia entre nodos y aristas con un ejemplo cotidiano. Una red neuronal convencional equivale a superponer varias lentes de forma predeterminada. La forma de las lentes es fija y solo se ajusta la intensidad de la superposición. Las redes de Kolmogorov-Arnold tallan la propia forma de la lente. Así, crean la refracción deseada con mayor libertad. En cambio, tallar la lente requiere más trabajo.

Por supuesto, esta estructura no es una solución universal. El aprendizaje puede ser más complejo. Cuanto más finamente se divida la curva, mayor será el número de valores que se deben ajustar. Según el problema, una red neuronal convencional puede resultar mejor. Conviene considerar esta nueva estructura como una opción más.

8.4.2 Capacidad de representación del frente de llama y reducción de parámetros

El frente de llama es una región delgada donde se concentran las reacciones. Allí, la temperatura se eleva drásticamente en una distancia corta. El gas a temperatura ambiente alcanza miles de grados en un instante. Las concentraciones también cambian de manera abrupta. Resulta difícil expresar computacionalmente los cambios tan pronunciados de este estrecho intervalo.

Las redes neuronales convencionales son vulnerables a cambios tan repentinos. Con funciones de activación fijas, resulta difícil trazar capas pronunciadas de forma suave. Si se fuerzan los ajustes, aparecen errores similares a ondulaciones en los alrededores. También tienen la propiedad de aprender lentamente la información de alta frecuencia. Debido a estas características, los resultados del cálculo del frente de llama se vuelven borrosos.

Comparemos estas ondulaciones con el dibujo de una bandera. Supongamos que dibujamos un poste recto únicamente con un pincel suave. Junto a los bordes afilados del poste aparecerán patrones ondulados. En el cálculo también surgen patrones espurios junto al frente de llama afilado. Estos patrones son falsos y no existen en la realidad. Tales patrones espurios distorsionan las predicciones con señales irreales.

Las redes de Kolmogorov-Arnold pueden mostrar fortalezas en este aspecto. Trazan mejor las capas pronunciadas mediante curvas flexibles. Permiten seleccionar y dividir con densidad solo el estrecho intervalo donde se ubica la llama. Al dejar con menor densidad las zonas innecesarias, reducen el desperdicio. Esta capacidad de adaptación para redistribuir los recursos se ajusta bien a los cálculos de combustión.

Un ejemplo real es ChemKANs. Los autores son Koenig, Kim y Deng. Esta investigación se publicó en 2025 en la revista Physical Chemistry Chemical Physics. Integró la cinética de reacción y el flujo de las leyes termodinámicas en redes de Kolmogorov-Arnold. El objetivo es aprender modelos de química de la combustión y acelerar los cálculos. Sin embargo, lo que validó este estudio no fue un frente de llama distribuido en el espacio, sino el problema de un reactor químico donde las sustancias reaccionan a lo largo del tiempo. La reconstrucción de frentes de llama espaciales o de ondas de choque en la cámara de combustión de un cohete sigue siendo una tarea pendiente.

Los resultados publicados son prácticos. Este modelo funcionó con una cantidad muy reducida de parámetros ajustables. Aprendió adecuadamente sin sobreajuste incluso con datos que contenían ruido. Soportó hasta un quince por ciento de ruido. También mostró resultados aproximadamente el doble de

rápidos en comparación con los cálculos estándar. El mismo equipo de investigación presentó además un estudio preliminar posterior para aprender las velocidades de reacción en función de la presión.

Aclaremos qué significa ser resistente al ruido. En las mediciones reales siempre se mezcla ruido, ya que los sensores no son perfectos. Un modelo vulnerable al ruido llega a memorizar incluso ese ruido. Esto se denomina sobreajuste. En cambio, un modelo resistente al ruido extrae y aprende únicamente las reglas reales.

Comparemos el sobreajuste con el estudio para un examen. Un estudiante que solo memoriza las respuestas se equivoca cuando cambian las preguntas. Un estudiante que comprende los principios resuelve también problemas nuevos. Un modelo resistente al ruido es como el estudiante que aprendió los principios. Los datos de telemetría e instrumentación de cohetes contienen mucho ruido. Por ello, esta propiedad de resistencia al ruido resulta valiosa en la instrumentación de cohetes.

Los puntos fuertes de este modelo deben describirse dentro del alcance verificado. Se han confirmado tres aspectos: la reducida cantidad de parámetros ajustables, la resistencia al ruido y una velocidad aproximadamente dos veces mayor en comparación con los cálculos estándar. Las cifras que prometen un rendimiento superior a este aún requieren validación. Por ello, este libro se limita a exponer únicamente este alcance confirmado.

¿Qué implicaciones tienen las redes de Kolmogorov-Arnold para los cohetes? El frente de llama también se forma como una capa delgada dentro de la cámara de combustión del cohete. Es necesario trazar adecuadamente esta capa delgada para predecir de forma precisa el calor y la presión. Si hay menos parámetros, el cálculo se aligera. Un modelo ligero resulta idóneo para su uso en cálculos en tiempo real. Por esta razón, esta estructura atrae la atención como candidata para el diagnóstico en tiempo real.

Examinemos la trascendencia de tener pocos parámetros. Los parámetros son el número de valores numéricos que ajusta el modelo. Si esta cantidad es grande, se requieren muchos datos para el aprendizaje. El cálculo también se vuelve pesado y consume una gran cantidad de memoria. Realizar la misma tarea con menos parámetros ofrece múltiples ventajas. Incluso permite su ejecución en computadoras pequeñas a bordo de un cohete.

También deben considerarse las limitaciones. Estos casos se validaron en sistemas de reacciones relativamente simples. Si las especies químicas aumentan a decenas, las conexiones se incrementan de forma exponencial. En consecuencia, la memoria y el tiempo de entrenamiento pueden aumentar significativamente. Por este motivo, continúan las investigaciones para aligerar aún más la estructura. La combustión turbulenta a alta presión en cohetes es mucho más compleja. Para aplicar esta estructura al diseño real de motores, aún queda camino por recorrer. Se requieren más códigos abiertos, experimentos comparativos y validaciones independientes. Por ahora, es más apropiado considerarla como una estructura candidata prometedora.

8.5 Operador neuronal de Fourier que transforma funciones en funciones

8.5.1 Principio de aprendizaje de operadores e independencia de la malla

Un operador es una regla que va de una función a otra función. Como la definición puede parecer difícil, ilustrémosla con un ejemplo. Supongamos que ingresamos la distribución de flujo en la entrada de la cámara de combustión. Como resultado, obtenemos la distribución de temperatura en toda la cámara de combustión. La regla que transforma esta imagen de entrada en una imagen de temperatura es el operador.

Comparemos el mapeo entre funciones con la cocina. Supongamos que ingresamos una lista completa de ingredientes y obtenemos la fotografía del plato terminado. Si modificamos ligeramente los

ingredientes, el plato también cambia. El aprendizaje de operadores aprende esta transformación global. En lugar de limitarse a los valores individuales de cada ingrediente, asimila las relaciones de conjunto. Por eso ofrece respuestas rápidas incluso cuando cambian las condiciones.

Una red neuronal convencional transforma unos pocos números en otros pocos números. El aprendizaje de operadores transforma una imagen en otra imagen. Por ello, procesa todo el campo de flujo de una sola vez. Una investigación pionera que amplió esta idea es DeepONet. El operador neuronal de Fourier también es un método que pertenece a este aprendizaje de operadores.

Analicemos por qué esta diferencia es tan importante. Si se debe recalcular todo desde cero para cada condición, el proceso resulta lento. En cambio, una vez aprendido el operador, responde rápidamente a múltiples condiciones. Incluso si cambian las condiciones de entrada, no es necesario volver a entrenarlo. En el diseño se prueban innumerables cambios de condiciones. Por ello, el aprendizaje de operadores llama la atención como herramienta de diseño.

La clave del operador neuronal de Fourier reside en la transformada de Fourier. La transformada de Fourier es una técnica que descompone formas complejas en la suma de múltiples ondas. Las ondas de baja frecuencia capturan las grandes estructuras del flujo. Las ondas de alta frecuencia contienen los detalles finos. Esta red neuronal aprende principalmente los componentes de baja frecuencia para representar con rapidez el flujo general.

Este enfoque tiene sus propias ventajas. La red neuronal trata el flujo como una suma de ondas. Sintetiza el flujo general mediante unas pocas bajas frecuencias para captarlo con rapidez. Resulta ventajoso para abarcar de un vistazo las relaciones en un dominio amplio. Sin embargo, este método no es una técnica numérica que resuelva directamente las derivadas de las ecuaciones. Al conservar solo las frecuencias bajas y recortar las altas, el cálculo se aligera. No obstante, dicho recorte puede difuminar partes pronunciadas como ondas de choque o frentes de llama. Por lo tanto, se debe tener precaución en las zonas donde los detalles son cruciales.

Comparemos este desenfoque con una fotografía. Una foto fuera de foco tiene bordes borrosos y las líneas nítidas aparecen difuminadas. Cuanto más se recortan las frecuencias altas, más se produce este tipo de distorsión. El límite afilado de la llama puede desdibujarse. Por ello, en estos casos se requiere un tratamiento específico que preserve las frecuencias altas.

Una propiedad destacada de este método es la independencia de la malla. Este modelo se define independientemente de la densidad de la malla computacional. Tras entrenarse con datos de baja resolución, puede aplicarse a una malla más fina. Esto se denomina superresolución sin ejemplos previos (zero-shot super-resolution). Se asemeja a restaurar una fotografía de baja resolución a una de alta resolución. No obstante, se trata de una propiedad teórica y de resultados observados en experimentos específicos; la precisión en una nueva resolución no queda garantizada automáticamente.

Las redes neuronales convencionales no pueden hacer esto. Las dimensiones de entrada y salida están fijadas desde el principio. Si se cambia la malla, es necesario construir un nuevo modelo. El operador neuronal de Fourier es diferente: trata el flujo como una función y no como una malla. Por ello, utiliza el mismo modelo aunque cambie la resolución.

Comparemos la transformada de Fourier con una radio. Una radio sintoniza diferentes emisiones dividiéndolas por frecuencia. Las bajas y altas frecuencias contienen informaciones distintas. El operador neuronal de Fourier también divide el flujo por frecuencias. El flujo principal se recoge en las bajas frecuencias, mientras que los detalles finos se hallan en las altas frecuencias. Esta red neuronal selecciona las bajas frecuencias para capturar con rapidez el flujo general.

La ventaja que ofrece la independencia de la malla en los cálculos de cohetes es notable. En las fases iniciales de desarrollo, se evalúan rápidamente varias alternativas con una malla gruesa. Tras seleccionar la mejor opción, se verifica con precisión usando una malla fina. En este proceso, no es necesario volver a entrenar el modelo. Al utilizar el mismo modelo en ambas resoluciones, se reducen simultáneamente el tiempo y los costos.

Sin embargo, esto no ocurre de forma automática en todos los problemas. Resulta complejo en zonas donde los valores presentan discontinuidades abruptas, como en las ondas de choque. También debe tenerse precaución con los frentes de llama delgados y la turbulencia cerca de las paredes. En tales regiones, la información de alta frecuencia es crucial. Solo tras un tratamiento específico y su debida validación se obtienen resultados confiables.

8.5.2 Operadores neuronales para flujos de combustión

La combustión es un problema complejo en el que se entrelazan el calor, el flujo y la química. Prosiguen las investigaciones orientadas a adaptar los operadores neuronales a este desafío. Diversos equipos ensayan diferentes enfoques. Aquí examinaremos los casos recientes ya validados centrándonos en sus conceptos.

Una de las líneas de trabajo consiste en aprender la química rígida mediante operadores. El operador neuronal de Fourier extendido (EFNO) es un ejemplo de ello. Los autores son Wong, Li, Zhang, Chen y Zhou. Esta investigación buscó resolver la química rígida incluso en condiciones no entrenadas. Se evaluó en la autoignición de hidrógeno en cero dimensiones y en llamas de chorro de hidrógeno y amoníaco en tres dimensiones. Mostró una excelente capacidad de generalización incluso en estados no incluidos en el entrenamiento. Este estudio sobre el operador extendido se publicó en la revista Combustion and Flame.

Desglosemos el término autoignición. Si el combustible y el oxígeno se calientan, se encienden por sí solos. Es un fenómeno en el que arden por sí mismos sin necesidad de un fósforo. La química de este instante es sumamente rápida y de gran rigidez. El operador extendido intentó aprender este momento difícil. Es un problema que suele usarse como banco de pruebas para la química rígida.

Examinemos por qué es importante el operador extendido. Por lo general, las redes neuronales solo funcionan bien en las condiciones en las que fueron entrenadas. Si la temperatura o la presión varían aunque sea un poco, fallan. El operador extendido intentó derribar esta barrera. En las pruebas, logró generalizar hasta cierto punto a condiciones no aprendidas. Esto demuestra su potencial para aplicarse también en nuevas condiciones de cohetes.

Otra dirección es aprender por completo el flujo turbulento reactivo tridimensional. Un estudio abordó esto mediante un operador neuronal de Fourier al que se añadió una estructura U-Net. La U-Net es una arquitectura de red neuronal que maneja simultáneamente la imagen general y los detalles. Reprodujo adecuadamente las características de frecuencia de la velocidad, la temperatura y la densidad en flujos turbulentos reactivos compresibles. Predijo el flujo con mucha mayor rapidez que los cálculos convencionales. Estos resultados demuestran el potencial de los modelos subrogados. Gran parte de las investigaciones de esta línea se encuentra todavía en una fase inicial previa a la revisión por pares.

Expliquemos también qué significa reproducir bien las características de frecuencia. En un flujo se mezclan vórtices grandes y pequeños. Cada tamaño de vórtice contiene una determinada cantidad de energía. El criterio clave radica en si se representa correctamente esta distribución de energía. Si solo se acierta en el flujo principal y se pasan por alto los vórtices pequeños, el resultado queda incompleto. Un buen operador reproduce ampliamente esta distribución tanto del flujo principal como de los pequeños vórtices.

La tercera dirección consiste en abordar la transferencia de calor por radiación mediante operadores. Una llama caliente también pierde calor en forma de luz. Esto se denomina transferencia de calor por radiación. Zhao y sus colaboradores propusieron un operador neuronal para resolver la transferencia de calor por radiación en incendios. Este estudio es una investigación preliminar centrada en incendios y no en cohetes. Por ello, no se extrapola como una demostración en cohetes, sino que se toma únicamente como un caso de referencia adyacente.

Sintamos la transferencia de calor por radiación a través de una fogata. Al pararse frente a una fogata, el rostro se siente caliente. La luz transporta el calor antes de que el aire circundante se caliente. La llama dentro de la cámara de combustión de un cohete también emite calor a través de la luz de esta manera. Esta radiación calienta las paredes y enfría la llama. Por tanto, para lograr una predicción precisa, es necesario incluir también este calor por radiación.

Incluir la radiación en los cálculos resulta complejo. La luz se propaga en todas direcciones y se entrecruza. La intensidad varía según la dirección. El intento de aprender este flujo complejo mediante operadores surgió de la investigación sobre incendios. Para trasladarlo a los cohetes, es necesario reajustar las condiciones y validarlo. Con todo, constituye un caso preliminar significativo que señala el rumbo futuro.

Estos casos demuestran ventajas comunes. Al aprender de los resultados de cálculos de alta fidelidad, se convierten en modelos subrogados rápidos. Permiten reconstruir cálculos de baja resolución en formatos de alta resolución. Se convierten en herramientas auxiliares para estimar el campo completo en situaciones con pocos sensores. Además, pueden utilizarse para explorar el espacio de diseño en poco tiempo.

También existe un punto de encuentro donde el aprendizaje de operadores se une con el aprendizaje informado por la física. Los operadores neuronales son rápidos, pero no siempre respetan la física. A estos se les pueden imponer restricciones de conservación como funciones de pérdida. Se trata de un intento de obtener simultáneamente la velocidad del operador y la fiabilidad de la física. Esta combinación vuelve a aparecer más adelante en el caso de la detonación rotatoria. Por consiguiente, estas dos herramientas no compiten entre sí, sino que forman una pareja complementaria.

Conviene señalar aquí cómo deben interpretarse estos casos. Este libro no incluye tablas comparativas de rendimiento de modelos específicos. Esto se debe a que las cifras de los estudios preliminares varían considerablemente según las condiciones. En su lugar, se explica el papel que desempeñan los operadores neuronales mediante casos verificados. Se da más peso a la utilidad del método que a las cifras.

También surgen casos de cálculos a gran escala. El equipo de investigación de DeepFlame incorporó redes neuronales para resolver la combustión supercrítica a gran escala. Los autores son Guo, Mao, Liu, Tan, Jia y Chen. Según los datos públicos, calcularon la combustión turbulenta de metano y oxígeno a una escala de decenas de miles de millones de celdas. Esto supera con creces las capacidades previas. Es una investigación preliminar orientada a la combustión de cohetes con más de cien inyectores.

Para aplicarlos al diseño real, existen numerosos elementos de verificación. Es necesario comprobar si se respeta la conservación de la masa y la energía. Hay que confirmar si la posición de la onda de choque es correcta. Se debe vigilar que la concentración de especies químicas no caiga en valores negativos. Solo al superar estas validaciones se convierten verdaderamente en herramientas. Los modelos subrogados son compañeros que asisten al cálculo de precisión, no sustitutos.

8.6 Reconstrucción de la interfaz de combustión por onda de choque en motores de detonación rotatoria

8.6.1 Desafíos en el análisis de motores de detonación rotatoria

El motor cohete de detonación rotatoria (RDRE) utiliza una cámara de combustión anular. En su interior, una onda de detonación gira velozmente describiendo un círculo. La detonación es una combustión explosiva en la que la onda de choque y la combustión se desplazan unidas. Esta onda viaja a varios miles de metros por segundo, varias veces más rápido que el sonido.

Este método tiene la ventaja de que la combustión misma eleva la presión. Los motores convencionales aumentan la presión mediante compresores o bombas antes de quemar el combustible. En la detonación, la onda de choque comprime instantáneamente el gas mientras lo quema. Por tanto, la presión se eleva por sí misma durante la combustión. Esto se conoce como combustión con ganancia de presión. Sin embargo, todavía se requiere presión de inyección para suministrar el combustible y el oxígeno. Aunque la presión aumente dentro de la cámara de combustión, se producen pérdidas en la inyección, el mezclado y el escape. Por ello, la ganancia a nivel de todo el motor varía según las condiciones. Dado que existe la posibilidad de aumentar la eficiencia si se utiliza adecuadamente, diversas instituciones de todo el mundo se han sumado a su desarrollo.

Examinemos la diferencia entre la detonación y la combustión ordinaria. La combustión ordinaria se propaga como un fuego que se extiende lentamente. La detonación se propaga como una explosión encabezada por una onda de choque. La onda de choque comprime y calienta el aire en un instante. Justo detrás de ella, el combustible se quema de inmediato. Esta estructura en la que la onda de choque y la llama están unidas genera una enorme presión.

El problema radica en que el interior es extremadamente complejo. La onda de choque frontal y la zona de reacción exotérmica están unidas entre sí. Se entrelazan ondas de choque oblicuas, superficies de contacto y capas de mezcla no quemadas. Estas estructuras interactúan en escalas de microsegundos (millonésimas de segundo). Por esta razón, se encuentra entre los problemas más difíciles de la mecánica de fluidos.

Comparemos esta interacción con un accidente de tráfico. Varios automóviles se ven involucrados simultáneamente en una intersección estrecha. El cambio en un solo punto se propaga en un instante a toda la zona. Dentro de la detonación, las ondas de choque y la llama se entrelazan de manera similar. Una pequeña perturbación altera la onda por completo. Por eso resulta tan difícil predecir y controlar el interior de este motor.

La instrumentación es aún más frustrante. Lo que se puede observar en las pruebas en tierra es limitado. Las señales de presión en la pared exterior de la cámara de combustión son prácticamente todo. En ocasiones, se capta la luz de la llama con cámaras de ultraalta velocidad. Esta luz se denomina quimioluminiscencia. Es la luz emitida espontáneamente por moléculas específicas al reaccionar. Sin embargo, tales mediciones no permiten conocer directamente el estado tridimensional interno.

Conviene entender también por qué se desarrollan los motores de detonación rotatoria. Los motores cohete ordinarios queman combustible a presión constante. La combustión por detonación quema aumentando la presión por sí misma. Existe la posibilidad de obtener mayor empuje con la misma cantidad de combustible. También surge la oportunidad de hacer el motor más corto y liviano. Por ello, este método se considera un candidato clave para la tecnología de propulsión de próxima generación.

Comparemos la dificultad de este motor con una cámara fotográfica. Supongamos que se toma una fotografía de un trompo que gira velozmente. Si el obturador es lento, el trompo se verá borroso y

distorsionado. La onda de detonación es mucho más rápida que un trompo. Por muy rápida que sea la cámara, solo capta la apariencia externa. Para ver el interior, es necesario rellenar las partes vacías mediante el cálculo.

Sintamos la velocidad de la onda de detonación a través de los números. Esta onda viaja a varios miles de metros por segundo. Es una velocidad superior a la de una bala. Tarda menos de un milisegundo en dar una vuelta completa a la cámara de combustión anular. En ese breve lapso, la presión y la temperatura fluctúan drásticamente. El cambio es tan rápido que el ojo humano no puede seguirlo en absoluto.

Por esta razón, los modelos de reconstrucción resultan indispensables. Se necesitan herramientas para recuperar el campo de flujo completo a partir de mediciones dispersas. Es preciso estimar en poco tiempo la distribución de presión, temperatura, velocidad y especies químicas. Este es precisamente el punto donde convergen el aprendizaje informado por la física y el aprendizaje de operadores. Los métodos presentados en este capítulo unen sus fuerzas aquí.

8.6.2 Reconstrucción y gemelos digitales mediante la fusión de datos y física

La primera herramienta de reconstrucción es la técnica de superresolución. Esta recupera mediciones dispersas o gruesas en forma de campos detallados. El estudio de 2026 realizado por Yao y Zhang intentó esto mediante una red generativa antagónica con consistencia de ciclo (CycleGAN). Una red generativa antagónica es una estructura donde la parte generadora y la parte discriminadora aprenden compitiendo entre sí. Se asemeja a la forma en que un falsificador y un perito se entrenan mutuamente. Este método se entrenó con datos de cálculo de motores que utilizan hidrógeno y aire. Demostró reconstruir el campo de flujo grueso haciéndolo ocho veces más detallado. Este resultado corresponde a una investigación publicada en una revista académica.

Veamos mediante una imagen por qué se necesita la superresolución. Al ampliar una fotografía antigua, los bordes se vuelven pixelados y rugosos como escalones. La tecnología de superresolución restaura con suavidad esta imagen rugosa. En los cálculos de combustión también se obtienen resultados con mallas gruesas. Al reconstruir este campo grueso en un campo detallado, la información aumenta. Es una forma de ahorrar costos computacionales mediante cálculos gruesos mientras se recuperan los detalles.

La siguiente herramienta es la fusión de multifidelidad. Esta entrelaza cálculos económicos de baja fidelidad con cálculos costosos de alta fidelidad. La idea consiste en obtener estimaciones precisas incluso con pocos datos de alta fidelidad. Un estudio denominado Cheap2Rich aplicó esto a motores de detonación rotatoria. Los autores son Bao, Zayak, Powers, Raman y Kurtz. Esta investigación de multifidelidad es todavía un estudio preliminar previo a la revisión por pares.

Comparemos las ventajas de la multifidelidad con la economía del hogar. Si solo se utilizan ingredientes costosos, el presupuesto resulta ajustado. Se utiliza una base con ingredientes económicos y se reservan los ingredientes de alta calidad únicamente para las partes esenciales. Con el cálculo ocurre lo mismo. Se establece el marco general mediante cálculos de baja fidelidad. Luego, solo las partes cruciales se refinan mediante cálculos de alta fidelidad. De este modo, es un método para alcanzar una alta precisión a un bajo costo.

En esta fusión también se incorpora el concepto de asimilación de datos. La asimilación de datos es un método que ajusta continuamente los cálculos con las mediciones reales. Es similar a cómo los pronósticos meteorológicos reciben observaciones para corregir sus predicciones. En los motores de detonación rotatoria también se reciben señales de los sensores. Con estas señales se corrigen

continuamente los cálculos. De este modo, las predicciones del modelo se aproximan más al estado real del motor.

Otra herramienta son los modelos generativos que reflejan las observaciones. Un modelo generativo es una red neuronal que genera por sí misma escenas verosímiles. Al suministrarle observaciones dispersas como condición, dibuja el campo completo. Aunque cambien las ubicaciones de observación, existe la posibilidad de utilizarlo sin necesidad de reentrenar. Wang y sus colaboradores propusieron tal intento para la reconstrucción de cámaras de combustión de detonación rotatoria. Por ello, resulta útil en situaciones donde las observaciones son escasas debido a la falta de sensores.

La última herramienta es la combinación de operadores neuronales y redes neuronales informadas por la física. El operador neuronal aproxima rápidamente el campo completo. La red neuronal informada por la física lo sujeta para que cumpla las leyes de conservación. Al combinar ambos métodos, se obtiene un estimador rápido y físicamente consistente. Cuando entra la señal del sensor, reconstruye el interior en poco tiempo.

Los resultados de esta reconstrucción dan paso al gemelo digital. El gemelo digital es una tecnología en la que un modelo computacional sigue el estado del dispositivo real. A partir del campo de flujo reconstruido, se pueden calcular la carga térmica en las paredes y la eficiencia de combustión. Permite detectar con anticipación el riesgo de ablación térmica en los inyectores. Asimismo, se vincula con el control para gestionar la inestabilidad de la combustión en poco tiempo.

Comparemos la inestabilidad de la combustión con un columpio. Si se empuja un columpio al compás del ritmo, oscila cada vez con mayor amplitud. Las fluctuaciones de presión dentro de la cámara de combustión también aumentan si entran en resonancia con el ritmo de la llama. Si estas vibraciones crecen, el motor puede destruirse. Por ello, es necesario detectar las vibraciones rápidamente y romper el ritmo. La reconstrucción y el control en tiempo real se encargan de detectar y mitigar estas vibraciones.

No obstante, aquí también se deben tener muchas precauciones. La mayoría de estos estudios de reconstrucción se encuentra en una fase preliminar. Las cifras que prometen grandes aceleraciones o alta precisión dependen estrictamente de las condiciones. Por esta razón, este libro no incluye dichas cifras en el cuerpo del texto. Esta sección se explica únicamente mediante conceptos de diseño y casos verificados.

Examinemos por qué es útil el gemelo digital. El interior de un motor real no se puede observar directamente durante las pruebas. El gemelo digital estima el estado interno a partir de las señales de los sensores. Detecta indicios peligrosos con mayor rapidez que los humanos. Puede enviar señales antes de que el problema se agrave. En cohetes reutilizables, esta alerta temprana resulta sumamente valiosa.

La autoactualización también conlleva peligros ocultos. Si un modelo se autocorrige basándose en sus propias predicciones, los errores pueden acumularse. Podría conducirse a sí mismo en una dirección errónea. Un estudio de 2026 señaló este problema y propuso un gemelo digital autoactualizable estabilizado por la física. Dicho estudio demostró el concepto en problemas térmicos generales y no en cohetes. Es un método en el que las leyes físicas sujetan las riendas de la actualización.

Analicemos por qué se necesita un limitador. La inteligencia artificial a veces produce valores absurdos. Si ese valor se introduce directamente en el control, resulta peligroso. El limitador evita que los valores sobrepasen el rango de seguridad. Es una valla que preserva los límites físicos fijados por los seres humanos. En el control autónomo, este tipo de valla es absolutamente indispensable.

Para aplicarlo al control real de motores, se requiere una mayor preparación. Es necesario acumular suficientes datos de pruebas a alta velocidad. Debe complementarse con validaciones independientes. También deben incorporarse limitadores que garanticen la seguridad. Esta tendencia demuestra que la tecnología aún está madurando. Con todo, la dirección es clara: la reconstrucción que combina datos y física se está consolidando.

Resumen

Este capítulo ha abordado diversos caminos para incorporar la inteligencia artificial a los cálculos de combustión. El punto de partida fue un único problema. En la combustión de cohetes, la química rápida y el flujo lento se entrelazan, lo que genera una gran carga computacional. El cálculo tradicional es preciso pero lento. La inteligencia artificial es rápida, pero resulta peligrosa si viola la física.

La respuesta a esto es el aprendizaje automático informado por la física. La red neuronal no se limita a imitar datos. Cumple también con leyes como la conservación de la masa y la energía. Las redes neuronales informadas por la física son el método representativo. Aunque los datos sean escasos, la física llena los vacíos. Esta fortaleza se manifiesta con gran claridad en los problemas inversos.

En este punto, conviene resumir la relación entre la inteligencia artificial y la física. La inteligencia artificial es fuerte en encontrar reglas a partir de datos. La física contiene leyes verificadas a lo largo del tiempo. Depender de solo una de ellas es insuficiente. Si solo se usan datos, se viola la física. Si solo se usa la física, la carga de cálculo es enorme. Por ello, emplear ambas juntas constituye el núcleo de todo este capítulo.

En los problemas rígidos, es necesario imponer la física con mayor rigor. A menudo, las penalizaciones no son suficientes. Se obtiene estabilidad al incorporar integradores temporales en la red neuronal. Mediante proyecciones que fuerzan la conservación, se evitan valores absurdos. CRK-PINN es un ejemplo de restricción blanda que impone la física en la función de pérdida.

Revisemos también las ecuaciones de los flujos reactivos multicomponente. El armazón lo componen cuatro balances: masa, cantidad de movimiento, especies químicas y energía. Este equilibrio es la regla que establece que nada se crea ni se destruye de la nada. El aprendizaje informado por la física introduce esta regla en la función de pérdida. Mediante diferenciación automática, mide con precisión los residuales de las ecuaciones. Gracias a ello, reconstruye el panorama completo incluso a partir de datos dispersos.

Las redes de Kolmogórov-Arnold presentan una nueva arquitectura. Colocan funciones flexibles sobre las conexiones. Pueden representar mejor los frentes de llama empinados. ChemKANs demostró este potencial en la química de la combustión. Con pocos valores, produjo resultados robustos frente al ruido.

Los operadores neuronales de Fourier mapean funciones en funciones. Pueden aprender a baja resolución y aplicarse a alta resolución. Se convierten en modelos sustitutos rápidos y herramientas de reconstrucción. La investigación que combina operadores extendidos y U-Net ha mostrado resultados en flujos reactivos. En campos adyacentes también han surgido estudios de operadores dedicados a la transferencia de calor por radiación. No obstante, todavía se debe tener precaución con las ondas de choque y los frentes de llama delgados.

El motor de detonación rotativa es el escenario donde convergen estos métodos. El objetivo es la reconstrucción para recuperar el interior a partir de mediciones dispersas. Se emplean la superresolución, la multifidelidad y la combinación de operadores con redes neuronales físicas. Sus

resultados conducen a gemelos digitales y control en tiempo real. Aún se encuentra en una etapa inicial y requiere mayor validación.

Resumamos los métodos de este capítulo en una sola frase. Todos son caminos para usar la física y los datos en conjunto. La física aporta confiabilidad y los datos aportan velocidad. Es difícil abordar la combustión de cohetes con solo uno de los dos. La respuesta está en el punto donde ambas fuerzas se encuentran. Este libro ha señalado ese lugar a través de casos verificados.

Conviene resumir también la naturaleza de los casos utilizados en este capítulo. Las redes neuronales informadas por la física de Raissi y los operadores neuronales de Fourier de Li son investigaciones fundamentales ampliamente citadas. CRK-PINN y ChemKANs son artículos de revistas científicas revisados por pares. DeepFlame, los operadores extendidos y la reconstrucción multifidelidad son investigaciones recientes. Gran parte de ellas son aún preprints en fase inicial. Que sean investigaciones iniciales significa que existe margen para que cambien en el futuro.

Vale la pena recordar también la actitud que atraviesa este capítulo. Este libro ha seleccionado los casos priorizando la evidencia. Los modelos que solo sonaban convincentes por su nombre se redujeron al nivel de conceptos. Las cifras de rendimiento no confirmadas no se incluyeron en el texto principal. Se dejó en claro que las investigaciones iniciales son investigaciones iniciales. Esto se hizo para ayudar al lector a distinguir entre hechos y expectativas.

Los desafíos futuros también son claros. Se necesitan modelos que se ajusten bien a una amplia gama de condiciones de una sola vez. Deben manejar de manera estable las ondas de choque y los frentes de llama delgados. Deben acumularse más casos validados con datos de motores reales. Los limitadores que garantizan la seguridad también deben evolucionar a la par. A medida que se resuelvan estos desafíos, las herramientas se acercarán más a la aplicación en campo.

La lección que deja este capítulo es clara. La inteligencia artificial transforma radicalmente la velocidad de los cálculos de combustión. Sin embargo, solo es confiable cuando la física y la validación van de la mano. La evidencia confirmada es más importante que los nombres llamativos de los modelos. Sobre este equilibrio, la tecnología de propulsión de próxima generación avanza paso a paso. Cuando la inteligencia artificial y la física se dan la mano, ese paso se vuelve firme.

Fuentes y referencias relacionadas

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3, 422-440. https://doi.org/10.1038/s42254-021-00314-5

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier Neural Operator for Parametric Partial Differential Equations. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.2010.08895

Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2023). Neural Operator: Learning Maps Between Function Spaces With Applications to PDEs. Journal of Machine Learning Research, 24, 1-97. https://www.jmlr.org/papers/v24/21-1524.html

Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence, 3, 218-229. https://doi.org/10.1038/s42256-021-00302-5

Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. SIAM Review, 63(1), 208-228. https://doi.org/10.1137/19M1274067

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljacic, M., Hou, T. Y., & Tegmark, M. (2025). KAN: Kolmogorov-Arnold Networks. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.2404.19756

Ji, W., Qiu, W., Shi, Z., Pan, S., & Deng, S. (2021). Stiff-PINN: Physics-Informed Neural Network for Stiff Chemical Kinetics. The Journal of Physical Chemistry A, 125(36), 8098-8106. https://doi.org/10.1021/acs.jpca.1c05102

Zhang, S., Zhang, C., & Wang, B. (2024). CRK-PINN: A physics-informed neural network for solving combustion reaction kinetics ordinary differential equations. Combustion and Flame, 269, 113647. https://doi.org/10.1016/j.combustflame.2024.113647

Koenig, B. C., Kim, S., & Deng, S. (2025). ChemKANs for combustion chemistry modeling and acceleration. Physical Chemistry Chemical Physics, 27(33), 17313-17330. https://doi.org/10.1039/D5CP02009C

Koenig, B. C., & Deng, S. (2025). Kolmogorov-Arnold Chemical Reaction Neural Networks for learning pressure-dependent kinetic rate laws. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2511.07686. https://arxiv.org/abs/2511.07686

Weng, Y., Li, H., Zhang, H., Chen, Z. X., & Zhou, D. (2025). Extended Fourier Neural Operators to learn stiff chemical kinetics under unseen conditions. Combustion and Flame, 272, 113847. https://doi.org/10.1016/j.combustflame.2024.113847

Wu, J., Wang, X., Zhang, G., Liu, J., Li, X., Zhang, Y., Zhang, H., Lyu, J., Wang, B., & Wu, Y. (2025). Physics-informed machine learning for combustion: A review. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2509.03347. https://arxiv.org/abs/2509.03347

Zhang, Z., Li, Z., Wang, Y., Yang, H., Peng, W., Teng, J., & Wang, J. (2026). Prediction of three-dimensional chemically reacting compressible turbulence based on implicit U-Net enhanced Fourier neural operator. Acta Mechanica Sinica, 42(7), 725274. https://doi.org/10.1007/s10409-025-25274-x

Guo, Z., Mao, R., Liu, L., Tan, G., Jia, W., & Chen, Z. X. (2025). Deep Learning-Enabled Supercritical Flame Simulation at Detailed Chemistry and Real-Fluid Accuracy Towards Trillion-Cell Scale. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC25). arXiv:2508.18969. https://doi.org/10.1145/3712285.3759850

DeepModeling. (2026). DeepFlame development repository. https://github.com/deepmodeling/deepflame-dev

Jiao, A., Jiang, W., Lu, X., Wang, Y., & Lu, L. (2026). Nested Fourier-enhanced neural operator for efficient modeling of radiation transfer in fires. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2604.13919. https://arxiv.org/abs/2604.13919

Bao, Y., Zajac, J., Powers, M., Raman, V., & Kutz, J. N. (2026). Cheap2Rich: A Multi-Fidelity Framework for Data Assimilation and System Identification of Multiscale Physics -- Rotating Detonation Engines. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2601.20295. https://arxiv.org/abs/2601.20295

Yao, S., & Zhang, W. (2026). Super-resolution reconstruction of high-fidelity rotating detonation flow fields based on cycle-consistent generative adversarial networks. Physics of Fluids, 38(4), 041704. https://doi.org/10.1063/5.0333031

Wang, X., Li, Z., Zhang, Y., Wen, H., & Wang, B. (2025). Observation-guided generative flow field reconstruction of rotating detonation combustor. Physics of Fluids, 37(11), 116110. https://doi.org/10.1063/5.0300394

Karkadakattil, A. (2026). A Physics-Stabilized Self-Updating Digital Twin Framework Using Physics-Informed Neural Networks for Thermal Field Prediction. Applied Computer Systems, 31(1), 30-40. https://doi.org/10.2478/acss-2026-0003

Shateri, A., Yang, Z., Yan, Y., Paul, M. C., & Xie, J. (2026). AI-Powered Surrogate Modelling for Multiscale Combustion: A Critical Review and Opportunities. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2604.25617. https://arxiv.org/abs/2604.25617

Mousavi, M., Caldwell, C., Baltes, J., Aljasem, M., & Lee, B. J. (2025). Physics-Informed Neural Networks in Clean Combustion: A Pathway to Sustainable Aerospace Propulsion. arXiv preprint (Preimpresión sin revisión por pares). arXiv:2509.08094. https://arxiv.org/abs/2509.08094

Poinsot, T., & Veynante, D. (2012). Theoretical and Numerical Combustion (3rd ed.). R.T. Edwards, Inc., Philadelphia, PA.

Capítulo 9. Optimización del diseño inverso topológico de canales de refrigeración regenerativa, orificios de refrigeración por película e inductores

Introducción

Si agarramos una olla caliente con la mano en verano, nos quemamos. Por eso fabricamos las asas de las ollas con madera o plástico. Es para evitar que el calor llegue a las manos. Los motores de cohete se enfrentan al mismo dilema. En el interior del motor, el combustible se quema y genera un gas extremadamente caliente. Este gas está lo bastante caliente como para fundir incluso el metal. Sin embargo, las paredes que contienen ese gas caliente son de metal. Proteger esta pared es un gran desafío en el diseño de motores de cohete.

Los cohetes resuelven este problema con ingenio. No queman inmediatamente el combustible que va a entrar al motor. Primero lo hacen fluir a través de delgados conductos practicados en el interior de la pared caliente. Al pasar por la pared, el combustible absorbe su calor. La pared se enfría y el combustible se calienta. El combustible calentado entra luego a la cámara de combustión y se quema. Es un método que protege la pared y precalienta el combustible a la vez. A esto se le llama refrigeración regenerativa (regenerative cooling). Se denomina regenerativa porque reutiliza el combustible calentado en lugar de desecharlo.

La refrigeración regenerativa es un método antiguo. Se ha utilizado desde los primeros cohetes. Sin embargo, era difícil diseñar los canales de refrigeración con libertad. Esto se debía a que el método de mecanizado del metal limitaba la forma de los canales. Recientemente, dos factores han eliminado esta restricción. Uno es la impresión 3D en metal. El otro es la inteligencia artificial, que encuentra rápidamente buenos diseños. La unión de ambos ha llevado el diseño de refrigeración a una nueva fase.

Este capítulo aborda las estructuras de gestión térmica de los motores de cohete. Examina los canales de refrigeración dentro de las paredes y los orificios de refrigeración por película, que forman una fina película refrigerante sobre la pared. También trata el inductor (inducer), el álabe frontal de la bomba que impulsa el combustible. Las tres estructuras plantean problemas de equilibrio entre el calor y el flujo. Si se intenta enfriar la pared con más fuerza, aumenta la resistencia que obstaculiza el flujo. Si se intenta reducir la resistencia, la refrigeración se debilita. Encontrar este punto de equilibrio únicamente con manos humanas resulta difícil.

Aquí es donde entra la inteligencia artificial. La inteligencia artificial que se trata en este capítulo no es la que dibuja imágenes o escribe textos. Es una herramienta que encuentra rápidamente buenas respuestas en espacios de diseño complejos. Algunas inteligencias artificiales sustituyen los lentos cálculos de flujo. Otras eligen por sí mismas qué experimento realizar a continuación. Otras reducen la forma compleja de los canales de refrigeración a un pequeño conjunto de números. Este capítulo explica cada una de esas herramientas de manera sencilla y detallada.

Conviene señalar por qué es importante lo que se expone en este capítulo. Las áreas de un motor de cohete donde surgen problemas con mayor frecuencia son la gestión térmica y las bombas. Si la pared se funde o la bomba vibra, el motor se detiene. Por eso, al desarrollar un nuevo motor, se invierte mucho tiempo y dinero en estas partes. Incluso si el diseño se puede realizar solo un poco más rápido y un poco más seguro, la ayuda es enorme. La inteligencia artificial aporta su contribución precisamente en este punto. Explora un amplio espacio de diseño más rápido que los humanos y selecciona buenos candidatos.

Hay un punto que debe aclararse de antemano. El diseño de refrigeración de motores de cohete es un campo con amplios márgenes de seguridad. Si la pared se funde una sola vez, el motor se destruye por completo. Por lo tanto, no se confía ciegamente en la respuesta proporcionada por la inteligencia artificial para su uso directo. La inteligencia artificial se utiliza para encontrar buenos candidatos con rapidez. La decisión final siempre se valida mediante cálculos de alta precisión y pruebas de combustión reales. Los casos de este capítulo también deben leerse sobre la base de ese principio.

9.1 Cómo los cohetes protegen las paredes calientes de la cámara de combustión

Lo que aborda esta sección es cómo los motores de cohete protegen sus paredes calientes. Primero se examina por qué la cámara de combustión de un cohete está tan caliente. A continuación, se explica qué es la refrigeración regenerativa y por qué resultaba difícil de fabricar. Por último, se observa cómo la impresión 3D en metal está transformando este diseño. Esta sección sirve de base para todo el contenido que sigue.

En la cámara de combustión principal de un motor de cohete, el combustible y el oxígeno se encuentran y arden. La temperatura del gas generado en ese momento supera holgadamente los 3.000 grados. El punto de fusión del acero es de unos 1.500 grados. El gas de combustión está muy por encima del punto de fusión del acero. Además, la presión dentro de la cámara de combustión también es alta. En los motores de alto rendimiento, la presión interna supera las cien veces la presión atmosférica. Un gas caliente y a alta presión empuja y calienta desde el interior la delgada pared metálica.

La alta presión también supone una gran carga. La elevada presión dentro de la cámara de combustión empuja la delgada pared hacia el exterior. Es igual a cuando se infla un globo con fuerza y su superficie se tensa. A esto se suma una gran diferencia de temperatura, con el interior caliente y el exterior frío. El lado caliente tiende a expandirse, mientras que el lado frío tiende a permanecer igual. Esta discrepancia genera un gran esfuerzo en la pared. En otras palabras, la presión y el calor castigan la pared al mismo tiempo. Por ello, la pared debe ser delgada y, al mismo tiempo, enfriarse bien y ser resistente.

El lugar donde el calor que incide sobre la pared alcanza su punto máximo es la garganta de la tobera. La tobera es un componente con forma de trompeta que expulsa gas hacia atrás para generar empuje. La parte de esa trompeta donde la anchura se estrecha se denomina garganta (throat) de la tobera. En este punto, el gas se acelera hasta alcanzar la velocidad del sonido. La temperatura del gas en sí desciende ligeramente al atravesar la tobera. Aun así, en la garganta de la tobera el calor que entra a la pared alcanza su valor máximo. Esto se debe a que el gas rápido roza con fuerza la pared estrecha. La cantidad de calor que ingresa a la pared por unidad de superficie se denomina flujo térmico (heat flux). El flujo térmico en la garganta de la tobera es sumamente grande. Es como si el calor de decenas de estufas de gran potencia se concentrara en un área pequeña del tamaño de la palma de una mano.

Imaginemos cuán grande es este flujo térmico. El calor emitido por un quemador de una estufa de gas doméstica se extiende sobre el área de una palma. En la garganta de la tobera, el calor de decenas de quemadores se concentra en un área equivalente. Ese calor se vierte sobre una sola capa de delgada pared metálica. La diferencia de temperatura entre la cara interior y la cara exterior de la pared alcanza varios cientos de grados. Esta enorme diferencia deforma el metal y provoca grietas. Si no se logra enfriar la pared, esta colapsa en cuestión de pocos segundos.

Si la pared recibe este calor directamente, se funde o se debilita. Por eso se utiliza la refrigeración regenerativa. Se crean múltiples conductos finos dentro de la pared. El combustible se hace fluir primero a través de estos conductos. Al pasar por ellos, el combustible absorbe el calor de la pared. La pared se enfría y el combustible se calienta. El combustible calentado entra en la cámara de combustión

y se quema. No se desperdicia calor, ya que se protege la pared y al mismo tiempo se precalienta el combustible. A esto también se le puede añadir la refrigeración por película. La refrigeración por película es un método que cubre la pared con una fina capa de fluido refrigerante a lo largo de su superficie. Si se utilizan ambos métodos juntos, la pared se puede proteger de forma más segura.

El problema radica en fabricar estos canales de refrigeración. Antiguamente, los conductos se creaban mecanizando un bloque de metal. El método de corte y mecanizado obligaba a que la forma de los canales fuera sencilla. Por lo general, se fresaban ranuras rectangulares rectas y paralelas. En tales canales resulta difícil variar sustancialmente la anchura o la dirección según la posición. Es complicado refrigerar con mayor intensidad solo la garganta de la tobera, donde se concentra gran parte del calor. Si se estrecha el canal a la fuerza, al combustible le cuesta pasar y la pérdida de carga aumenta. Cuando la pérdida de carga es grande, la bomba que impulsa el combustible debe ejercer mayor potencia.

La impresión 3D en metal está superando esta limitación. La impresión 3D en metal es un método en el que se extiende una fina capa de polvo metálico y se funde con un láser. La pieza se fabrica acumulando capas una a una. Este método permite crear conductos internos complejos que no pueden fabricarse mediante mecanizado. También es posible un diseño en el que los canales se ramifican como ramas de árboles y luego se vuelven a unir. Se pueden construir canales que se curvan siguiendo la superficie curva de la pared. Es posible disponer los canales de manera densa donde se concentra el calor y de forma más espaciada donde no. La libertad de diseño se ha ampliado enormemente.

Veamos un ejemplo de cómo la impresión 3D transforma los canales de refrigeración. El método anterior creaba los canales mecanizando ranuras en la pared interior y colocando una cubierta exterior sobre ellas. Había que fabricar varias piezas y unirlas. Si hay muchas uniones, esas zonas pueden volverse débiles. La impresión 3D produce la pared con los canales integrados como una sola pieza. Al reducirse las uniones, la pared se vuelve más resistente. Los pasos de fabricación también disminuyen, ahorrando tiempo y material. Por ello, se ha vuelto más fácil modificar el diseño y probarlo en repetidas ocasiones.

Los materiales para los canales de refrigeración también han avanzado al mismo tiempo. La Administración Nacional de Aeronáutica y el Espacio (NASA) desarrolló una aleación de cobre llamada GRCop-42. El cobre es un metal que conduce bien el calor. El GRCop-42 es una aleación de cobre con pequeñas adiciones de cromo y niobio. Transfiere bien el calor y resiste altas temperaturas. Esta aleación permite fabricar mediante impresión 3D cámaras de combustión completas con canales integrados. La NASA llevó a cabo pruebas de combustión reales con cámaras de combustión fabricadas con esta aleación. Los resultados de este desarrollo y las pruebas están recogidos en informes públicos. La NASA impulsó este material y la tecnología de fabricación dentro de sus iniciativas de tecnología de propulsión de próxima generación. Es un ejemplo en el que los materiales, la fabricación y el diseño avanzaron como un conjunto integrado.

Si también se utiliza la refrigeración por película, la pared se puede proteger de manera más segura. Se perforan varios orificios pequeños en la pared. A través de esos orificios se hace fluir poco a poco un fluido frío. El fluido que sale forma una fina película a lo largo de la pared. Esta película se interpone entre el gas caliente y la pared. Si la refrigeración regenerativa enfría la pared desde el interior, la refrigeración por película cubre la superficie de la pared con una capa protectora. Ambos métodos se complementan. En las inmediaciones de la garganta de la tobera, donde se concentra gran cantidad de calor, a menudo se superponen ambos métodos.

La forma de los canales de refrigeración influye en gran medida en el rendimiento de la gestión térmica. Si los canales son estrechos y densos, el área de contacto con la pared aumenta. Si el área de contacto es grande, se absorbe más calor. Sin embargo, si el canal es demasiado estrecho, al combustible le cuesta

pasar. Se requiere más energía para impulsar el combustible. A esta pérdida de energía se le llama pérdida de carga. Si la pérdida de carga es grande, la bomba debe girar con más fuerza. Un buen canal de refrigeración enfría bien la pared manteniendo baja la pérdida de carga. Estos dos factores guardan una relación de compromiso mutuo.

El nuevo método no está exento de dificultades. Un estudio publicado en 2025 analizó un caso de fallo de una cámara de combustión de GRCop-42 impresa en 3D. En una prueba de combustión realizada en febrero de 2021, la cámara de combustión falló tras solo 9 segundos. Al examinar la causa, el problema residía en un punto donde el proceso de impresión se había detenido momentáneamente. En ese lugar se formaron muchos más microporos de lo habitual. Cuando hay muchos poros, esa zona se debilita. Este caso demuestra cuán sensible es el nuevo método de fabricación al control de calidad. A cambio de obtener libertad de diseño, se debe supervisar el proceso de fabricación con mayor minuciosidad.

La mayor libertad de diseño también plantea nuevos dilemas. Al poder variar la forma de los canales de manera prácticamente infinita, se volvió más difícil elegir cuál es la mejor forma. El método en el que una persona dibuja a mano algunas formas y realiza cálculos no permite explorar todo el amplio espacio de diseño. Solo para calcular con precisión el flujo y el calor de un solo canal, una computadora requiere mucho tiempo. La inteligencia artificial entra en juego para explorar este vasto espacio de manera rápida e inteligente. A partir de la siguiente sección, examinaremos estos métodos uno a uno.

9.2 Modelos sustitutos para predecir variaciones bruscas del combustible en canales de refrigeración

Lo que aborda esta sección es por qué resulta difícil predecir el comportamiento del combustible en los canales de refrigeración. El metano o el queroseno utilizados como refrigerantes cambian drásticamente de propiedades a altas presiones y temperaturas. El análisis de alta precisión que calcula este flujo tan variable es lento. Por ello, se utilizan modelos sustitutos de inteligencia artificial para reemplazar los cálculos. Aquí se explican los principios y las limitaciones de dichos modelos sustitutos.

Como refrigerantes de cohetes se utilizan metano (methane) o RP-1. El RP-1 es queroseno refinado para uso en cohetes. Es una especie de primo del queroseno que usamos comúnmente. Estos combustibles experimentan presiones muy altas dentro de los canales de refrigeración. Cuando la presión supera cierto valor, la frontera entre líquido y gas se desdibuja. Tomando el agua como ejemplo, es un estado en el que desaparece la distinción en la que el agua hierve en una olla formando burbujas. A este estado se le denomina estado supercrítico (supercritical). El combustible en estado supercrítico presenta propiedades intermedias, no siendo ni puramente líquido ni puramente gas.

El término supercrítico puede sonar difícil. Expliquémoslo de manera sencilla. Cuando se hierve agua en una olla, el agua se convierte en burbujas y entra en ebullición. El líquido y el gas están claramente divididos. Sin embargo, si la presión se eleva mucho, esta ebullición desaparece. El líquido pasa gradualmente a presentar propiedades similares a las de un gas. La frontera definida donde se forman burbujas se desvanece. El estado en el que esta frontera desaparece es el estado supercrítico. Dado que el combustible en los canales de refrigeración está sometido a una alta presión, fluye en este estado.

Las propiedades del estado supercrítico cambian drásticamente con la temperatura. Al pasar por cierto intervalo de temperatura, la densidad cae en picado. En el mismo intervalo, el calor específico se dispara repentinamente. El calor específico es la cantidad de calor necesaria para elevar la temperatura en 1 grado. El punto donde la densidad y el calor específico fluctúan de este modo se denomina línea pseudocrítica (pseudo-critical line). Cerca de este punto, las propiedades del combustible cambian drásticamente como si estuviera hirviendo. Aunque en realidad no hierve, se asemeja a la ebullición, por lo que se le llama pseudoebullición (pseudo-boiling).

El aumento repentino del calor específico también tiene un aspecto positivo para la refrigeración. Si el calor específico es grande, el combustible puede almacenar mucho calor. La misma cantidad de combustible transporta más calor. Sin embargo, si la densidad cae drásticamente, el combustible se expande y se acelera. Si este flujo acelerado fluctúa cerca de la pared, la forma en que se transfiere el calor se desestabiliza. Los aspectos positivos y negativos aparecen juntos en el mismo punto. Por eso resulta complicado predecir la refrigeración en este intervalo.

La pseudoebullición causa problemas en la refrigeración. El combustible debe absorber bien el calor de la pared para que esta se enfríe. Sin embargo, cerca de la línea pseudocrítica, la capacidad de absorber calor puede caer repentinamente. Una capa de combustible calentada cerca de la pared bloquea el calor como si fuera una manta delgada. Este fenómeno se denomina deterioro de la transferencia de calor (heat transfer deterioration). Cuando se produce el deterioro de la transferencia de calor, la temperatura de la pared en ese punto se dispara de repente. Un estudio de 2020 resumió que la pseudoebullición y el deterioro de la transferencia de calor ocurren simultáneamente al calentar combustible para cohetes. Un estudio de 2024 investigó por qué ocurre este deterioro mediante un modelo microscópico de flujo de dos estados. Este deterioro de la transferencia de calor también se relaciona con el problema de la coquización, que se abordará más adelante. Esto se debe a que, si la pared se calienta repentinamente, el combustible se descompone más rápido en ese lugar.

Existe un método preciso para calcular estos flujos. Se trata de la dinámica de fluidos computacional (computational fluid dynamics, CFD). Divide el flujo en pequeñas celdas de una malla y resuelve la velocidad, la presión y la temperatura de cada celda de acuerdo con las leyes de la física. Las leyes fundamentales son tres leyes de conservación. La primera es la conservación de la masa. La cantidad de fluido que entra y la cantidad de fluido que sale deben coincidir. La siguiente es la conservación del momento lineal. Las fuerzas que empujan o tiran del fluido deben coincidir con el cambio de velocidad. La última es la conservación de la energía. El calor entrante y el saliente deben coincidir con el cambio de temperatura. Resolver estas tres leyes conjuntamente en todas las celdas es lo que hace la CFD. La CFD es precisa pero lenta. Incluso para resolver con precisión una sola sección de un canal de refrigeración, una computadora puede tardar casi una hora. En el diseño de canales de refrigeración, es necesario verificar miles de variaciones en la forma del canal y el caudal. Si cada cálculo toma una hora, el diseño nunca terminaría.

Por eso se utilizan modelos subrogados (surrogate model). Un modelo subrogado es un modelo de aproximación rápida que reemplaza cálculos lentos. Primero, se calculan previamente múltiples casos mediante CFD para recopilar datos. Con esos datos se entrena una red neuronal artificial. La red neuronal recibe como entrada la forma del canal, el caudal, la presión y la temperatura de entrada. Luego, arroja directamente la temperatura de la pared, el flujo de calor y la pérdida de carga. Investigadores del Centro Aeroespacial Alemán (DLR) crearon un modelo subrogado basado en redes neuronales de este tipo. Lo que utilizaron en la práctica fue una red neuronal estándar con varias capas ocultas. Este modelo completó en menos de un segundo cálculos que con CFD tomaban una hora. Al principio solo abordaron canales rectos, pero luego ampliaron el alcance para incluir canales curvos.

El proceso para crear un modelo subrogado es el siguiente. Primero, se define el rango de las variables de diseño. Este comprende el ancho del canal, la altura del canal, el caudal y la temperatura de entrada. Dentro de ese rango, se seleccionan varios casos y se calculan mediante CFD. Los resultados de los cálculos se agrupan en pares de entrada y salida. Estos pares se convierten en el libro de texto de la red neuronal. La red neuronal observa los pares y aprende las reglas para pasar de la entrada a la salida. Una vez finalizado el entrenamiento, responde con rapidez ante nuevas entradas. Se parece a la forma en que una persona aprende tipos de problemas resolviendo múltiples ejemplos.

Existen más métodos para aumentar la precisión de los modelos subrogados. Estas técnicas no fueron utilizadas directamente por el estudio del DLR mencionado antes, sino que son recursos ampliamente empleados para hacer más robustos los modelos subrogados. Uno de ellos consiste en incorporar conocimiento físico. Si una red neuronal aprende únicamente a partir de datos, puede generar respuestas que violen las leyes de la física. Por lo tanto, se insertan en las reglas de aprendizaje las leyes de conservación de la masa, el momento lineal y la energía. Si la red neuronal infringe esas leyes, se le impone una penalización. Una red neuronal construida de este modo se denomina red neuronal informada por la física (physics-informed neural network, PINN). A esto también se le puede añadir una estructura llamada conexión residual (residual connection). Una conexión residual es una vía que omite información de las capas anteriores y la envía directamente a las capas posteriores. Esta estructura ayuda a entrenar redes neuronales profundas de forma estable. Incorporar tales técnicas a los modelos subrogados de canales de refrigeración es la dirección que propone este libro.

También es necesario introducir con precisión los valores de las propiedades de los fluidos supercríticos. El Instituto Nacional de Estándares y Tecnología de los Estados Unidos (NIST) recopila y publica valores como la densidad y el calor específico de diversos fluidos según la temperatura y la presión. Para fluidos puros de un solo componente, como el metano, estos valores estándar se pueden conectar y utilizar directamente. De esta manera, no se pasan por alto los cambios bruscos cerca de la línea pseudocrítica. Sin embargo, en el caso del queroseno RP-1, la situación es diferente. Dado que el RP-1 es una mezcla de varios hidrocarburos, su composición varía ligeramente con cada refinado. Por ello, no se pueden usar directamente tablas de fluidos puros y es necesario crear un modelo de mezcla sustituta independiente que ajuste sus propiedades. Si los valores de las propiedades son inexactos, el modelo subrogado también arrojará temperaturas de pared erróneas. Dado que los fluidos supercríticos tienen propiedades muy sensibles, estos datos de base son fundamentales.

También existen investigaciones sobre cómo elegir adecuadamente el refrigerante en sí. Un estudio de 2026 diseñó un combustible sustituto que imita el queroseno para cohetes. Se utilizó un algoritmo genético para encontrar una combinación que se ajustara bien a las propiedades en estado supercrítico. Un algoritmo genético es un método que mejora las soluciones seleccionando y mezclando combinaciones prometedoras. Cuanto más preciso sea el modelo de combustible utilizado en los cálculos de refrigeración, más precisa será la predicción. La inteligencia artificial también contribuye a la tarea de seleccionar combinaciones de combustibles.

Hay que tener cuidado al trabajar con modelos subrogados. No se debe depender de nombres de modelos usados como marcas ni de cifras de rendimiento no verificadas. Por esta razón, este libro explica mediante el concepto general de modelo subrogado de CFD en lugar de usar nombres de modelos específicos. Si los principios son los mismos, lo que se aprende es idéntico sin importar el nombre.

Este modelo subrogado también se ha aplicado a motores reales. El Centro Aeroespacial Alemán examinó el rendimiento de los canales de refrigeración de un motor de desarrollo propio utilizando este método. Mejoraron la predicción de la temperatura de la pared y la incorporaron al diseño. Al acelerar los cálculos, pudieron comparar múltiples diseños de canales en poco tiempo. De este modo, exploraron ampliamente el espacio de diseño sin quedar atados a cálculos precisos y lentos.

Los modelos subrogados tienen limitaciones claras. Un modelo subrogado solo es confiable dentro del rango en el que fue entrenado. Puede cometer grandes errores ante formas de canal desconocidas o condiciones de operación no vistas durante el entrenamiento. Dado que las propiedades de los fluidos supercríticos son muy sensibles, este riesgo es aún mayor. Por ello, no se debe definir el diseño final basándose únicamente en el modelo subrogado. Con el modelo subrogado se seleccionan rápidamente

los candidatos prometedores. Luego, se eligen solo esos candidatos para verificarlos de nuevo mediante CFD de alta precisión y pruebas de combustión reales. El modelo subrogado no es una herramienta para dar la respuesta definitiva, sino para filtrar candidatos rápidamente.

9.3 Aprendizaje activo para encontrar orificios de refrigeración por película con pocos cálculos

Veamos primero por qué es necesaria la refrigeración por película. La refrigeración regenerativa enfría la pared desde el interior. Sin embargo, en zonas donde el calor se concentra en exceso, la refrigeración desde el interior de la pared puede no ser suficiente. Esto se debe a que la superficie misma de la pared entra en contacto directo con los gases calientes y se calienta. Por lo tanto, se añade una capa delgada de película protectora sobre la superficie. Esto es la refrigeración por película. La refrigeración regenerativa y la refrigeración por película protegen la pared conjuntamente desde dentro y desde fuera.

La refrigeración por película (film cooling) es un método para crear una fina capa de fluido frío justo delante de una pared caliente. Se perforan múltiples orificios pequeños en la pared. A través de esos orificios se expulsa fluido frío de forma oblicua. El fluido expulsado se extiende finamente a lo largo de la pared y la cubre como una manta. Esta película separa los gases calientes de la pared. La pared no entra en contacto directo con los gases calientes, por lo que se calienta menos. Se asemeja al aire que fluye de abajo hacia arriba en el parabrisas de un automóvil para desempañarlo.

El rendimiento de la refrigeración por película está determinado por varios parámetros. Uno de ellos es la intensidad del fluido inyectado. Esto se conoce como relación de soplado (blowing ratio). Es el valor resultante de dividir la intensidad del flujo del fluido frío entre la del gas caliente. El ángulo de inclinación del orificio, la separación entre orificios y la forma del orificio también determinan el rendimiento. Incluso si estos valores varían ligeramente, el rendimiento de la película cambia drásticamente. Que la película se adhiera bien a la pared o se eleve y se desprenda depende de estos valores.

Si la relación de soplado es demasiado grande, también surgen problemas. Si el fluido se inyecta con demasiada fuerza, la película se despega de la pared. El chorro inyectado no logra cubrir la pared y se eleva hacia arriba. Entonces, la pared vuelve a entrar en contacto con el gas caliente. Si la relación de soplado es demasiado pequeña, la película es tan delgada que se disipa rápidamente. Por lo tanto, existe un valor óptimo para la intensidad. Solo al coordinar adecuadamente el ángulo, la separación y la intensidad se logra una buena película.

Una de las razones por las que la película se desprende son los vórtices. Cuando el fluido sale expulsado por el orificio, se genera un par de vórtices que giran a izquierda y derecha. Estos vórtices arrastran el gas caliente hacia la pared. Como resultado, la película de refrigeración creada con tanto esfuerzo se levanta y se desprende. La pared queda expuesta nuevamente al gas caliente. Si se diseña adecuadamente el orificio, estos vórtices se pueden debilitar. Por eso, hoy en día se utilizan orificios en forma de abanico, cuyos extremos se abren hacia los lados. El fluido inyectado se extiende ampliamente hacia los lados y cubre la pared de manera uniforme.

Los orificios en forma de abanico extienden la película de refrigeración más ampliamente que los orificios circulares. Los orificios circulares inyectan el fluido en un solo chorro. Este chorro tiende a levantarse fácilmente de la pared. En cambio, los orificios en forma de abanico ensanchan su extremo hacia los lados para expulsar el fluido ampliamente. El fluido ampliamente dispersado se adhiere mejor a la pared. Al aumentar la superficie de adhesión, cubre la pared de manera uniforme. Por ello, las

turbinas de gas y los motores aeronáuticos utilizan frecuentemente orificios en forma de abanico. Cuánto se ensanche el orificio y con qué ángulo se perfore determinarán el rendimiento.

Para encontrar una buena forma de orificio, es necesario calcular el flujo. Si se quiere observar detalladamente incluso los vórtices, se requiere un análisis de alta precisión. Este análisis se denomina simulación de grandes remolinos (large eddy simulation, LES). La LES es un método que calcula directamente los grandes torbellinos dentro del flujo. Es precisa, pero el costo computacional es muy alto. Resolver una sola condición lleva mucho tiempo. Resulta difícil calcular cientos de combinaciones variando el ángulo, la separación y la intensidad del orificio. Dado que el presupuesto de cálculo es limitado, no se pueden calcular condiciones al azar de forma indiscriminada.

Aquí es donde los procesos gaussianos (Gaussian process) resultan útiles. Un proceso gaussiano es un método para realizar predicciones con pocos datos. Su característica distintiva es que no solo proporciona valores predichos, sino que también indica qué tan incierta es la predicción. En regiones con abundantes datos, predice con gran confianza. En regiones sin datos, señala que existe incertidumbre. Es semejante a dibujar un mapa donde los caminos ya transitados se trazan con detalle y los no transitados se dibujan de forma borrosa. Este método conecta los valores mediante una regla suave denominada función de Matérn (Matern).

Veamos también cómo el proceso gaussiano conecta los valores. Es de esperar que condiciones cercanas tengan rendimientos similares. Si se modifica muy poco el ángulo del orificio, el rendimiento de refrigeración también cambiará solo un poco. La representación matemática de esta idea es el kernel. El kernel es una medida que evalúa cuán cercanas son dos condiciones. Si están cerca, asume que sus valores son similares; si están lejos, considera que sus valores pueden ser diferentes. Con esta métrica, conecta suavemente las condiciones calculadas previamente y estima los valores de las condiciones que aún no han sido calculadas.

El aprendizaje activo (active learning) añade a esto una selección inteligente. Elige por sí mismo qué condición calcular a continuación mediante LES. Al elegir, evalúa dos aspectos simultáneamente. Uno son las zonas que parecen ofrecer un buen rendimiento. El otro son las zonas de gran incertidumbre. Explorar zonas con buen rendimiento prometedor permite acercarse a una mejor solución. Explorar zonas inciertas aclara las partes borrosas del mapa. La regla que equilibra ambos aspectos se denomina función de adquisición (acquisition function). Una regla utilizada con frecuencia consiste en medir cuánto se espera mejorar respecto al estado actual.

La ventaja de este método radica en determinar inteligentemente el orden de los cálculos. No calcula condiciones arbitrarias al azar. Selecciona y calcula primero las condiciones que ofrecen mayor aprendizaje. Así, con un número reducido de cálculos de LES, se aproxima a una forma de orificio óptima. Reduce cálculos costosos y acota el diseño con rapidez. Este enfoque se estableció primero en el diseño de orificios de refrigeración para álabes de turbinas de gas. Los estudios validados hasta ahora se centran principalmente en turbinas de gas. Mediante optimización bayesiana multiobjetivo, se mejoraron simultáneamente el rendimiento de refrigeración y la uniformidad. Un estudio reportó haber aumentado el rendimiento de refrigeración respecto a diseños anteriores bajo la misma condición de pérdida de carga.

La inteligencia artificial también se utiliza en el propio cálculo del flujo. Hacia 2024, diversos estudios desarrollaron redes neuronales para predecir la eficiencia de la refrigeración por película. Ciertos estudios predijeron la eficiencia de refrigeración incluso al variar la curvatura de la pared y el ángulo del orificio. Estos modelos de predicción proporcionan valores similares a los de cálculos de alta precisión con mucha mayor rapidez. Resultan útiles para comparar rápidamente múltiples formas de

orificios en las fases iniciales del diseño. También es posible combinar ambos enfoques, de modo que el proceso gaussiano elija el siguiente cálculo y la red neuronal lo estime rápidamente.

La refrigeración por película en las cámaras de combustión de cohetes también aprovecha estos mismos métodos. Sin embargo, en los cohetes las condiciones son mucho más severas. El gas es más caliente y más rápido. Por ello, los valores óptimos de orificios obtenidos en motores aeronáuticos no deben transferirse directamente a los cohetes. Si cambian el tamaño del orificio, la velocidad del fluido, la curvatura de la pared o el tipo de combustible, los resultados varían. Aunque se puede trasladar una buena metodología de diseño, los valores numéricos óptimos específicos deben buscarse de nuevo en cada caso. La inteligencia artificial es una herramienta que agiliza esa nueva búsqueda. La geometría final debe validarse nuevamente bajo las condiciones propias del cohete.

9.4 Diseño de la forma de los canales de refrigeración mediante inteligencia artificial

El diseño de los canales de refrigeración es un problema en el que se entrelazan sólidos y fluidos. El calor del gas caliente entra primero en la pared metálica. Dicho calor se transmite a través del metal hacia los canales de refrigeración. En el canal, el combustible que fluye absorbe ese calor. La conducción térmica a través del metal y la convección absorbida por el fluido se encuentran en un mismo lugar. El cálculo que resuelve ambos fenómenos juntos se denomina transferencia de calor conjugada (conjugate heat transfer, CHT). En los canales de refrigeración de cohetes, los sólidos y los fluidos están fuertemente acoplados, por lo que este cálculo resulta crucial.

Es fácil de entender si pensamos en una ventana en invierno. El interior de la habitación es cálido y el exterior es frío. La temperatura de la cara interna del vidrio difiere de la de la cara externa. El calor fluye a través del vidrio y, en el exterior, el aire frío absorbe el calor. La conducción térmica dentro del vidrio y la convección del aire exterior se encuentran en la superficie de la ventana. La pared de un cohete es idéntica. El calor entra desde el lado del gas caliente y el combustible absorbe el calor desde el lado de los canales de refrigeración. Ambos flujos se encuentran en la pared y alcanzan el equilibrio.

La clave de la transferencia de calor conjugada reside en el compromiso en la frontera. Pensemos en la superficie de contacto entre el metal y el combustible. El calor que sale del metal hacia esa superficie debe ser igual al calor que el combustible recibe en ella. Esto se debe a que el calor no se crea ni se destruye. La temperatura de la pared debe calcularse respetando este principio. Este cálculo nos indica en qué punto de la pared la temperatura alcanza su valor máximo y cómo refrigerar esa zona.

El cálculo de la transferencia de calor conjugada requiere una gran carga computacional. Esto se debe a que la conducción térmica en el sólido y la convección en el fluido deben resolverse simultáneamente. Además, el fluido es turbulento. La turbulencia es un flujo en el que se mezclan vórtices grandes y pequeños. Cuanto más detalladamente se resuelva este flujo, mayor será el tiempo de cálculo. Incluso para resolver con precisión un solo canal de refrigeración, la computadora tarda mucho tiempo. Si existen miles de candidatos de diseño, ese cálculo debe repetirse miles de veces. Por ello, se necesita una técnica que reduzca el cálculo.

Ahora es el momento de diseñar la forma del canal. Para ello se utiliza la optimización topológica (topology optimization). La optimización topológica es un método que determina si un espacio definido debe rellenarse con material o dejarse como espacio vacío. En los canales de refrigeración, se convierte en el problema de decidir qué partes se dejan como metal y cuáles como canales vacíos por donde fluye el combustible. El ser humano no define la forma del canal de antemano. Con solo fijar el objetivo, la optimización genera por sí misma la forma del canal. El objetivo es reducir la temperatura máxima de la pared, evitando al mismo tiempo que la pérdida de carga aumente excesivamente. Ya existen

investigaciones que diseñan canales de refrigeración mediante optimización topológica. Un ejemplo representativo es el método basado en densidad, que ajusta gradualmente la distribución del material considerando conjuntamente el flujo y el calor.

Este problema maneja demasiadas variables. Hay cientos de miles de celdas que dividen finamente el espacio. En cada celda es necesario decidir si es metal o espacio vacío. El número de combinaciones es incontable. Explorar este vasto espacio tal cual tomaría un tiempo infinito. Por lo tanto, se requiere una técnica para simplificar el problema. Aquí es donde entra en juego el autoencoder convolucional (convolutional autoencoder, CAE).

Un autoencoder es una red neuronal que comprime datos complejos a un tamaño reducido y luego los reconstruye. La parte frontal comprime la forma del canal en un pequeño conjunto de números. Este conjunto de números se denomina vector latente (latent vector). La parte posterior reconstruye la forma del canal a partir de dicho vector latente. Con un entrenamiento adecuado, es posible capturar la forma esencial del canal con solo unos pocos números. Es similar a comprimir una fotografía en un archivo pequeño y volver a abrirla. La convolución es un método de cálculo que examina conjuntamente las regiones adyacentes de una imagen. Es idónea para procesar datos distribuidos espacialmente, como la forma de los canales.

Los casos validados en los que este método de reducción del espacio de diseño mediante autoencoders se aplica directamente a los canales de refrigeración de cohetes aún son escasos. Este libro propone esa dirección. La optimización topológica en sí es un método probado, y la idea es incorporar autoencoders sobre ella para acelerar la búsqueda.

Esta compresión facilita el diseño. El optimizador no manipula directamente cientos de miles de celdas. En su lugar, solo ajusta el pequeño vector latente. Al modificar ligeramente el vector latente, la parte posterior genera una forma de canal plausible. En la práctica, se reduce el amplio espacio de diseño a un espacio pequeño para explorarlo. El optimizador busca un buen vector latente en este espacio reducido. El canal generado por dicho vector se convierte en un excelente canal de refrigeración.

El uso de vectores latentes ofrece otra ventaja adicional. Las formas de canal generadas son, en general, plausibles. Esto se debe a que el autoencoder ha aprendido a partir de datos de canales reales. Genera nuevas formas dentro del marco de las formas aprendidas durante el entrenamiento. Por ello, se producen menos formas extrañas que carezcan de sentido físico. Se reducen los esfuerzos inútiles que surgen al buscar aleatoriamente en un espacio amplio. Dado que las soluciones se buscan únicamente en un espacio reducido y coherente, la búsqueda es rápida. No obstante, el mero hecho de asemejarse a los datos de entrenamiento no garantiza un buen canal. Es necesario verificar por separado si el canal tiene continuidad de principio a fin, si las paredes no son demasiado delgadas y si la resistencia al flujo y la resistencia estructural son adecuadas. Estas condiciones deben establecerse por separado como restricciones dentro de la optimización.

También debe incluirse la condición de que la forma sea fabricable. La impresión 3D fabrica acumulando capas hacia arriba. Las secciones que se inclinan demasiado hacia los lados sin soporte inferior tienden a pandearse durante la fabricación. Por esta razón, se impone una restricción para que las paredes del canal no se inclinen más allá de cierto ángulo. Si se respeta esta restricción, se obtiene un canal autosoportado sin necesidad de soportes. Al incorporar estas condiciones dentro de la optimización, el diseño resultante puede imprimirse de inmediato. Esto evita que un buen diseño en el cálculo resulte imposible de fabricar en la realidad.

También existen investigaciones que sustituyen el cálculo de transferencia de calor conjugada por inteligencia artificial. Un estudio de 2024 construyó una red neuronal apilando capas de codificadores y decodificadores. Esta red neuronal recibía como entrada la forma del componente y trazaba

directamente la distribución de temperatura. Sustituyó la ejecución repetida de cálculos precisos de transferencia de calor conjugada por predicciones rápidas. Otro estudio de 2025 abordó el sólido y el fluido dividiéndolos en redes neuronales distintas. Entrenó ambas redes neuronales para que hicieran coincidir sus valores en la frontera. De este modo, se capturan mejor las variaciones bruscas de temperatura en la superficie de contacto entre la pared y el combustible. Dichos estudios trataron problemas generales de transferencia de calor conjugada. Si se utilizan estos modelos de predicción rápida dentro de la optimización topológica de los canales de refrigeración de cohetes, la búsqueda del diseño se acelera. Dicha integración es una dirección que se irá perfeccionando en el futuro.

Los canales generados por la optimización topológica presentan formas difíciles de dibujar a mano por una persona. Las ramas se bifurcan y convergen, y se curvan siguiendo la superficie de la pared. A veces se disponen densamente aletas inclinadas similares a espinas de pescado. Estas formas distribuyen el calor de manera uniforme y guían el flujo suavemente. Aunque a los ojos humanos no parezcan atractivas, pueden ser formas excelentes desde el punto de vista del cálculo. La inteligencia artificial encuentra estas formas inesperadas.

Existen casos industriales de este enfoque. Una empresa aplicó esta tecnología de optimización topológica a los canales de refrigeración de cohetes y anunció que logró reducir notablemente la temperatura máxima de la pared. Recientemente, también han surgido investigaciones que emplean superficies regladas en los canales de refrigeración. Esta superficie se denomina superficie mínima triplemente periódica (triply periodic minimal surface, TPMS). La TPMS es una estructura en la que superficies suavemente curvadas, semejantes a películas de pompas de jabón, se repiten periódicamente en el espacio. Esta estructura pliega una gran superficie dentro de un espacio reducido. Una mayor superficie permite un mayor intercambio térmico. Un estudio de 2025 fabricó canales TPMS que envuelven la superficie curva de la pared y reportó que tanto el rendimiento de refrigeración como la uniformidad de la temperatura alrededor de la pared mejoraron conjuntamente. En 2026, también apareció un estudio que mejoró la transferencia de calor mediante canales de refrigeración con relleno reticular. Estos canales de superficie curva y canales reticulares solo pueden fabricarse con impresión 3D. No obstante, los casos empresariales difieren en su método de validación respecto a los artículos académicos. En este libro se abordan únicamente como casos de referencia para mostrar las tendencias industriales. El rendimiento final solo debe tomarse como base de diseño cuando se disponga de datos de ensayos independientes.

9.5 Optimización multiobjetivo para reducir la coquización por acumulación de hollín

El RP-1, un combustible derivado del queroseno, presenta una desventaja cuando se utiliza como refrigerante. Al pasar cerca de paredes calientes, el combustible se descompone por efecto del calor. Esto se denomina pirólisis. Entre los fragmentos descompuestos, el carbono se adhiere a la pared. Es similar a cuando el aceite se deja durante mucho tiempo en una sartén y se quema formando una costra negra. Esta costra de carbono que se adhiere a la pared se denomina coquización (coking). La coquización se acumula gradualmente en el interior de los canales de refrigeración.

El fenómeno por el cual el combustible se descompone térmicamente se denomina pirólisis. Es un proceso en el que moléculas grandes reciben calor y se dividen en fragmentos más pequeños. A mayor temperatura, más rápida es la descomposición. Entre los fragmentos descompuestos, aquellos con alto contenido de carbono se aglomeran entre sí. El carbono aglomerado se adhiere a la pared y forma una costra. Por ello, cuanto más caliente está la pared, con mayor facilidad se produce la coquización. La coquización depende conjuntamente de la temperatura de la pared, el tipo de combustible y el tiempo de paso.

La coquización se produce con facilidad en la garganta de la tobera. La garganta de la tobera es el lugar donde el flujo térmico alcanza su valor máximo. Al estar la pared tan caliente, la descomposición del combustible también es intensa. La coquización causa dos problemas. Uno es que la costra de carbono no conduce bien el calor. La costra bloquea el contacto entre la pared y el combustible, haciendo que la pared se caliente aún más. El otro problema es que la costra estrecha el canal. Al estrecharse el canal, resulta más difícil el paso del combustible y aumenta la pérdida de carga. Cuando ambos problemas coinciden, la pared entra en una situación de peligro. En los motores reutilizables, la coquización se convierte en un problema mayor debido a que se acumula con cada prueba sucesiva.

Existe una regla en la relación entre la temperatura y la coquización. Incluso un pequeño aumento de temperatura acelera enormemente las reacciones químicas. Si la temperatura de la pared se eleva, la descomposición del combustible se acelera de forma notable. En consecuencia, la coquización también se acumula más rápido. Por ello, con solo reducir ligeramente la temperatura de la pared se puede disminuir considerablemente la coquización. Una buena refrigeración enfría la pared, y una pared más fría reduce la coquización. La refrigeración y la supresión de la coquización están vinculadas de este modo.

El deterioro de la transferencia de calor de la sección anterior se entrelaza con esta coquización. Cuando el combustible supercrítico cruza la línea pseudocrítica, la transferencia de calor cerca de la pared se degrada repentinamente. En consecuencia, la temperatura de la pared en ese punto se dispara. Al calentarse la pared, el combustible se descompone con mayor rapidez y la coquización se acumula velozmente. La coquización acumulada bloquea aún más el calor, calentando todavía más la pared. Si esta retroalimentación continúa, la pared puede colapsar de repente. El deterioro de la transferencia de calor y la coquización se potencian mutuamente.

Para reducir la coquización, es necesario disminuir la temperatura de la pared. Si la pared está menos caliente, el combustible se descompone menos. Además, debe observarse conjuntamente el flujo cerca de la pared. Si el combustible roza rápidamente la pared al pasar, se reducen las oportunidades de que la costra se adhiera. La fuerza con la que el fluido roza la pared en sus inmediaciones se denomina esfuerzo cortante de pared. Esta fuerza debe ser suficientemente intensa para que se deposite menos costra. El diseño para la supresión de la coquización es un problema de control conjunto de la temperatura de la pared y el esfuerzo cortante de pared.

El problema radica en que existen múltiples objetivos que entran en conflicto entre sí. Está el objetivo de reducir la temperatura de la pared. Está el objetivo de disminuir la pérdida de carga. Está el objetivo de suprimir la coquización. Estos objetivos no mejoran simultáneamente. Si se estrecha el canal para acelerar el flujo, la pared se enfría pero la pérdida de carga aumenta. Si se ensancha el canal para ralentizar el flujo, la pérdida de carga disminuye pero la refrigeración se debilita. Ganar por un lado implica perder por el otro. Un problema en el que entran en conflicto múltiples objetivos de esta manera se denomina optimización multiobjetivo.

Este tipo de compromiso nos resulta familiar en la vida cotidiana. Pensemos en cuando elegimos un bolso o una mochila. Es difícil conseguir uno que sea ligero, resistente y barato al mismo tiempo. Si es ligero tiende a ser frágil, y si es resistente suele ser pesado o caro. Por ello, elegimos un punto de equilibrio adecuado para su uso. Lo mismo ocurre con el diseño de canales de refrigeración. No es posible maximizar simultáneamente la refrigeración, la reducción de la pérdida de carga y la supresión de la coquización. Ganar en un aspecto implica ceder un poco en otro.

La optimización multiobjetivo no tiene una respuesta única. En su lugar, se obtiene un conjunto de múltiples soluciones óptimas que intercambian ventajas y desventajas entre sí. Este conjunto se denomina frontera de Pareto (Pareto front). Cada una de las soluciones sobre la frontera de Pareto

presenta un punto de equilibrio diferente. Algunas soluciones priorizan la refrigeración, mientras que otras priorizan la reducción de la pérdida de carga. El diseñador examina este conjunto y elige el punto de equilibrio adecuado para el motor. En lugar de que una persona fije una única solución de antemano, es un método en el que se despliegan buenas alternativas candidatas para luego elegir.

Aquí es donde se utiliza la optimización bayesiana (Bayesian optimization). Es un método que amplía los procesos gaussianos y el aprendizaje activo de la sección anterior a múltiples objetivos. En primer lugar, traza un mapa aproximado de los objetivos con una cantidad reducida de cálculos. A partir de ese mapa, selecciona qué condiciones calcular a continuación. Se calculan primero las condiciones que contribuyan significativamente a expandir la frontera de Pareto. Esto permite encontrar excelentes puntos de equilibrio realizando pocos análisis de combustión costosos. Si se incluye la coquización como objetivo, se obtiene un diseño que sopesa conjuntamente la temperatura de la pared, el esfuerzo cortante y la pérdida de carga. Sin embargo, todavía es difícil encontrar casos validados de optimización que incluyan la coquización como objetivo en canales de refrigeración de cohetes. Dado que la física de la pirólisis y la deposición es bien conocida, este libro propone la dirección de incorporar dicha física como objetivo de diseño.

La coquización se convierte en un desafío aún mayor en los motores reutilizables. Sin embargo, esto no significa que los motores desechables estén libres de coquización. Incluso durante una sola combustión, si los depósitos estrechan los canales o deterioran la transferencia de calor, la pared puede dañarse. Independientemente de si el motor es reutilizable o no, lo importante es qué tan rápido se acumulan los depósitos dentro del tiempo de la misión. En los motores reutilizables, a esto se suma el problema de la acumulación. A medida que se repiten las pruebas de combustión, los depósitos se acumulan poco a poco. Los depósitos acumulados empeoran gradualmente la refrigeración. Por eso, los motores reutilizables deben diseñarse para suprimir la coquización desde el principio. La optimización mediante inteligencia artificial ayuda a concebir diseños que anticipan incluso estas variaciones a lo largo del tiempo.

Este método también se combina a menudo con datos de pruebas de cohetes. El centro de pruebas de Lampoldshausen del Centro Aeroespacial Alemán es un lugar donde se realizan pruebas de combustión reales de motores cohete para obtener datos. Al disponer de estos datos de ensayo, el modelo base de optimización se vuelve más preciso. Los datos suplen el comportamiento real de coquización que los cálculos por sí solos pasan por alto. Es un flujo de trabajo que refina el diseño alternando entre pruebas y cálculos.

Los canales con una relación de aspecto alta también se emplean para suprimir la coquización. La relación de aspecto es el valor que se obtiene al dividir la altura del canal por su anchura. Los canales estrechos y profundos tienen una relación de aspecto alta. Estos canales tienen una mayor superficie de contacto con la pared, por lo que absorben bien el calor. Al enfriarse más la pared, se reduce la descomposición del combustible. En ocasiones, también se utilizan canales helicoidales o en espiral. Si se hace girar el flujo en espiral, este roza con mayor fuerza cerca de la pared. Sin embargo, estos canales pueden aumentar la pérdida de presión, por lo que ambos factores deben sopesarse conjuntamente.

En esta sección también hay aspectos que requieren precaución. Las cifras de reducción de coquización solo son fiables cuando están respaldadas por pruebas reales. Por ello, este libro se centra en los principios del método en lugar de en cifras específicas. En un motor real, la composición del combustible, la rugosidad de la superficie, la acción catalítica de las paredes y el tiempo de funcionamiento influyen en la coquización. Hay tantas variables que la predicción resulta difícil. Los valores previstos de coquización no respaldados por pruebas deben tratarse con cautela. La inteligencia

artificial se utiliza para reducir los candidatos prometedores, y la confirmación definitiva se realiza mediante pruebas de combustión reales.

9.6 Aprendizaje automático mejorado por la física para diseñar inductores de bombas

Los motores cohete deben impulsar con fuerza el combustible y el oxígeno hacia la cámara de combustión. Quien realiza esta tarea es la turbobomba (turbopump). La turbobomba es una bomba que gira a una velocidad extremadamente alta. Sin embargo, los álabes principales de la bomba no pueden funcionar adecuadamente si la presión de entrada es baja. Por ello, se coloca un álabe auxiliar delante del álabe principal. Este álabe auxiliar es el inductor. El inductor aspira suavemente el líquido primero y eleva ligeramente la presión. Gracias a esto, el álabe principal situado detrás puede girar de forma estable. Es un componente que actúa como el guardián de la bomba.

Veamos primero por qué es necesaria la turbobomba. La presión dentro de la cámara de combustión es sumamente alta. Para empujar el combustible hacia su interior, se requiere una presión aún mayor. El combustible a baja presión no puede entrar en una cámara de combustión de alta presión. Es lo mismo que el agua, que no puede fluir espontáneamente de un lugar bajo a uno alto. Por tanto, se utiliza una bomba para elevar sustancialmente la presión del combustible. La turbobomba del cohete logra esto mediante una rotación muy rápida. En motores grandes, la turbobomba gira decenas de miles de veces por minuto.

El inductor es aún más exigente cuando se maneja hidrógeno líquido. El hidrógeno líquido (liquid hydrogen) es un líquido sumamente frío, a unos 253 grados bajo cero. Es ligero debido a su baja densidad. Cuando el inductor gira a gran velocidad, se genera una zona donde la presión cae bruscamente en la parte delantera de los álabes. Los líquidos hierven si la presión desciende lo suficiente. Aunque no esté caliente, si la presión es baja, se forman burbujas de vapor. Es el mismo principio por el cual el agua hierve a menor temperatura en la montaña. Este fenómeno de formación de burbujas dentro de un líquido en movimiento se denomina cavitación (cavitation).

La cavitación perturba gravemente al inductor. Cuando las burbujas formadas fluyen hacia una zona donde la presión vuelve a ser alta, colapsan en un instante. Al desintegrarse las burbujas, se generan ondas de presión extremadamente intensas. Estas ondas erosionan poco a poco la superficie del álabe. Con el tiempo, el álabe sufre picaduras y desgaste. La cavitación también degrada el rendimiento de la bomba. Si el número de burbujas aumenta, la presión que impulsa la bomba colapsa repentinamente. Superar este punto de colapso pone en peligro a todo el motor. Por ello, el inductor debe diseñarse para retrasar al máximo la aparición de la cavitación.

Existe una propiedad que resulta de gran ayuda. Los líquidos criogénicos presentan un efecto de supresión termodinámica. Para que se forme una burbuja, esta debe absorber calor latente de vaporización del líquido circundante. Esto provoca que la temperatura alrededor de la burbuja descienda ligeramente. Al bajar la temperatura, la presión de ebullición en ese punto también disminuye. En consecuencia, a la burbuja le cuesta más seguir creciendo. Es como si la burbuja frenara su propio crecimiento. En el hidrógeno líquido, este efecto de supresión es notable, lo que retrasa ligeramente la cavitación. El diseño del inductor debe tener en cuenta este efecto en sus cálculos. También es necesario incorporar este efecto en las reglas físicas del aprendizaje mejorado por la física. De lo contrario, se subestimaría erróneamente el rendimiento frente a la cavitación respecto a la realidad.

El diseño de un inductor consiste en definir la forma de los álabes. El ángulo del álabe, la curvatura y el grado de flecha hacia adelante determinan el rendimiento. Hay tres objetivos principales: debe elevar

adecuadamente la presión, debe tener una alta eficiencia y debe retrasar la cavitación. Estos objetivos también entran en conflicto entre sí. Si se inclina el álabe para elevar fuertemente la presión, la cavitación se genera con facilidad. Si se tumba el álabe para evitar la cavitación, la presión aumenta en menor medida. Se trata del mismo dilema de equilibrio tratado en las secciones anteriores.

La capacidad del inductor para elevar la presión tiene un límite. Supongamos que operamos la bomba reduciendo gradualmente la presión de entrada. Hasta cierto punto, la bomba resiste bien. Sin embargo, cuando la presión de entrada cae por debajo de cierto valor, la cavitación se intensifica. En ese instante, la presión generada por la bomba colapsa de forma abrupta. Este punto de ruptura se conoce como punto de corte de altura manométrica. La altura manométrica es la magnitud de la fuerza con la que la bomba impulsa el fluido. En las bombas reales, se distingue entre el punto donde aparecen las primeras burbujas y el punto donde la altura manométrica colapsa por completo. Entre ambos, a menudo se toma como referencia el punto en el que la altura cae un 3 por ciento. Este valor se denomina NPSH3. La cavitación incipiente, el NPSH3 y el colapso total de la altura son etapas distintas. Un buen inductor retrasa este punto de corte hacia presiones de entrada más bajas. Esto le permite operar de manera estable incluso a presiones muy reducidas. Esta propiedad repercute directamente en el peso del cohete. Si el inductor soporta presiones bajas, es posible mantener una menor presión en el tanque de combustible. Si la presión del tanque es baja, las paredes del tanque pueden fabricarse más delgadas. Al aligerar el tanque, se reduce el peso de todo el cohete. En definitiva, el diseño de un pequeño álabe llega a transformar el peso total del cohete.

Aquí es donde se aplica el aprendizaje automático mejorado por la física (physics-enhanced machine learning). El aprendizaje automático basado exclusivamente en datos puede violar las leyes físicas. Por ello, se incorporan las leyes de conservación de fluidos y la física de la cavitación en el entrenamiento. El campo de flujo se aproxima rápidamente mediante el operador neuronal de Fourier (Fourier neural operator, FNO). Este operador es una red neuronal capaz de predecir el panorama completo del flujo de una sola vez y con gran rapidez. Mientras que una red neuronal convencional predice el valor en un solo punto, este operador predice la totalidad del campo de flujo. Al descomponer y procesar el flujo en ondas de diversas escalas, captura tanto el flujo general como las pequeñas perturbaciones. Por tanto, es ideal para estimar con rapidez valores distribuidos en un espacio amplio, como los campos de flujo. Permite evaluar velozmente el flujo y el riesgo de cavitación variando la geometría del álabe. Es muchísimo más rápido que ejecutar cálculos de alta precisión en cada ocasión. El operador neuronal de Fourier es, en sí mismo, un método validado en diversos problemas de dinámica de fluidos. Su aplicación al diseño inverso de inductores de hidrógeno líquido es la dirección que propone este libro.

En el diseño de inductores, la inteligencia artificial analiza múltiples informaciones. Observa la distribución de presión en la parte delantera del álabe. Examina dónde disminuye drásticamente la presión y cuánto podría desarrollarse la cavitación en ese lugar. También analiza en qué medida se recupera la presión en la parte posterior del álabe. Integrando esta información, busca una geometría óptima de los álabes. Una forma de álabe con flecha hacia adelante contribuye a retrasar la cavitación. La inteligencia artificial genera múltiples variantes de estas geometrías y las compara rápidamente.

La inteligencia artificial también se utiliza para diagnosticar la cavitación. Un estudio de 2025 grabó la cavitación de un inductor con una cámara de alta velocidad y analizó las imágenes mediante redes neuronales. Se emplearon conjuntamente una red neuronal para analizar patrones espaciales y otra para observar variaciones temporales. De este modo, es posible detectar el instante en que la cavitación se vuelve inestable. La inteligencia artificial capta cambios rápidos difíciles de percibir a simple vista. Esta investigación corresponde al ámbito del diagnóstico de la cavitación. Su propósito difiere tanto de los modelos subrogados que calculan el flujo como del diseño inverso que busca la forma óptima de los álabes. Es necesario distinguir estas tres aplicaciones sin mezclarlas.

El diseño de un inductor no concluye con el cálculo del flujo. Dado que los álabes giran a altísima velocidad, soportan fuerzas mecánicas considerables. Las ondas de choque generadas por el colapso de las burbujas de cavitación también impactan contra los álabes. Por lo tanto, se debe verificar simultáneamente si los álabes pueden resistir estas fuerzas. Si el diseño de canales de refrigeración es un equilibrio entre calor y flujo, el diseño del inductor es un equilibrio entre flujo, esfuerzos mecánicos y cavitación. La inteligencia artificial se emplea para sopesar todos estos objetivos de manera conjunta.

Aquí también las limitaciones son evidentes. Las pruebas con hidrógeno criogénico son complejas y costosas. Existen muy pocas instalaciones de prueba capaces de manejar líquidos a más de 250 grados bajo cero. Por ello, los datos de validación para el diseño de inductores no son abundantes. Las cifras de reducción de la cavitación solo son fiables cuando están respaldadas por pruebas reales de bombas. Por esta razón, este libro no incluye valores no verificados. En su lugar, se enfoca en qué información analiza el modelo mejorado por la física para asistir en el diseño. Aunque la inteligencia artificial reduzca la lista de candidatos óptimos de álabes, la decisión final debe tomarse mediante pruebas reales de la bomba.

Resumen

Este capítulo ha examinado los métodos para diseñar estructuras de gestión térmica de motores cohete mediante inteligencia artificial. Se abordaron tres estructuras: los canales de refrigeración internos en las paredes, los orificios de refrigeración por película que crean una capa protectora sobre la pared, y el inductor previo a la bomba que impulsa el combustible. Las tres estructuras comparten un problema de la misma naturaleza: el dilema del equilibrio entre calor y flujo. Ganar mucho en un aspecto implica perder en el otro. La inteligencia artificial se ha introducido para encontrar con rapidez este punto de equilibrio.

En la base de estos avances han tenido lugar dos grandes transformaciones. La primera es la impresión 3D de metales. Ha hecho posible fabricar canales internos complejos que no podían elaborarse mediante mecanizado tradicional, ampliando enormemente la libertad de diseño. La segunda es el surgimiento de nuevos materiales, como la aleación de cobre GRCop-42, que conducen el calor de forma excepcional. El encuentro entre la libertad de diseño y los materiales de alto rendimiento ha hecho viables nuevas estructuras de refrigeración. No obstante, dado que este nuevo método de fabricación es sumamente sensible al control de calidad, es imprescindible supervisar minuciosamente el proceso de producción.

La inteligencia artificial se ha empleado de múltiples formas. Los modelos subrogados sustituyeron los lentos cálculos de fluidos para aproximar con rapidez el comportamiento del combustible supercrítico. El aprendizaje activo mediante procesos gaussianos permitió ahorrar costosos cálculos numéricos al tiempo que identificaba orificios óptimos de refrigeración por película. Los autoencoders convolucionales redujeron geometrías de canales complejas a un espacio numérico compacto, facilitando el diseño topológico. La optimización bayesiana multiobjetivo encontró el equilibrio entre metas contrapuestas para suprimir la coquización. Por último, el aprendizaje automático mejorado por la física y los operadores neuronales de Fourier asistieron en el diseño para mitigar la cavitación en inductores.

Los métodos de este capítulo están estrechamente interconectados. El modelo subrogado estima rápidamente el flujo. El aprendizaje activo selecciona el siguiente cálculo a partir de dicha estimación. El autoencoder reduce el espacio de diseño para facilitar la exploración. La optimización bayesiana halla el punto de equilibrio entre múltiples objetivos. El aprendizaje mejorado por la física garantiza el

cumplimiento de las leyes físicas. No se trata de emplear una sola herramienta, sino de integrarlas en conjunto. Es tarea del diseñador seleccionar y combinar los métodos adecuados según el problema.

Existe un principio fundamental que lo articula todo: la inteligencia artificial no es una herramienta para decidir la respuesta final en nuestro lugar, sino un instrumento para reducir rápidamente el abanico de candidatos prometedores en un amplio espacio de diseño. Los modelos subrogados solo son confiables dentro del rango de datos con el que fueron entrenados. Las cifras de rendimiento presentadas en casos corporativos difieren de la validación académica. No se deben emplear valores de procedencia no verificada. La refrigeración de motores cohete es un ámbito en el que, si la pared se funde una sola vez, el motor entero queda destruido. Por ello, la decisión final siempre debe basarse en cálculos de alta fidelidad y pruebas de combustión reales.

Los métodos de este capítulo comparten un rasgo común: todos son recursos para ahorrar cálculos costosos. Los cálculos de alta precisión son exactos, pero lentos. Si se ejecutaran esos cálculos lentos en cada iteración, el diseño nunca concluiría. Por ello, la inteligencia artificial asume la tarea de realizar estimaciones rápidas. Los modelos subrogados reemplazan los cálculos directos, el aprendizaje activo selecciona la secuencia de cálculo y los autoencoders reducen la dimensión del problema. El tiempo ahorrado se aprovecha para explorar un mayor número de candidatos. Cuantos más candidatos se evalúan, mayor es la probabilidad de aproximarse a la solución óptima.

La dirección futura también se basa en este principio. El aprendizaje automático que incorpora un conocimiento físico más profundo está en desarrollo. El flujo que conecta el diseño y el diagnóstico también se está ampliando. La inteligencia artificial propone candidatos de diseño, las pruebas verifican esos candidatos y los datos de dichas pruebas entrenan de nuevo a la inteligencia artificial. Este ciclo impulsa poco a poco el diseño de la gestión térmica de los motores cohete. Si se resumiera este capítulo en una sola frase, sería la siguiente. Sobre una mayor libertad de diseño, la inteligencia artificial ayuda a resolver rápidamente el difícil problema de equilibrio entre el calor y el flujo. Y esa respuesta siempre se confirma mediante pruebas. Las herramientas se han vuelto más rápidas, pero el principio de verificación sigue siendo el mismo. El siguiente capítulo aborda cómo la inteligencia artificial diagnostica y controla el motor cuando realmente se opera un motor fabricado de esta manera.

Fuentes y referencias relacionadas

Pizzarelli, M. (2017). Regenerative cooling of liquid rocket engine thrust chambers: Modeling and simulation. Journal of Propulsion and Power, 33(4), 802-821. https://doi.org/10.2514/1.B36261

Waxenegger-Wilfing, G., Dresia, K., Deeken, J., & Oschwald, M. (2020). Heat transfer prediction for methane in regenerative cooling channels with neural networks. Journal of Thermophysics and Heat Transfer, 34(2), 347-357. https://doi.org/10.2514/1.T5865

Dresia, K., Waxenegger-Wilfing, G., Riccius, J., Deeken, J., & Oschwald, M. (2023). Improved wall temperature prediction for the LUMEN rocket combustion chamber with neural networks. Aerospace, 10(5), 450. https://doi.org/10.3390/aerospace10050450

Waxenegger-Wilfing, G., Dresia, K., Deeken, J., & Oschwald, M. (2021). Machine learning methods for the design and operation of liquid rocket engines: Research activities at the DLR Institute of Space Propulsion. arXiv. https://arxiv.org/abs/2102.07109

Gradl, P. R., Protz, C. S., Cooper, K., Ellis, D., & Evans, L. J. (2019). GRCop-42 development and hot-fire testing using additive manufacturing powder bed fusion for channel-cooled combustion chambers. AIAA Propulsion and Energy 2019 Forum. https://doi.org/10.2514/6.2019-4228

NASA NTRS. (2019). GRCop-42 development and hot-fire testing using additive manufacturing powder bed fusion for channel-cooled combustion chambers. Document ID 20190030433. https://ntrs.nasa.gov/citations/20190030433

Gradl, P., Mireles, O. R., Katsarelis, C., Smith, T. M., Sowards, J., Park, A., et al. (2023). Advancement of extreme environment additively manufactured alloys for next generation space propulsion applications. Acta Astronautica, 211, 483-497. https://doi.org/10.1016/j.actaastro.2023.06.035

NASA. (2023). GRCop-42 additive manufacturing for high heat flux thrust chamber liners. NASA Technical Reports Server. https://ntrs.nasa.gov/citations/20230007103

NASA. (2023). Rapid Analysis and Manufacturing Propulsion Technology (RAMPT) project. NASA. https://www.nasa.gov/rapid-analysis-and-manufacturing-propulsion-technology/

Demeneghi, G., Williams, B. B., Katsarelis, C., Tilson, W., & Gradl, P. (2025). Additively manufactured GRCop-42 copper-alloy combustion chamber failure analysis: The role of build interruptions. Engineering Failure Analysis, 174, 109509. https://doi.org/10.1016/j.engfailanal.2025.109509

Ferster, K. K., et al. (2025). An open tool for the design of cooling channels within regeneratively cooled liquid rocket engines. AIAA SciTech 2025 Forum. https://doi.org/10.2514/6.2025-2127

NIST. (2026). Thermophysical properties of fluid systems. NIST Chemistry WebBook. https://webbook.nist.gov/chemistry/fluid/

Banuti, D. T. (2020). Pseudo-boiling and heat transfer deterioration while heating supercritical liquid rocket engine propellants. Journal of Supercritical Fluids, 165, 104895. https://doi.org/10.1016/j.supflu.2020.104895

Zhu, B., Zhang, Q., Ma, L., & Shi, H. (2025). Analysis of heat transfer deterioration mechanism of supercritical fluid based on VOF two-phase method and pseudo-boiling theory. International Journal of Heat and Mass Transfer, 240, 126643. https://doi.org/10.1016/j.ijheatmasstransfer.2024.126643

Xiao, R., Zhang, P., Chen, L., & Hou, Y. (2023). Machine learning based prediction of heat transfer deterioration of supercritical fluid in upward vertical tubes. Applied Thermal Engineering, 228, 120477. https://doi.org/10.1016/j.applthermaleng.2023.120477

Gong, K., Zhang, Y., Cao, Y., Feng, Y., & Qin, J. (2024). Deep learning approach for predicting the flow field and heat transfer of supercritical hydrocarbon fuels. International Journal of Heat and Mass Transfer, 219, 124869. https://doi.org/10.1016/j.ijheatmasstransfer.2023.124869

Ko, S., Gil, Y., & Park, S. (2026). Development of RP-3 surrogate fuels via multi-objective genetic algorithm for regenerative cooling CFD with supercritical property fidelity. Aerospace, 13(4), 307. https://doi.org/10.3390/aerospace13040307

Zhang, H., Gou, J., Yin, P., Su, X., & Yuan, X. (2023). Film-cooling hole optimization and experimental validation considering the lateral pressure gradient. Frontiers in Mechanical Engineering, 8, 973293. https://doi.org/10.3389/fmech.2022.973293

Wang, Y., Wang, Z., Qian, S., Qiu, X., Shen, W., Zhang, X., Lyu, B., & Cui, J. (2024). Data-driven framework for prediction and optimization of gas turbine blade film cooling. Physics of Fluids, 36(3), 035160. https://doi.org/10.1063/5.0186087

Yao, G., Li, D., Zhu, J., Tao, Z., & Qiu, L. (2024). Hybrid AI framework for the predictions of film cooling effectiveness distribution with various surface curvatures and compound angles. Applied Thermal Engineering, 257, 124147. https://doi.org/10.1016/j.applthermaleng.2024.124147

Navah, F., Lamarche-Gagnon, M.-E., & Ilinca, F. (2023). Thermofluid topology optimization for cooling channel design. arXiv:2302.04745. https://arxiv.org/abs/2302.04745

Gao, C., Xu, W., Zhu, X., Cui, J., Luo, T., Wang, D., Sun, L., Ling, W., Li, X., & Zhou, W. (2025). Enhanced regenerative cooling performance with conformal TPMS channels. Energy, 322, 135530. https://doi.org/10.1016/j.energy.2025.135530

Fu, R., Miao, X., Gao, L., Duan, L., Xu, Y., Guo, M., Luo, P., & Hu, J. (2026). Numerical study of flow and heat-transfer enhancement in lattice-core regenerative cooling channels for liquid rocket engines. International Communications in Heat and Mass Transfer, 175, 111069. https://doi.org/10.1016/j.icheatmasstransfer.2026.111069

Ebbs-Picken, T., Romero, D. A., Da Silva, C. M., & Amon, C. H. (2024). Deep encoder-decoder hierarchical convolutional neural networks for conjugate heat transfer surrogate modeling. Applied Energy, 372, 123723. https://doi.org/10.1016/j.apenergy.2024.123723

Lee, J., Shin, S., Choi, H., Lee, A., Park, B., & Lee, S. (2025). Extended multiphysics-informed neural network for conjugate heat transfer problems. International Journal of Heat and Mass Transfer, 246, 127098. https://doi.org/10.1016/j.ijheatmasstransfer.2025.127098

Lee, S., Yoon, Y., & Song, S. J. (2025). Deep learning-based identification of turbopump inducer cavitation instability using high-speed camera images. AIAA SciTech 2025 Forum. https://doi.org/10.2514/6.2025-0760

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier neural operator for parametric partial differential equations. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.2010.08895

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

DLR Institute of Space Propulsion. (2026). Test facilities at Lampoldshausen and rocket propulsion research. DLR. https://www.dlr.de/en/ra

Capítulo 10. Diagnóstico de fallos en tiempo real, gemelo digital y control adaptativo de propulsión en tiempo real basado en aprendizaje por refuerzo para motores cohete de propelente líquido (LRE)

Introducción

Cuando un conductor maneja un automóvil durante mucho tiempo, conoce su estado por el sonido. Si el motor produce un ruido extraño, acude al taller. Si se enciende una luz de advertencia en el panel de instrumentos, reduce la velocidad. Las personas observan y juzgan múltiples señales a la vez de esta manera. Con los motores de cohete ocurre lo mismo. Sin embargo, un motor de cohete es mucho más rápido, más caliente y más peligroso que un automóvil.

Un motor cohete de propelente líquido quema combustible líquido y oxidante líquido para generar una gran fuerza. Este motor maneja una enorme cantidad de calor y presión en un breve lapso. Si un solo componente se desajusta ligeramente, todo el conjunto puede venirse abajo. Por eso es importante vigilar continuamente su estado mientras el motor permanece encendido. Esta labor de supervisar el estado durante todo el funcionamiento del motor se denomina monitorización del estado en tiempo real.

Este capítulo aborda cómo la inteligencia artificial supervisa y regula los motores de cohete. La inteligencia artificial se denomina en inglés Artificial Intelligence, abreviado como AI. Trata principalmente cuatro aspectos. El primero es el diagnóstico para detectar con rapidez anomalías en las señales de los sensores. El siguiente es el gemelo digital (Digital Twin), que modela computacionalmente el interior invisible del motor. Otro es el método para cuantificar la incertidumbre, incluso cuando el motor se desgasta o presenta tolerancias de fabricación. El último es el control inteligente que coordina y regula múltiples válvulas simultáneamente.

Cuando estas tecnologías se integran, el motor cuida de su propio estado de salud. Esta tendencia es indispensable en la era de los cohetes reutilizables. Si se utiliza un mismo motor varias veces, su condición varía ligeramente en cada uso. Para un ser humano resulta difícil ajustarlo manualmente en cada ocasión. La inteligencia artificial es una herramienta que interpreta rápidamente esos cambios y responde a ellos. A partir de ahora, se explicarán paso a paso los principios de funcionamiento de estas tecnologías.

Antes de leer este capítulo, conviene adoptar una premisa. La inteligencia artificial no es magia. Es una herramienta que examina datos para descubrir patrones. Si los datos son deficientes, los resultados también lo serán. Por ello, este capítulo expone la utilidad de cada tecnología junto con sus limitaciones. No se emplean cifras de rendimiento no verificadas. En su lugar, se dedica el esfuerzo a explicar los principios de manera accesible. El objetivo es que el lector comprenda la estructura básica de la tecnología.

10.1 Desafíos de la monitorización del estado en tiempo real de los LRE

Contenido de esta sección

Esta sección explica por qué resulta tan difícil vigilar un motor de cohete en tiempo real. Primero se describe lo extremo que es el entorno de un motor cohete de propelente líquido. A continuación, se señalan las limitaciones de los métodos de monitorización utilizados hasta ahora. Por último, se expone la necesidad de la monitorización basada en inteligencia artificial.

Primero aclararemos un término empleado aquí. La sigla LRE, que aparece con frecuencia en este capítulo, es la abreviatura de motor cohete de propelente líquido (Liquid Rocket Engine). En lo sucesivo, se denominará motor cohete de propelente líquido o simplemente motor. Este motor se utiliza ampliamente para elevar cargas pesadas al espacio.

Un motor cohete de propelente líquido es una máquina compleja en la que múltiples componentes funcionan conjuntamente. La cámara de combustión es el recinto donde se queman el combustible y el oxidante. Los inyectores son componentes que pulverizan y mezclan finamente el combustible y el oxidante. La turbobomba es la bomba que impulsa los propelentes a alta presión. Las válvulas abren y cierran el paso del propelente. Los canales de refrigeración disipan el calor para evitar que las paredes se fundan. Estos componentes deben operar como un solo cuerpo para que el motor funcione correctamente.

El problema radica en que este entorno es sumamente severo. La presión dentro de la cámara de combustión alcanza decenas de atmósferas. La temperatura de la llama supera los 3000 grados. La turbobomba gira decenas de miles de revoluciones por minuto. Estos valores pertenecen a una dimensión completamente distinta a la de una olla a presión doméstica o el motor de un automóvil. Incluso una pequeña vibración se propaga con una fuerza inmensa. Por eso, el motor opera siempre bajo el riesgo constante de sufrir averías.

Este entorno hostil resulta también implacable para los sensores. Los sensores son los ojos que perciben el estado. Sin embargo, si un sensor se encuentra cerca de una zona de alta temperatura, puede dañarse a sí mismo. Si la vibración es intensa, los cables pueden romperse. Es como tener que observar el motor con la vista empañada. Por ello, resulta difícil confiar en el valor de un solo sensor. Es necesario analizar y filtrar múltiples sensores de manera conjunta. Aunque un sensor presente anomalías, se verifica con otros sensores.

El volumen de señales tampoco es despreciable. Un solo sensor de muestreo denso genera decenas de miles de valores por segundo. En un motor se instalan de decenas a cientos de estos sensores. Esto produce una avalancha de millones de valores por segundo. Un ser humano no puede examinar visualmente semejante cantidad de datos. Además, la evaluación debe completarse en milisegundos. Un milisegundo es una milésima de segundo (1/1000 de segundo). Es necesario tomar decisiones cientos de veces en el lapso de un parpadeo. Esta velocidad y volumen no pueden gestionarse sin la ayuda de las máquinas.

Veamos un ejemplo de la rapidez con la que una anomalía del motor puede transformarse en un accidente. En la cámara de combustión, la presión comienza a oscilar de forma periódica y pronunciada. Esta vibración puede multiplicarse en cuestión de pocos milisegundos. La vibración intensificada golpea los inyectores y las paredes. Si los impactos se repiten, el metal se fisura. A través de las grietas se escapan gases calientes. Este proceso se desencadena en un parpadeo. Por eso, es fundamental captar los primeros indicios de anomalía lo antes posible.

Otro factor de dificultad radica en que las condiciones de las pruebas en tierra y del vuelo son distintas. En las pruebas en tierra, el motor se fija rígidamente. Se pueden instalar numerosos sensores. Si surge

algún problema, se puede apagar de inmediato. Durante el vuelo, la situación cambia. Por razones de peso, no es posible montar muchos sensores. Tampoco hay personas cerca que puedan vigilarlo. Por lo tanto, el motor en vuelo debe tomar decisiones de manera autónoma con un número reducido de sensores, y la inteligencia artificial asiste en esta determinación.

El diagnóstico de motores de cohete persigue el mismo objetivo que el de otras máquinas. Sin embargo, sus condiciones son incomparablemente más severas. El margen de tiempo es breve, los datos son escasos y el peligro es elevado. Por eso no es posible trasladar directamente los métodos de otros campos. Es preciso reelaborar y perfeccionar las metodologías para adaptarlas a los cohetes. Las tecnologías que presenta este capítulo son el resultado de dicho refinamiento, un trabajo que continúa desarrollándose en la actualidad.

Las pruebas en tierra son igualmente peligrosas. La prueba en la que se fija el motor al suelo y se enciende se denomina prueba de fuego estático (hot-fire). En este ensayo, el motor produce la misma potencia que en vuelo real. En el espacio, la situación es todavía más delicada. Para que el cohete vuelva a aterrizar, el motor debe apagarse y reencenderse en el aire. En ese instante, el propelente se agita dentro del tanque. Este vaivén se denomina chapoteo o sloshing. Los tiempos de apertura de las válvulas también presentan ligeras desviaciones. Estas perturbaciones repentinas surgen y desaparecen en fracciones de segundo, por lo que resulta fácil pasarlas por alto.

Resulta fácil de comprender si se imagina la escena del aterrizaje de un cohete. Durante el despegue, el motor entrega su empuje máximo. Al aterrizar, debe reducir sustancialmente la potencia para posarse con suavidad. En este proceso, el motor modula su empuje en un amplio rango. El líquido dentro del tanque de propelente también se agita simultáneamente. Las válvulas se abren y se cierran a gran velocidad. La temperatura y la presión experimentan cambios drásticos. En este instante tan breve y riguroso es muy probable que surjan anomalías. Por ello, es imperativo vigilar atentamente este momento sin perder ningún detalle.

El reencendido del motor en pleno vuelo también resulta sumamente complejo. En órbita, la gravedad terrestre sigue siendo considerable. Sin embargo, dado que el cohete y el propelente caen juntos en caída libre, apenas se experimenta peso. Este estado se conoce como microgravedad. En condiciones de microgravedad, el líquido no permanece en su lugar dentro del tanque. Por ello, se debe acumular el propelente en un extremo mediante una pequeña aceleración previa o recurriendo a la tensión superficial. Si se enciende el motor sin este procedimiento, el flujo resulta irregular. La presión y el caudal sufren picos repentinos. En tales estados transitorios, el diagnóstico puede desorientarse con facilidad. Se podría catalogar como anomalía lo que en realidad es un comportamiento normal. Por esta razón, el sistema de diagnóstico debe ser entrenado para tolerar y discernir estos estados transitorios.

Hasta ahora, la monitorización de motores ha empleado predominantemente el método de valores límite. Es un sistema que apaga el motor de inmediato en cuanto cierto parámetro rebasa una línea predeterminada. A este umbral se le denomina línea roja (redline). Por ejemplo, si la presión excede un valor fijado, se detiene el motor. Este método es fácil de comprender y veloz, por lo que se ha utilizado desde hace mucho tiempo como un procedimiento fiable.

El método de la línea roja se parece al disyuntor o cuadro de fusibles de una casa. Si se consume demasiada electricidad, el disyuntor salta. Es una modalidad que desconecta todo incondicionalmente al superar el límite. Este enfoque es excelente para prevenir catástrofes de gran magnitud. La decisión es simple, por lo que resulta rápida y concluyente. Los motores de cohete también incorporan este corte de seguridad como norma básica. En cuanto un valor cruza la línea de peligro, el motor se apaga de inmediato.

Sin embargo, el método de la línea roja pasa por alto con facilidad las señales leves de anomalía. Las señales de los sensores siempre vienen acompañadas de ruido. Si el ruido es elevado, la alarma puede sonar incluso cuando todo opera con normalidad. A esto se le llama falsa alarma. Cuando las falsas alarmas son frecuentes, los operadores dejan de confiar en ellas. Por el contrario, una microvibración en los inyectores o un desgaste incipiente en los sellos permanecen ocultos por debajo de los valores límite. Estas señales tenues apenas son perceptibles hasta que crecen, y cuando ya son evidentes, resulta demasiado tarde para actuar.

El diagnóstico mediante inteligencia artificial busca subsanar esta deficiencia. Un ser humano solo puede vigilar uno o dos indicadores a la vez. La inteligencia artificial observa simultáneamente las señales de cientos de sensores. Analiza conjuntamente el comportamiento y la interrelación de múltiples señales. Gracias a ello, puede detectar indicios de fallo antes de que los valores sobrepasen la línea límite. Un artículo de revisión publicado en una revista académica aeroespacial en 2026 resume esta evolución. Dicho estudio explica que los métodos basados en datos han crecido considerablemente en el diagnóstico de motores cohete de propelente líquido. Al mismo tiempo, recalca la dificultad que plantea la escasez de datos sobre averías reales.

En resumen, la monitorización de motores de cohete se enfrenta a tres barreras. Las señales son rápidas y ruidosas. Las anomalías leves se manifiestan tardíamente. Además, es necesario mantener un equilibrio entre las falsas alarmas y las omisiones. La inteligencia artificial constituye un esfuerzo por sortear simultáneamente estas tres barreras. A partir de la próxima sección, se examinarán en detalle y uno a uno dichos métodos.

La monitorización con inteligencia artificial no descarta el método de la línea roja. Ambos enfoques se emplean conjuntamente. La línea roja es la última línea defensiva para evitar accidentes graves. La inteligencia artificial detecta los indicios tempranos por delante de ella. Constituye, en efecto, una red de seguridad estratificada. Lo que se escapa de la red exterior lo atrapa la red interior. La seguridad de un cohete se garantiza superponiendo múltiples métodos de este modo, sin depender exclusivamente de uno solo.

10.2 Detección temprana de anomalías mediante la observación conjunta de múltiples sensores

Contenido de esta sección

Esta sección trata sobre las arquitecturas de inteligencia artificial que detectan oportunamente anomalías a partir de las señales de los sensores del motor. Primero se explican el concepto de telemetría y las señales de ultraalta frecuencia. Seguidamente, se presenta una red neuronal que evalúa tanto la evolución temporal como las relaciones entre canales. A esta estructura se la denomina red de atención y convolución temporal (TCAN). Esta arquitectura no es un modelo validado exclusivamente para cohetes, sino un enfoque empleado en diversas disciplinas. Por último, se examinan también las limitaciones que este tipo de diagnóstico aún debe superar.

Primero aclararemos los términos. La telemetría (telemetry) es la transmisión y recepción de señales que reflejan el estado de un dispositivo a distancia. En el motor se incorporan sensores para medir la presión, la vibración, la temperatura y el caudal. Los datos de estos sensores llegan a la computadora en tiempo real. Ultraalta frecuencia significa que estas señales se registran a intervalos sumamente estrechos. Ciertos sensores toman de decenas a cientos de miles de lecturas por segundo. Solo con un muestreo tan denso es posible capturar las vibraciones de alta velocidad.

Para medir la presión se emplean elementos piezoeléctricos. Un elemento piezoeléctrico es un material que genera una pequeña corriente eléctrica al ser sometido a presión. Cuando la presión cambia con rapidez, la señal eléctrica también varía velozmente. Por ello, este sensor registra con precisión las fluctuaciones rápidas de presión. Los acelerómetros miden la magnitud de las vibraciones en los componentes. Al utilizar ambos sensores en conjunto, se supervisan a la vez la presión y la vibración. Esto equivale a examinar el estado del motor desde múltiples perspectivas.

Una monitorización con lecturas lentas deja escapar las anomalías veloces. Supongamos que intentamos fotografiar el aleteo de un colibrí con una cámara antigua. Si el obturador es lento, las alas saldrán borrosas y desdibujadas. Se necesita un obturador veloz para congelar el instante del vuelo. Con los sensores pasa lo mismo. Si el intervalo de muestreo es amplio, las oscilaciones rápidas se desvanecen. Por tanto, para detectar las peligrosas vibraciones de alta frecuencia, es imprescindible un muestreo denso.

Veamos por qué se requiere una supervisión tan rápida. Las anomalías de un motor de cohete se agravan en un abrir y cerrar de ojos. Existe un fenómeno en el que la presión oscila de forma rítmica cerca de los inyectores. Esto se denomina chugging. Es una inestabilidad en la que el motor traquetea a baja frecuencia. En casos extremos, sacude y destruye válvulas y tuberías. Esta oscilación se propaga en instantes. Ocurre demasiado rápido para que el ojo humano lo detecte y reaccione. Por consiguiente, la máquina debe realizar la evaluación en escalas de milisegundos.

La anomalía de un componente puede manifestarse inicialmente en la señal de otro componente distinto. Por ejemplo, el eje de la turbobomba vibra primero. Poco después, esa oscilación desencadena una variación en la presión de la cámara de combustión. Para una persona resulta difícil advertir a simple vista este retardo temporal y su correlación. Dado que las redes neuronales analizan múltiples sensores de forma integrada, tienen la capacidad de aprender este desfase temporal y las interrelaciones entre canales.

La estructura que se introduce aquí combina dos métodos. Uno es la convolución temporal. La convolución temporal es un procedimiento para extraer patrones de la secuencia temporal de las señales. Permite observar tendencias prolongadas ensanchando progresivamente el alcance desde ventanas cortas hasta ventanas largas. El otro es la atención (attention). La atención es un mecanismo que decide en qué partes de las diversas señales concentrarse prioritariamente. Asigna mayor ponderación a los canales que resultan críticos en el momento presente. Al articular estos dos métodos, se examinan simultáneamente las oscilaciones breves y las relaciones entre canales. A esta configuración combinada se la conoce como red de atención y convolución temporal (TCAN).

Expliquemos con un poco más de detalle cómo la convolución temporal observa flujos largos. La convolución es un método para escanear una señal a través de una pequeña ventana. Si la ventana es estrecha, solo observa un instante breve. Las anomalías de los cohetes también pueden manifestarse a lo largo de un período prolongado. Por lo tanto, es necesario observar un intervalo de tiempo amplio. El método utilizado en este caso consiste en ampliar el intervalo de la ventana. Esto se denomina convolución dilatada. Escanea un tiempo amplio observando los valores de forma alternada, uno sí y otro no. Es el mismo principio que subir escalones de dos en dos para llegar más alto con la misma cantidad de pasos. De esta manera, se puede observar un flujo largo con menos cálculos.

Tomemos también como ejemplo la forma en que la atención distribuye la concentración. Pensemos en la escena de escuchar selectivamente la voz de un amigo en un aula ruidosa. El oído se concentra en el sonido importante entre muchos otros. La atención hace lo mismo. Otorga peso al canal que muestra claramente una anomalía en ese momento. Si la situación cambia, el peso también cambia. Por lo tanto,

selecciona y observa sobre la marcha las señales importantes entre múltiples sensores. La convolución temporal observa el flujo temporal y la atención selecciona los canales.

Este método de combinación se utiliza ampliamente en investigaciones recientes. Un estudio de 2025 propone una red neuronal convolucional temporal que añade atención. Dicho estudio explica que captura tanto las dependencias cortas como las largas. Otro estudio del mismo año combina la convolución temporal con la atención de grafos. Modela múltiples sensores como un grafo y aprende sus relaciones en conjunto. Estos estudios no son exclusivos para cohetes. Sin embargo, se adaptan bien al problema de detectar simultáneamente anomalías en múltiples sensores. El diagnóstico de motores cohete también es el mismo tipo de problema, en el que se deben hallar anomalías en múltiples sensores a la vez. Lo que se ha comprobado es que esta estructura detecta bien las anomalías en datos de series temporales generales. Todavía no existe una demostración empírica de que se aplique directamente a los sensores y tipos de fallas de los cohetes. Este libro propone esta estructura como una dirección para transferirla al diagnóstico de cohetes.

Otras investigaciones en el campo de los cohetes muestran una dirección similar. El diagnóstico mediante aprendizaje profundo que lee múltiples señales de sensores a la vez va en aumento. Un estudio de 2022 utiliza una red de memoria a corto y largo plazo interpretable. Esta red neuronal clasifica las anomalías del motor a partir de datos de múltiples sensores. Un estudio de 2024 aborda la detección de fallas basada en el aprendizaje de propiedades. Todas estas tendencias representan intentos de agrupar múltiples señales para detectar anomalías rápidamente.

También es importante el lugar donde se calculan estos diagnósticos. Si la señal se envía a una computadora lejana para su evaluación, toma tiempo. Para obtener una respuesta en milisegundos, el cálculo debe realizarse junto al motor. Este método de computar directamente al lado del dispositivo se denomina computación perimetral (edge computing). La red neuronal se implementa y ejecuta en un chip pequeño y dedicado. En cuanto recibe la señal, toma una decisión en el acto. Solo así es posible detectar el peligro en milisegundos. Por ello, la red neuronal de diagnóstico debe ser pequeña y rápida. Aunque una red neuronal grande sea precisa, su utilidad disminuye si es lenta.

El modelo de diagnóstico pasa generalmente por tres etapas. En primer lugar, acondiciona uniformemente las señales de los sensores. Esto se denomina estandarización. Se debe a que cada sensor tiene magnitudes y unidades diferentes. A continuación, corta una ventana temporal breve y la introduce en el modelo. Por último, clasifica el estado en varios tipos: si es normal, si se trata de una anomalía del sensor, una anomalía de la bomba o una oscilación de la combustión. El modelo de diagnóstico repite estas tres etapas sin descanso y a una velocidad extremadamente rápida.

Existen varios tipos de anomalías que este diagnóstico intenta captar. Está el chugging, que palpita con vibraciones de baja frecuencia cerca de los inyectores. Están las oscilaciones de alta frecuencia que resuenan con vibraciones más rápidas. Está el problema de fugas debido al desgaste de los sellos de la turbobomba. También existe el problema del aumento de vibraciones por el desgaste de los cojinetes. Cada anomalía deja una forma de señal diferente. El modelo de diagnóstico aprende estas diferencias en los patrones. Al observar qué patrón aparece, determina el tipo de anomalía.

Aquí hay un punto con el que se debe tener precaución. Algunos textos que tratan sobre esta estructura citan cifras como la precisión o el tiempo de latencia. Sin embargo, tales cifras no están confirmadas en artículos publicados. Utilizar cifras no verificadas como si fueran hechos puede inducir a error al lector. Por lo tanto, este libro no utiliza tales cifras. En su lugar, explica esta estructura únicamente como un principio para el diagnóstico temprano de fallas. La clave radica en observar múltiples sensores de forma conjunta para detectar las anomalías a tiempo.

En los motores reales, al gestionar los resultados del diagnóstico se clasifica el nivel de riesgo de la falla. Si se trata de una señal crítica en la que la presión supera con creces el umbral de peligro, se apaga el motor de inmediato con esa sola señal. Forzar una confirmación múltiple ante tales señales retrasaría la parada y aumentaría la gravedad del accidente. Por el contrario, ante indicios leves se toman medidas cuando múltiples señales apuntan a la misma anomalía. Esto se hace para evitar que el juicio erróneo de un único modelo arruine una prueba en perfecto estado. Por ello, la salida del diagnóstico se utiliza conjuntamente con la lógica de seguridad. Debe contarse tanto con la supervisión humana como con dispositivos de seguridad basados en reglas.

Para confiar en un modelo de diagnóstico, es necesario saber por qué emitió ese juicio. Una red neuronal es como una caja cuyo interior es difícil de ver. Da una respuesta, pero no explica el motivo. En entornos peligrosos como los cohetes, el motivo es fundamental. Por ello, aumentan las investigaciones que muestran qué sensor condujo a la decisión. Esto se denomina diagnóstico interpretable. Los humanos observan ese motivo y deciden si confiar en el diagnóstico. Poder conocer la razón es indispensable para utilizar la inteligencia artificial en decisiones de seguridad.

Otra dificultad reside en la escasez de datos. Los datos reales de fallas en motores cohete son extremadamente raros, ya que una falla representa un accidente muy costoso. Por esta razón, abundan los datos de funcionamiento normal, pero casi no hay datos de fallas. La inteligencia artificial difícilmente puede aprender bien con pocos datos. Para resolver esto, los investigadores emplean diversos métodos. Uno de ellos consiste en generar datos sintéticos que imitan situaciones de falla mediante modelos físicos. Otro consiste en introducir deliberadamente fallas predeterminadas en modelos de prueba. Este procedimiento de insertar fallas a propósito se denomina inyección de fallas. La generación de datos sintéticos y la inyección de fallas tienen significados diferentes. La escasez de datos sigue siendo un gran desafío en este campo.

También existe un método para transferir lo aprendido en un motor a otro. Esto se denomina aprendizaje por transferencia. Es similar a cómo quien sabe andar en bicicleta aprende a conducir una motocicleta con mayor rapidez. Se transfiere el conocimiento previamente adquirido a una nueva situación. De este modo, aprende rápido aunque haya pocos datos del nuevo motor. Un artículo de revisión publicado en 2026 en una revista académica aeroespacial también señala esta dificultad: que los datos reales de fallas son escasos y los tipos de fallas son diversos. Dicho artículo sintetiza las investigaciones sobre diagnóstico centrándose en estos tres problemas prácticos.

El objetivo del diagnóstico no consiste únicamente en prevenir accidentes. También se utiliza para realizar el mantenimiento adecuado. Si no hay anomalías, no se reemplazan piezas innecesariamente. Si se aprecian indicios de anomalía, se intervienen de antemano. Este método de mantenimiento basado en la observación del estado se denomina mantenimiento basado en la condición. Los motores reutilizables reducen costos mediante este método. El diagnóstico es una herramienta que protege tanto la seguridad como los costos.

10.3 Gemelo digital que dibuja el interior invisible de la cámara de combustión

Esta sección aborda las oscilaciones peligrosas dentro de la cámara de combustión. Estas oscilaciones se denominan inestabilidad de la combustión. En primer lugar, se explica cómo el sonido y la llama intensifican mutuamente su fuerza. A continuación, se examina el método para reconstruir por computadora el flujo invisible de presión. Este método utiliza una red neuronal que respeta las ecuaciones físicas. Por último, se analizan la utilidad y las limitaciones de este método.

Aclaremos primero los términos. La inestabilidad de la combustión es un fenómeno en el que la presión oscila fuertemente dentro de la cámara de combustión. Si la llama y el sonido se coordinan, esta oscilación aumenta. Una pequeña oscilación de presión perturba la liberación de calor de la llama. La llama perturbada incrementa a su vez la oscilación de presión. Cuando esta retroalimentación se repite, la oscilación crece como una bola de nieve. En casos graves, se destruyen las paredes de la cámara de combustión y los inyectores. En la historia de los cohetes, este fenómeno ha sido un problema crónico.

Esta oscilación fue un gran problema desde los inicios del desarrollo de los cohetes. Incluso los grandes motores que llevaron al ser humano a la Luna sufrieron por esta causa. La presión resonaba fuertemente dentro de la cámara de combustión y destruía los motores. Los investigadores solucionaron esto modificando repetidamente la forma de los inyectores. Así de compleja es la gestión de las oscilaciones de combustión. Todavía hoy se comprueba este problema cada vez que se construye un motor nuevo.

Existen varias formas de oscilación. Pensemos en la cámara de combustión como un cilindro. El sonido puede resonar a lo largo del cilindro. También puede resonar girando a lo largo de la circunferencia. O puede resonar atravesando el diámetro. Cada modo tiene una frecuencia diferente. Si se sabe qué modo se está amplificando, responder resulta más sencillo. Por ello, conocer la forma y la intensidad de la oscilación es fundamental.

Existe un principio clásico que explica esta retroalimentación. Si se libera más calor en el instante en que la presión aumenta, la oscilación se intensifica. Este principio se conoce como el criterio de Rayleigh. Resulta sencillo si se imagina empujar un columpio. Si se empuja cuando el columpio va hacia adelante, este sube cada vez más alto. Si se empuja en el momento opuesto, el columpio se detiene. Por el mismo principio, la llama y el sonido se amplifican o se atenúan mutuamente. El momento en que intercambian fuerza determina la estabilidad y la inestabilidad.

Esta retroalimentación también se asemeja a la relación entre un diapasón y un micrófono. Si en un auditorio se acerca un micrófono a un altavoz, se produce un sonido agudo. Esto se debe a que el sonido del altavoz entra al micrófono y se amplifica nuevamente. Esta retroalimentación hace crecer el sonido como una bola de nieve. En la cámara de combustión, la llama actúa como el altavoz y el sonido como el micrófono. Si las fases de ambos coinciden, el sonido se amplifica de forma incontrolable. Por ello, se requiere un diseño que interrumpa esta retroalimentación.

El problema radica en que resulta difícil observar directamente el interior de la cámara de combustión. El interior está demasiado caliente y la presión es sumamente alta. No se pueden colocar sensores en cualquier lugar. Apenas se pueden instalar sensores de presión en unos pocos puntos de las paredes. Sin embargo, las oscilaciones peligrosas ocurren en toda la cámara de combustión. Solo con unos pocos puntos en la pared resulta difícil conocer el panorama completo. Por ello, se necesita un método para reconstruir la totalidad a partir de unas pocas observaciones.

Aquí es donde entra en juego el gemelo digital. Un gemelo digital es un modelo computacional que reproduce fielmente el dispositivo real. Mientras el motor real opera, el modelo también se ejecuta a la par. Recibe los valores de los sensores y ajusta continuamente el modelo a la realidad. De este modo, es posible inferir el estado incluso en lugares donde no hay sensores. Recibe este nombre porque es un modelo que opera en paralelo con la realidad, como un gemelo.

El gemelo digital tiene similitudes con el navegador GPS de un automóvil. El navegador recibe las condiciones reales del camino y redibuja la ruta. El gemelo digital también recibe las señales del motor real y redibuja su estado. La diferencia radica en la velocidad y la precisión. El gemelo de un cohete debe sincronizar su estado cada milisegundo. Solo así puede seguir el ritmo de las oscilaciones rápidas. Para alcanzar esta velocidad, el modelo debe hacerse ligero.

Descubrir el interior a partir de unos pocos sensores en la pared es un problema inverso. Normalmente se conoce la causa y se calcula el resultado. Aquí solo se conoce el resultado: la presión en la pared. A partir de ese resultado se rastrea la causa: el campo de presión interno. Este tipo de problema que busca la causa a partir del resultado se denomina problema inverso. Los problemas inversos son complejos porque pueden tener múltiples soluciones. Las ecuaciones físicas descartan las demás y conservan solo la respuesta acorde con la física.

El método para dotar de contenido a este gemelo digital es la red neuronal basada en la física. Esta red neuronal se denomina red neuronal informada por la física (Physics-Informed Neural Network, PINN). En adelante, nos referiremos a ella como red neuronal física. Las redes neuronales comunes aprenden basándose únicamente en datos. La red neuronal física aprende respetando tanto los datos como las ecuaciones de la física. Si infringe las ecuaciones durante el entrenamiento, se le impone una penalización. Por ello, arroja respuestas conformes con la física aun con pocos datos.

Veamos un poco más en detalle la forma en que aprende la red neuronal física. Las redes neuronales comunes solo reducen la diferencia con la respuesta correcta. La red neuronal física reduce dos factores a la vez: por un lado, la diferencia con los valores de los sensores; por el otro, el grado de incumplimiento de las ecuaciones físicas. Aprende a disminuir conjuntamente ambas penalizaciones. Así, las respuestas respetan la física incluso en los lugares donde no hay sensores. Las reglas físicas rellenan, por así decirlo, los vacíos de los datos. Es similar a cómo las líneas del dibujo guían los colores en un libro para colorear.

Este método puede convertirse en una herramienta de supervisión económica. Para calcular con precisión el interior de la cámara de combustión, una computadora de gran tamaño debe operar durante mucho tiempo. Durante una prueba no se dispone de ese tiempo. La red neuronal física, una vez entrenada, ofrece respuestas rápidamente. Por lo tanto, puede representar el estado interno incluso mientras se realiza la prueba. Con esta rápida representación se observan los indicios de oscilaciones peligrosas. En la práctica, permite observar el interior sin necesidad de detener la prueba.

La ecuación que describe los modos acústicos de la cámara de combustión es la ecuación de Helmholtz. Esta ecuación indica de qué manera resuena el sonido dentro de un espacio cerrado. Es similar al principio por el cual ciertas notas resuenan con fuerza dentro de un instrumento de viento. La red neuronal física aprende respetando esta ecuación y las condiciones de contorno. Recibe como entrada los valores de unos pocos sensores en la pared. A partir de ello, reconstruye el flujo de presión en toda la cámara de combustión. Muestra, como en una imagen, en qué lugares las oscilaciones se intensifican fuertemente.

El sonido en la cámara de combustión varía según cómo responda en los límites. La cara del inyector y la garganta de la tobera reflejan en cierta medida el sonido y en cierta medida lo absorben. El grado de esta reflexión y absorción se denomina condición de contorno acústica. La red neuronal física aprende respetando también estas condiciones de contorno. Solo así dibuja el sonido de una cámara de combustión real y no el de una habitación vacía. El criterio de Rayleigh visto anteriormente y estas condiciones de contorno determinan la inestabilidad de la combustión en los cohetes.

Esta dirección es muy activa en las investigaciones recientes. Las redes neuronales físicas tienen la ventaja de reconstruir el campo de presión acústica a partir de pocas observaciones. Esto se debe a que la presión solo se mide en unos pocos puntos y la velocidad apenas se puede medir. Un estudio de 2024 aprende la interacción termoacústica de una cámara de combustión mediante redes neuronales físicas. La termoacústica se refiere al fenómeno en el que el calor y el sonido se entrelazan. Este estudio extrae de los datos valores como la amortiguación de la presión y el acoplamiento de liberación de calor. Un estudio de 2026 estima las propiedades de absorción acústica de la pared de la cámara de combustión

utilizando redes neuronales físicas de descomposición de dominio. Estos estudios abordan el aprendizaje termoacústico y la estimación de propiedades de contorno en cámaras de combustión convencionales. Todavía no existe una demostración práctica que reconstruya en tiempo real todo el campo de presión con pocos sensores en una cámara de combustión real de un motor cohete de propelente líquido. Este libro propone dicha reconstrucción en tiempo real como la dirección a seguir en el futuro.

Las redes neuronales físicas también se utilizan en la investigación de la combustión limpia para la propulsión aeroespacial. Un artículo de 2026 considera esta tendencia como un camino hacia la propulsión sostenible. Conecta las ecuaciones de la física con los datos para describir la combustión con mayor precisión. Este tipo de investigación se extiende no solo a los cohetes, sino también a los motores aeronáuticos. La idea de redes neuronales que respetan la física se está extendiendo a diversos campos.

Los gemelos digitales se utilizan de forma pionera en el mantenimiento de motores aeronáuticos. Cierto motor aeronáutico de gran tamaño detectó diversas anomalías por adelantado mediante un gemelo digital. Se sabe que predijo el desgaste de los álabes de la turbina y problemas de combustión. Este caso demuestra que los modelos de monitorización son verdaderamente útiles. Los motores cohete operan en condiciones más severas que los motores aeronáuticos y disponen de menos datos. Por ello, trasladar este enfoque a los cohetes requiere una mayor verificación.

Las redes neuronales físicas también tienen puntos débiles. Incorporar ecuaciones al aprendizaje hace que el entrenamiento sea más complejo. Es necesario calibrar adecuadamente los pesos entre la penalización física y la penalización de datos. Si los pesos se desajustan, la respuesta se sesga hacia un lado. Cuanto más compleja es la combustión que se aborda, más difícil resulta este ajuste. Por lo tanto, las redes neuronales físicas no son una solución universal. Es necesario seleccionar las ecuaciones y la arquitectura adecuadas para cada problema. Para esta selección, el conocimiento humano sigue siendo indispensable.

En una perspectiva más amplia, el aprendizaje automático basado en la física crece rápidamente en toda la investigación sobre combustión. Un artículo de revisión publicado en 2025 resume esta tendencia. Este artículo lo explica dividiéndolo en química de la combustión, flujos reactivos y diversos otros problemas. Sostiene que incorporar el conocimiento físico al aprendizaje permite obtener respuestas físicamente consistentes incluso con pocos datos. Otro artículo de revisión de 2026 se centra en la combustión turbulenta de la propulsión aeroespacial. Explica que une el rigor de la física con el poder de los datos.

La ventaja de este método radica en el uso conjunto de datos y física. Con pocos sensores, resulta difícil llenar los vacíos únicamente con datos. Las ecuaciones físicas reducen esos vacíos. Así, es posible describir el estado completo incluso con unos pocos sensores económicos. El hecho de que pueda ejecutarse con rapidez para fines de monitorización también reviste una gran utilidad.

Las limitaciones también son claras. Las reacciones de llama y la turbulencia son sumamente complejas. La ecuación de Helmholtz es un modelo reducido que simplifica el sonido. No puede captar todos los detalles de la combustión real. Por ello, este gemelo digital se utiliza para una monitorización rápida. Los juicios precisos deben complementarse con análisis de alta fidelidad y pruebas reales. Cifras detalladas como frecuencias específicas o amplitudes de presión no han sido confirmadas en artículos publicados. Por esa razón, este libro no incluye tales valores. En su lugar, solo explica los principios de cómo funciona este método y en qué aspectos resulta útil.

Para que un gemelo digital funcione bien, debe calibrarse continuamente con la realidad. Si el modelo se desvía poco a poco de la realidad, la representación se distorsiona. Por eso, el modelo se corrige constantemente con los valores de los sensores. Si esta corrección se retrasa, el gemelo pierde el rastro

de la realidad. Si la corrección es demasiado frecuente, sigue incluso el ruido y se vuelve inestable. Encontrar el equilibrio entre ambos es la clave técnica. Un buen gemelo no pierde de vista la realidad sin dejarse arrastrar por el ruido.

La monitorización no es la única forma de controlar las oscilaciones de combustión. También se incorporan en las paredes estructuras que absorben el sonido. A estas estructuras se las denomina dispositivos de absorción acústica. También se interrumpe la retroalimentación cambiando la disposición de los inyectores. La monitorización mediante inteligencia artificial es una herramienta que asiste a dicho diseño. Informa de antemano bajo qué condiciones crecen las oscilaciones. El diseño y la monitorización deben ir de la mano para dominar las oscilaciones.

10.4 Cómo medir el envejecimiento y las diferencias individuales de los motores

Esta sección aborda el problema del envejecimiento de los motores y las variaciones entre unidades individuales. Primero, se explica por qué los motores reutilizables cambian poco a poco. A continuación, se examina el método para medir la incertidumbre de dichos cambios. Este método es una red neuronal que proporciona no solo una respuesta, sino también el rango de esa respuesta. Por último, se destaca por qué esta incertidumbre es crucial para la seguridad. Conviene aclarar algo de antemano. Los estudios citados aquí corresponden a la predicción de vida útil en maquinaria general. No han sido verificados con datos de motores cohete, y su aplicación a cohetes representa una posibilidad futura.

Los motores reutilizables utilizan el mismo motor múltiples veces. Es comparable a pasar de desechar un vaso de papel tras un solo uso a utilizar una taza de cerámica muchas veces. Este enfoque reduce sustancialmente el coste del transporte espacial. Sin embargo, cada vez que se usa, las piezas se desgastan poco a poco. La superficie del inyector se desgasta por abrasión. Los álabes de la turbina sufren erosión. Los canales de refrigeración se deforman por fatiga térmica. A este fenómeno en el que el rendimiento cambia con el paso del tiempo se le denomina degradación por envejecimiento.

Otra variable es la variación de fabricación. Aunque se fabriquen con el mismo plano, las dimensiones reales varían ligeramente. Existe un margen de error diminuto en cada pieza. Debido a este margen de error, la respuesta varía de un motor a otro. Cuando el envejecimiento y las variaciones de fabricación se superponen, surge la individualidad de cada motor. Dado que el resultado difiere levemente entre motores aun con la misma orden a las válvulas, el controlador debe ser consciente de esta diferencia.

Este cambio afecta directamente al rendimiento del motor. El desgaste y la erosión reducen paulatinamente el empuje. Si el empuje disminuye, la trayectoria del cohete se altera. También cambia la eficiencia con la que se lleva a cabo la combustión. La velocidad característica de escape es un valor que refleja esta eficiencia de combustión. Si este valor cae, se genera menos fuerza con la misma cantidad de propelente. Por lo tanto, es importante medir con antelación el envejecimiento y las variaciones.

Conviene explicar un poco el concepto de velocidad característica de escape. Este valor es un indicador de cuán eficientemente combustiona la cámara de combustión. Se determina a partir de la presión de la cámara de combustión, el área de la garganta de la tobera y el caudal de propelente. Es un valor distinto de la velocidad de escape real expulsada por la tobera o de la eficiencia de la tobera. Es un criterio que mide únicamente el grado de eficiencia de la combustión. Si este valor es alto, la combustión es eficiente. Si es bajo, el propelente se quema de forma menos eficiente. Cuando los inyectores se desgastan, la mezcla empeora y este valor desciende. Por eso, a través de los cambios en este valor, es posible estimar la salud del motor.

Este cambio se puede comparar con unas zapatillas deportivas muy usadas. Las zapatillas nuevas tienen una suela firme. Con el uso prolongado, la suela se desgasta y resbala. Aunque se corra con la misma fuerza, se avanza menos. Del mismo modo, el rendimiento del motor decae poco a poco a medida que se usa. El problema radica en que esa degradación difiere en cada motor. Algunos motores se desgastan con rapidez y otros lo hacen lentamente. Por eso, es necesario examinar el estado de cada motor por separado.

Conocer el envejecimiento es la clave para una reutilización segura. Si se desconoce el estado, se pueden cometer dos errores. Se puede desechar demasiado pronto un motor que todavía es utilizable. Esto supone un desperdicio. Por el contrario, se puede forzar el uso de un motor que ya está exhausto. Esto conduce a accidentes. Si se conoce el estado con precisión, se utiliza el motor todo lo posible sin rebasar la línea de peligro, evitando ambos errores al mismo tiempo.

Aquí es donde se utilizan las redes neuronales bayesianas variacionales. En inglés se denomina Variational Bayesian Neural Network. En adelante, nos referiremos a ellas como redes neuronales bayesianas. Una red neuronal ordinaria solo produce una única respuesta. Por ejemplo, responde con una cifra única diciendo que la pérdida de empuje es del 3 por ciento. La red neuronal bayesiana proporciona también el rango de la respuesta. Indica conjuntamente la respuesta y el grado de certeza, por ejemplo, señalando que existe una alta probabilidad de que la reducción del empuje se sitúe entre el 2 y el 4 por ciento.

Ilustremos de forma sencilla el modo en que una red neuronal bayesiana genera un rango. Supongamos que, en lugar de preguntar a una sola persona, se pregunta a varias. Si las distintas respuestas son similares, la certeza es alta. Si las respuestas difieren mucho entre sí, la certeza es baja. La red neuronal bayesiana recoge estas diversas opiniones dentro de un único modelo. Por ello, entrega tanto el valor central de la respuesta como la dispersión de su rango. Si el rango es estrecho, significa que el modelo está seguro; si es amplio, significa que aún no lo sabe con claridad.

Existen dos tipos de incertidumbre. Una proviene de las fluctuaciones inherentes a los propios datos. Es una fluctuación imposible de eliminar, como el ruido de los sensores. La otra se debe a que el modelo no ha aprendido lo suficiente. Aumenta ante situaciones que nunca antes ha experimentado. Cuando se enfrenta por primera vez a un nuevo tipo de degradación, esta incertidumbre se amplía. Distinguir entre ambos tipos facilita la toma de medidas. Permite determinar si conviene recopilar más datos o realizar más pruebas.

Las redes neuronales bayesianas son más difíciles de entrenar que las redes neuronales ordinarias. Esto se debe a que deben aprender una distribución de respuestas y no un único valor. Calcular con exactitud esta distribución requiere una enorme cantidad de cómputo. Por ello, los investigadores emplean métodos de aproximación. Transforman la distribución en una forma más manejable para aprenderla. A este método se le denomina inferencia variacional. El término variacional en el nombre proviene de esta aproximación. Al recurrir a una aproximación, es imprescindible comprobar los resultados con datos reales.

La predicción del envejecimiento no es un problema exclusivo de los cohetes. Los motores de aviación, los aerogeneradores y la maquinaria industrial comparten la misma preocupación. Esto se debe a que todas son máquinas destinadas a un uso prolongado. Por ello, las investigaciones en estos campos se nutren mutuamente. Los métodos que han demostrado su eficacia en otras máquinas se trasladan a los cohetes. Las exigentes condiciones de los cohetes también dan origen a nuevos métodos. Este intercambio de métodos entre disciplinas acelera el desarrollo de la tecnología.

A este grado de certeza se le llama incertidumbre. Conocer la incertidumbre es fundamental para la seguridad. Es necesario actuar de forma diferente cuando el modelo responde con seguridad y cuando

duda. Si el modelo indica que no sabe con certeza, el ser humano toma mayores precauciones. Se establece un margen de seguridad más amplio. Se verifica mediante pruebas adicionales. Funciona con la misma lógica que el pronóstico del tiempo al anunciar la probabilidad de lluvia. Si la probabilidad es del cincuenta por ciento, se lleva un paraguas y se observa la situación.

Esta dirección se expande con rapidez en múltiples campos. La investigación sobre la predicción de la vida útil restante de los componentes es un ejemplo representativo. A la vida útil que queda se la denomina vida útil remanente. Un estudio de 2024 predice esta vida útil mediante redes neuronales bayesianas. Entrega la respuesta como una distribución de probabilidad en lugar de un único valor. Un estudio de 2025 aborda múltiples incertidumbres conjuntamente mediante redes neuronales recurrentes bayesianas. Estos estudios no están dedicados exclusivamente a los cohetes. Sin embargo, el problema de gestionar el estado de máquinas que envejecen es el mismo. El mismo concepto puede aplicarse a la predicción del envejecimiento de los motores cohete.

Los datos de motores aeronáuticos constituyen un objetivo en el que este método ha sido ampliamente probado. Los investigadores comparan y ponen a prueba sus métodos con datos públicos sobre la degradación de motores. Comparan las predicciones de diversos métodos utilizando el mismo conjunto de datos. De este modo, es posible determinar qué método es superior. Aunque los datos de cohetes son escasos, se consolida la base con estos datos públicos. Posteriormente, el método consolidado se traslada y verifica con datos de cohetes.

El controlador recibe esta información sobre la incertidumbre y toma decisiones. Si la incertidumbre es alta, opera de manera conservadora. Ajusta el empuje a un nivel ligeramente menor o amplía los márgenes de seguridad. Si la incertidumbre es baja, extrae un mayor rendimiento. De esta manera, es posible operar adaptándose al estado del motor. Es más seguro que el método de gestionar todos los motores con una única regla fija.

La información sobre la incertidumbre también se utiliza en las decisiones de mantenimiento. Si el modelo predice el envejecimiento con seguridad, resulta fácil planificar. Si el modelo muestra incertidumbre, se desmonta el componente para inspeccionarlo directamente. Así se reducen los mantenimientos a ciegas. Al realizar el mantenimiento conociendo el estado real, se ahorran componentes. Al mismo tiempo, no se pasan por alto las piezas peligrosas. Conocer la incertidumbre protege simultáneamente el ahorro y la seguridad.

También existen limitaciones. El hecho de que la inteligencia artificial proporcione probabilidades no significa que estas sean correctas por sí solas. Si el modelo aprende de forma errónea, transmitirá una falsa certeza. Por ello, es imprescindible calibrar y validar con datos de pruebas reales. Se trata de comprobar si la incertidumbre estimada por el modelo se corresponde con el error real. Las probabilidades que no hayan pasado por esta verificación carecen de fiabilidad. En entornos de alto riesgo como los cohetes, esta verificación debe ser aún más rigurosa. Cifras como los intervalos de confianza detallados no han sido confirmadas en artículos publicados. Por esa razón, este libro no incluye tales valores. En su lugar, solo explica el marco conceptual para medir la incertidumbre.

10.5 Control por aprendizaje por refuerzo para la regulación coordinada de múltiples válvulas (LUMEN, DLR P8.3)

Esta sección aborda el control inteligente que regula de manera coordinada múltiples válvulas del motor. En primer lugar, se explica en qué consiste el aprendizaje por refuerzo. A continuación, se analiza por qué el control de los motores cohete es complejo. Seguidamente, se presenta un caso experimental real llevado a cabo en Alemania. Por último, se señalan los desafíos pendientes antes de poder emplear esta tecnología en vuelo.

El aprendizaje por refuerzo es un método que aprende mediante ensayo y error. Es similar a la forma en que un bebé aprende a caminar. Se puede concebir como recibir un castigo si se cae y una recompensa si camina bien. Se premian las buenas acciones y se penalizan las malas. Tras múltiples intentos, se descubre la manera de maximizar las recompensas. Al método que añade redes neuronales a este proceso se le denomina aprendizaje por refuerzo profundo (Deep Reinforcement Learning, DRL). En adelante, nos referiremos a él como aprendizaje por refuerzo profundo.

En el aprendizaje por refuerzo chocan dos intenciones. Una es el deseo de utilizar un método bueno ya conocido. La otra es el deseo de probar un método nuevo. Si solo se usa lo que ya se conoce, no se puede encontrar un camino mejor. Si solo se prueba lo nuevo, es peligroso y uno se pierde. Por lo tanto, se necesita un equilibrio entre ambas. Este equilibrio puede verse como un tira y afloja entre la intención de explorar lo nuevo y la de explotar lo conocido. En las etapas iniciales del aprendizaje se experimenta mucho y, a medida que se avanza, se utiliza lo que ya se conoce.

En los cohetes, este proceso de prueba y error no puede realizarse con un motor real. Esto se debe a que no es posible aprender destruyendo motores. Por eso, primero se aprende con un modelo informático. A este modelo se le llama simulador. En el simulador no importa fallar miles de veces. Se puede probar libremente para encontrar un buen control. Es similar a jugar varias partidas de un juego para dominar los trucos.

Veamos por qué es difícil controlar el motor de un cohete. Los valores dentro del motor están estrechamente interconectados. Al abrir la válvula del comburente, la presión aumenta. Cuando la presión sube, la rotación de la turbobomba cambia. La proporción entre combustible y comburente también varía al mismo tiempo. A esta proporción se le llama relación de mezcla. Si se altera un valor, otros valores cambian sucesivamente en cadena. Por eso no es posible regular una sola válvula de forma aislada. Se necesita un control que gestione múltiples valores simultáneamente.

En los cohetes que aterrizan, este control es aún más complejo. Para que el cohete descienda de nuevo, el empuje debe reducirse drásticamente. A veces se reduce por debajo de un tercio del empuje máximo. Esta regulación amplia del empuje se denomina estrangulamiento profundo (deep throttling). Significa estrangular profundamente el flujo. Incluso en ese momento, la relación de mezcla debe mantenerse adecuada. Si la relación de mezcla se desajusta, la combustión empeora y los componentes se dañan. Es necesario ajustar múltiples valores al mismo tiempo, con rapidez y de forma segura.

El Centro Aeroespacial Alemán investiga este problema con motores reales. El Centro Aeroespacial Alemán se abrevia como DLR. El DLR construyó un motor experimental llamado LUMEN. LUMEN es un motor alimentado por bomba que utiliza oxígeno líquido y metano. Su empuje es del orden de 25 kilonewtons. Es un tamaño que corresponde al de un motor cohete pequeño. Este motor operó por primera vez en marzo de 2024 en el banco de pruebas P8.3. El P8.3 es una instalación de pruebas ubicada en el centro del DLR en Lampoldshausen. LUMEN sirve como plataforma para probar tecnologías de motores reutilizables. Allí se investigan conjuntamente la monitorización del estado y el control inteligente.

LUMEN utiliza un ciclo específico. Este ciclo se denomina expansor con purga (expander bleed). Es un método en el que una parte del combustible calentado en los canales de refrigeración se utiliza para accionar la turbina. Este método tiene una estructura relativamente simple y segura. Por ello, resulta adecuado para probar diversas tecnologías. LUMEN es un banco de pruebas abierto que comparten la industria y los centros de investigación. Permite aplicar y probar nuevos métodos de control en un motor real.

El banco de pruebas P8.3 comenzó a operar en 2021. Este banco de pruebas se construyó ampliando una instalación anterior. Antiguamente solo se podían probar cámaras de combustión. El P8.3 permite

probar el motor completo e incluso las turbobombas. De este modo, los métodos de control pueden verificarse en condiciones reales. LUMEN y el P8.3 constituyen la base de la investigación sobre el control mediante inteligencia artificial.

El DLR aplica el control por inteligencia artificial a este motor. Los documentos oficiales del DLR explican que se modela la transferencia de calor en los canales de refrigeración mediante redes neuronales. Señalan que se ajusta el empuje y se monitoriza la inestabilidad de la combustión mediante control por inteligencia artificial. La tesis doctoral de Kai Dresia, publicada en 2025, aborda esta investigación en detalle. Esta tesis controla el motor cohete mediante aprendizaje por refuerzo profundo. Compara sus resultados con el control predictivo basado en modelos convencional. Informa haber obtenido un error medio de aproximadamente el 1,3 por ciento al controlar múltiples valores de forma simultánea. Este valor es el resultado presentado en la tesis de Dresia. Proviene de estudios de simulación y de bancos de pruebas, y no constituye un valor demostrado en vuelo para todo el rango operativo de un motor completo.

El DLR aplicó el control por redes neuronales a pruebas reales en investigaciones previas. En 2023, verificó el control por redes neuronales en las pruebas de la turbobomba de LUMEN. También llevó a cabo investigaciones para modelar la transferencia de calor en los canales de refrigeración mediante redes neuronales. En 2024, presentó los resultados de las pruebas de componentes y el estado de avance de LUMEN. Estas investigaciones acumuladas conducen al control de motores reales. No se trata de un éxito puntual, sino de una acumulación a lo largo de varios años.

Veamos con más detalle el marco en el que el controlador gestiona estados y acciones. El controlador lee el estado en cada instante y produce una acción. Dicha acción genera el estado siguiente. En el estado siguiente, vuelve a producir una acción. Este proceso continúa de forma ininterrumpida. A un problema en el que estados y acciones se suceden de este modo se le denomina problema de decisión secuencial. El aprendizaje por refuerzo es idóneo para abordar estas decisiones secuenciales. No aprende un juicio aislado, sino un flujo continuo de decisiones.

Veamos el principio de cómo razona un controlador de aprendizaje por refuerzo. El controlador lee primero el estado del motor. Son valores como la presión de la cámara de combustión, la posición de las válvulas, la rotación de la bomba y la temperatura. A estos valores se les llama estado. Al observar el estado, el controlador decide cuánto mover las válvulas. A este movimiento se le denomina acción. Si el resultado de la acción es bueno, recibe una recompensa; si es malo, recibe una penalización. Si se acerca al empuje objetivo, se otorga una recompensa. Si la relación de mezcla se desvía o las vibraciones aumentan, se aplica una penalización. A la regla que contiene estas recompensas y penalizaciones se le llama función de recompensa. El controlador perfecciona la manipulación de las válvulas para maximizar las recompensas recibidas.

Diseñar adecuadamente la función de recompensa determina el éxito o el fracaso del control. Las reglas de recompensa y penalización deben reflejar fielmente los objetivos. Si solo se recompensa el empuje, el controlador puede ignorar las vibraciones. Por ello, el empuje, la relación de mezcla y las vibraciones se incluyen conjuntamente en las recompensas y penalizaciones. También se penaliza mover las válvulas de forma excesivamente brusca, ya que las maniobras bruscas sobrecargan los componentes. Es necesario calibrar bien el peso de los diversos objetivos. Este ajuste de pesos es la tarea central del diseño de control.

La operación automática de motores reutilizables también es un tema central en la industria. El motor más reciente de una empresa estadounidense está orientado a la reutilización repetida. Se sabe que este motor redujo el número de componentes y se modificó para facilitar su fabricación. Se informa que eliminó los escudos térmicos para reducir el peso y la complejidad. Estos cambios responden a la

tendencia de utilizar los motores con frecuencia y realizar su mantenimiento con rapidez. No obstante, se trata de un caso de desarrollo empresarial y no de valores de rendimiento validados académicamente. Este libro cita tales ejemplos únicamente como referencia de tendencias.

La ventaja de este método radica en que aborda el problema interconectado en su totalidad. Para un ser humano es difícil diseñar manualmente reglas para cada válvula. Dado que los valores están entrelazados, las reglas se vuelven complejas. El aprendizaje por refuerzo aprende estas relaciones complejas por sí mismo a partir de los datos. Examina múltiples valores conjuntamente y encuentra un punto de equilibrio. Sin embargo, un controlador entrenado no se adapta automáticamente a un envejecimiento o deterioro que no ha experimentado. Existen varias formas de responder a los cambios en el motor. Existe el método de leer primero el estado para reflejarlo en el control. Existe el control robusto, que evita oscilaciones ante variaciones. También existe el método de reentrenar el controlador con nuevos datos. La adaptación en línea durante la operación genera el nuevo riesgo de cambios de control no verificados. Por ello, la vía a utilizar se decide evaluando la seguridad.

El control por aprendizaje por refuerzo puede compararse con los métodos tradicionales. Un método ampliamente utilizado durante mucho tiempo es el control predictivo basado en modelos. Este método anticipa el futuro mediante un modelo matemático del motor y determina la posición de las válvulas. Funciona bien si se dispone de un modelo preciso. Sin embargo, si el motor envejece o el modelo no coincide exactamente, el rendimiento disminuye. Dado que el aprendizaje por refuerzo aprende directamente de los datos, es flexible ante tales cambios. Ambos métodos tienen ventajas y desventajas recíprocas. Por ello, los investigadores los comparan y los combinan.

Veamos por qué es importante mantener una relación de mezcla adecuada. La relación de mezcla es precisamente la proporción entre combustible y comburente. La temperatura de la llama suele alcanzar su punto máximo cerca de la proporción estequiométrica entre combustible y comburente. Si se desvía de esta proporción, la temperatura disminuye, ya sea que sobre combustible o comburente. Con frecuencia, los cohetes operan deliberadamente con un ligero exceso de combustible. Esto se hace para reducir la temperatura de la llama en contacto con las paredes y protegerlas. Sin embargo, si se añade demasiado combustible, la fuerza disminuye. En los combustibles derivados del queroseno, en este caso se acumula hollín. Los motores que utilizan hidrógeno casi no producen hollín. De este modo, la relación de mezcla adecuada varía según la combinación de propelentes. Por lo tanto, la relación de mezcla debe mantenerse incluso mientras se varía el empuje. Esta es la razón por la que se necesita un control que ajuste múltiples valores al mismo tiempo.

Aquí las cifras deben tratarse con cautela. Valores detallados de rendimiento como tiempos rápidos de estrangulamiento profundo, rangos de error de la relación de mezcla o tasas de reducción de vibraciones no han sido confirmados en publicaciones abiertas. El uso de valores de rendimiento no verificados distorsionaría los hechos exagerándolos. Por eso este libro no utiliza tales cifras. El valor confirmado es un error medio de aproximadamente el 1,3 por ciento en el control multivariable. Solo este valor se toma como fundamento.

Conviene señalar lo que significa este error medio del 1,3 por ciento. Este valor es el resultado de ajustar múltiples variables simultáneamente. Significa que se cumplieron varios objetivos juntos y no solo uno. Con reglas diseñadas manualmente por humanos resulta difícil alcanzar este equilibrio. La inteligencia artificial aprendió las relaciones complejas y encontró el equilibrio. No obstante, este valor es un resultado obtenido bajo las condiciones limitadas de un estudio específico. No es un valor que pueda extrapolarse directamente a todos los motores. Si las condiciones cambian, debe verificarse nuevamente.

Aún queda camino por recorrer antes de aplicar esta tecnología directamente a motores de vuelo. El controlador de aprendizaje por refuerzo aprende primero en simulaciones. Dentro del ordenador, equivocarse no implica peligro. A continuación, se realizan pruebas conectando el hardware. Esta prueba se denomina prueba de hardware en el bucle. Se ejecutan conjuntamente los componentes reales y la red neuronal. Posteriormente, se comprueba la seguridad mediante pruebas de combustión en tierra. Solo después de superar todas estas etapas se puede plantear el vuelo, ya que la certificación de cohetes es sumamente rigurosa.

Una gran dificultad radica en la discrepancia entre el simulador y la realidad. El modelo computacional no es idéntico al motor real. Un control bien aprendido en el modelo puede fallar en la realidad. A esta diferencia se le llama brecha entre la simulación y la realidad. Para reducir esta brecha, es necesario hacer el modelo más preciso. El modelo debe ajustarse continuamente con datos de pruebas reales. Bancos de pruebas como LUMEN se utilizan para estos ajustes. Solo con datos reales se puede confiar en la inteligencia artificial.

Otro principio consiste en disponer dispositivos de seguridad por capas. Existe la posibilidad de que el controlador de aprendizaje por refuerzo emita una orden anómala. Para esos casos, se incorpora conjuntamente un dispositivo de seguridad basado en reglas. Si la orden del controlador supera una línea de peligro, el dispositivo de seguridad la bloquea. La inteligencia artificial eleva el rendimiento y las reglas garantizan el límite inferior de seguridad. Solo con estas dos capas es posible aplicar la inteligencia artificial a un motor real.

Las válvulas también tienen límites físicos. Una válvula no puede abrirse ni cerrarse más rápido que una velocidad determinada. A esta velocidad máxima se le llama límite de velocidad de accionamiento de la válvula. Si la inteligencia artificial emite una orden brusca que supera este límite, las tuberías sufren un impacto. El golpe de ariete, en el que el líquido se detiene repentinamente y golpea la tubería, es un ejemplo de ello. Por tanto, las órdenes del controlador se restringen dentro de la velocidad que la válvula puede alcanzar. Además, se instala un mecanismo de conmutación de seguridad que transfiere el control inmediatamente a un controlador clásico si la inteligencia artificial se desestabiliza. El controlador clásico protege el motor de forma segura mediante reglas largamente verificadas.

El valor del control por aprendizaje por refuerzo se hace patente en la reutilización. Cuando un motor se utiliza varias veces, su estado varía en cada ocasión. Las reglas de control fijas no pueden seguir estos cambios. El controlador de aprendizaje por refuerzo lee el estado y ajusta las válvulas en consecuencia. Trata a un motor desgastado según su desgaste y a un motor nuevo según su novedad. Esta flexibilidad reduce los costes de los cohetes reutilizables. Por ello, diversas instituciones de todo el mundo concentran esfuerzos en esta investigación.

Sin embargo, el control por inteligencia artificial se encuentra todavía en la fase de pruebas en tierra. Aún no se utiliza de forma generalizada en motores de vuelo reales. La certificación de vuelo exige un largo historial de pruebas y registros. Es necesario demostrar que el nuevo control es seguro en cualquier circunstancia. Esta demostración requiere una gran cantidad de datos y tiempo. Por ello, el control por inteligencia artificial se desarrolla actualmente en laboratorios y bancos de pruebas. A medida que se acumulen estos resultados, algún día darán el salto al vuelo.

10.6 Conclusión y perspectivas futuras

Esta sección examina cómo se integran las cuatro tecnologías anteriores en un solo conjunto. A continuación, resume por qué esta integración es necesaria en la era de los cohetes reutilizables. Por último, señala los principios de seguridad necesarios para la gestión de motores mediante inteligencia artificial.

La operación inteligente de un motor cohete de propelente líquido conecta cuatro tecnologías. El diagnóstico por sensores examina múltiples señales conjuntamente para detectar indicios sutiles a tiempo y hallar anomalías tempranas. El gemelo digital reconstruye el flujo de presión interno no visible mediante una red neuronal que respeta las ecuaciones de la física. El modelo bayesiano cuantifica la incertidumbre del envejecimiento y las desviaciones, proporcionando tanto la respuesta como su margen de variación. El controlador de aprendizaje por refuerzo regula múltiples válvulas simultáneamente y encuentra el punto de equilibrio entre los valores interconectados.

Estas cuatro tecnologías intercambian datos entre sí. Cuando el diagnóstico detecta una anomalía, el controlador responde. Si el gemelo digital informa sobre el estado interno, el diagnóstico se vuelve más preciso. Si el modelo bayesiano comunica la incertidumbre, el controlador actúa con mayor cautela. De esta manera, las cuatro tecnologías giran en un único bucle. Es una estructura en la que el motor vigila su propio estado y se autorregula. A esta estructura se le denomina gestión autónoma de la salud.

Estas cuatro tecnologías compensan mutuamente sus debilidades. Es difícil conocer la causa exacta de una anomalía únicamente mediante el diagnóstico. Cuando el gemelo digital muestra el interior, la causa se aclara. El controlador por sí solo no conoce el envejecimiento del motor. Cuando el modelo bayesiano informa sobre la incertidumbre, el control se vuelve seguro. La brecha de una tecnología la llena otra. Por eso, no se manejan por separado, sino que se integran en una sola. Solo al unirlas se completa la gestión autónoma.

Esta corriente aún no es un producto terminado. La mayor parte se encuentra en fase de investigación y pruebas. Varios resultados son también investigaciones iniciales que no han pasado por una revisión por pares. Por ello, es prematuro dar por sentados los logros. Este libro deja claro ese punto. Ahora es la etapa de comprender la estructura y la dirección de la tecnología. Se necesita más tiempo para que se utilice ampliamente en motores de vuelo reales.

Esta corriente es indispensable en la era de los cohetes reutilizables. Si un motor se utiliza varias veces, su estado cambia cada vez. Es difícil abordar todos los cambios solo con reglas fijas. Los valores cambian con demasiada rapidez y están tan entrelazados que una persona no puede ajustarlos manualmente en cada ocasión. La inteligencia artificial es una herramienta para leer y responder rápidamente a estos cambios. Para reducir el costo del transporte espacial, esta gestión autónoma debe estar presente.

Sin embargo, en el control mediante inteligencia artificial deben existir necesariamente dispositivos de restricción. Uno de ellos es la explicabilidad: las personas deben poder entender por qué el modelo tomó esa decisión. Otro es la verificación independiente: el juicio del modelo debe confirmarse nuevamente mediante otro método. El restante es la transición a un estado seguro en caso de fallo: incluso si el modelo falla, el motor debe poder detenerse de forma segura. Sin estos tres mecanismos, resulta difícil confiar el motor a la inteligencia artificial.

La inteligencia artificial no es una herramienta para confiar a ciegas en el motor en nuestro lugar. La inteligencia artificial es una herramienta para observar el motor con más detalle. Los humanos comprenden el motor más a fondo valiéndose de los ojos de la inteligencia artificial. El juicio final y la responsabilidad siguen recayendo en las personas. Sobre este principio, la inteligencia artificial hace que los cohetes sean más seguros y económicos.

Resumen

Repasemos los puntos clave de este capítulo. El motor cohete de propelente líquido es una máquina extremadamente rápida, caliente y peligrosa. Para utilizar este motor de forma segura, se requieren una monitorización y un control en tiempo real. El método tradicional de valores límite suele pasar por alto anomalías leves. La inteligencia artificial examina varios sensores a la vez para subsanar esta deficiencia. El objetivo es detectar señales sutiles antes de que los valores crucen la línea de peligro.

El diagnóstico temprano analiza conjuntamente el flujo temporal y las relaciones entre canales. Una estructura que combina convolución temporal y atención es representativa. Esta estructura detecta anomalías de forma temprana en múltiples sensores. La decisión debe completarse en milisegundos en un pequeño chip junto al motor. Sin embargo, las cifras de rendimiento específicas para cohetes aún son difíciles de confirmar. Por eso, este libro las explica únicamente a través de sus principios. El diagnóstico se utiliza junto con la supervisión humana y dispositivos de seguridad basados en reglas.

El gemelo digital dibuja el interior invisible de la cámara de combustión. Utiliza redes neuronales físicas que respetan las ecuaciones de la física. A partir de unos pocos sensores en la pared, reconstruye todo el flujo de presión. Es de gran utilidad para monitorizar la inestabilidad de la combustión. Sin embargo, no puede capturar toda la complejidad de las llamas. Se emplea para una monitorización rápida y se complementa con análisis precisos. El gemelo digital cumple su función solo si se calibra continuamente con señales reales.

La red neuronal bayesiana mide la incertidumbre del envejecimiento y de las desviaciones. Proporciona una respuesta junto con su margen de incertidumbre. Si el margen es estrecho, significa que tiene alta confianza; si es amplio, significa que no lo sabe con certeza. Esta información se utiliza para que el controlador decida si debe actuar con cautela. También se usa para determinar cuándo realizar el mantenimiento. No obstante, las probabilidades del modelo deben verificarse con datos reales.

El controlador de aprendizaje por refuerzo regula varias válvulas simultáneamente. El DLR alemán investiga esto en el motor LUMEN y en el banco de pruebas P8.3. Los datos públicos informan un error medio de aproximadamente el 1,3 por ciento en el control de múltiples variables. Algunas cifras detalladas de rendimiento no se incluyen debido a la falta de confirmación. Para utilizar esta tecnología en vuelo, aún quedan varias etapas de verificación.

Repasemos nuevamente un principio que atraviesa este capítulo. La inteligencia artificial es una herramienta para mejorar el rendimiento. Sin embargo, la seguridad la protegen las reglas y los seres humanos. Ambas capas deben coexistir para aplicar la inteligencia artificial a un motor real. No confiamos en cifras no verificadas. Distinguimos entre la investigación inicial y los logros comprobados al leer. Esta actitud es el camino correcto para tratar con tecnologías peligrosas como los cohetes.

Cuando estas cuatro tecnologías se integran en una sola, el motor cuida de su propia salud. Las cuatro tecnologías intercambian datos entre sí y funcionan en un único circuito. Esta gestión autónoma es la base de la era de los cohetes reutilizables, ya que si un motor se utiliza varias veces, su estado cambia en cada ocasión. Sin embargo, para garantizar la seguridad, deben estar presentes la explicabilidad, la verificación independiente y los mecanismos de transición a un estado seguro.

Las tecnologías vistas en este capítulo responden a diferentes preguntas. El diagnóstico pregunta qué está fallando en este momento. El gemelo digital pregunta cómo es el interior que no se ve. El modelo bayesiano pregunta cuánto se puede confiar en este juicio. El controlador pregunta qué se debe hacer entonces. Al reunirse estas cuatro preguntas, el estado del motor y la respuesta se conectan como una sola entidad. Una buena gestión de motores con inteligencia artificial aborda estas cuatro preguntas en conjunto.

Por último, se expone la postura adoptada ante los datos de este capítulo. Gran parte de las investigaciones recientes presentadas en este capítulo se encuentran en una fase inicial. Varios artículos son estudios preliminares que no han pasado por una revisión por pares. Los casos de motores de las empresas solo muestran las tendencias de la industria; su peso es diferente al de los logros comprobados académicamente. Por eso, este capítulo ha puesto el énfasis en los valores y principios confirmados, omitiendo las cifras que no pudieron comprobarse. Esperamos que el lector comprenda adecuadamente la dirección de la tecnología. La inteligencia artificial es una herramienta en el largo viaje para hacer que los cohetes sean más seguros y económicos.

Fuentes y referencias relacionadas

Yang, S., Wu, J., Xie, C., Zhong, Z., Cheng, Y., Wang, B., Liu, Y., & Lu, J. (2026). Data-Driven Intelligent Fault Diagnosis: A Review from the Perspectives of Three Practical Issues in Liquid Rocket Engines. International Journal of Aeronautical and Space Sciences, 27(4), 3438-3463. https://doi.org/10.1007/s42405-026-01131-9

Wang, T., Ding, L., & Yu, H. (2022). Research and Development of Fault Diagnosis Methods for Liquid Rocket Engines. Aerospace, 9(9), 481. https://doi.org/10.3390/aerospace9090481

Deng, L., Cheng, Y., & Shi, Y. (2022). Fault Detection and Diagnosis for Liquid Rocket Engines Based on Long Short-Term Memory and Generative Adversarial Networks. Aerospace, 9(8), 399. https://doi.org/10.3390/aerospace9080399

Urgolo, A., Pill, I., Waxenegger-Wilfing, G., & Freiberger, M. (2024). Property Learning-Based Fault Detection for Liquid Propellant Rocket Engine Control Systems. OASIcs, 125, 15:1-15:20. https://doi.org/10.4230/OASIcs.DX.2024.15

Waxenegger-Wilfing, G., Dresia, K., Deeken, J., & Oschwald, M. (2021). Machine Learning Methods for the Design and Operation of Liquid Rocket Engines: Research Activities at the DLR Institute of Space Propulsion. arXiv. https://arxiv.org/abs/2102.07109

NASA. (2009). Distributed Health Monitoring System for Reusable Liquid Rocket Engines. NASA Technical Reports Server. https://ntrs.nasa.gov/api/citations/20090040299/downloads/20090040299.pdf

Wang, C., Liu, Q., Qu, L., Chang, W., Wang, T., Duan, P., & Cao, Y. (2025). Self-attention-enhanced Temporal Convolutional Networks for Anomaly Detection in Time Series. Proceedings of the 2025 4th International Conference on Cryptography, Network Security and Communication Technology. https://doi.org/10.1145/3723890.3723911

Wang, D., Zhang, H., & Yu, Z. (2026). MST-MFGAT: An adaptive temporal convolution with feature graph attention network for time series anomaly detection. Neurocomputing, 664, 132134. https://doi.org/10.1016/j.neucom.2025.132134

Harrje, D. T., & Reardon, F. H. (1972). Liquid Propellant Rocket Combustion Instability. NASA SP-194. https://ntrs.nasa.gov/citations/19720026079

Culick, F. E. C. (2006). Unsteady Combustion in Liquid Rocket Engines: Dynamics, Instabilities, and Control. NATO RTO-EN-AVT-143.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Mariappan, S., Nath, K., & Karniadakis, G. E. (2024). Learning thermoacoustic interactions in combustors using a physics-informed neural network. Engineering Applications of Artificial Intelligence, 138, 109388. https://doi.org/10.1016/j.engappai.2024.109388

Li, A., Du, J., & Liu, Y. (2026). Domain-decomposition physics-informed neural network approach for the acoustic boundary impedance estimation of a combustion chamber model. Journal of Vibration and Control. https://doi.org/10.1177/10775463261433496

Wu, J., Wang, X., Zhang, G., Liu, J., Li, X., Zhang, Y., Zhang, H., Lyu, J., Wang, B., & Wu, Y. (2025). Physics-informed machine learning for combustion: A review. arXiv. https://arxiv.org/abs/2509.03347

Mousavi, M., Caldwell, C., Baltes, J., Aljasem, M., & Lee, B. J. (2025). Physics-Informed Neural Networks in Clean Combustion: A Pathway to Sustainable Aerospace Propulsion. arXiv:2509.08094. https://arxiv.org/abs/2509.08094

Selvam D., C., Devarajan, Y., Raja, T., Tiwari, A., Sunil Kumar, M., Sharma, K., Agrawal, A. P., Khandual, A., & Mehar, K. (2026). Physics-Informed machine learning for turbulent combustion in aerospace propulsion: bridging physical rigour and data intelligence. Engineering Applications of Computational Fluid Mechanics, 20, 2678066. https://doi.org/10.1080/19942060.2026.2678066

Ochella, S., Dinmohammadi, F., & Shafiee, M. (2024). Bayesian neural networks for uncertainty quantification in remaining useful life prediction of systems with sensor monitoring. Journal of Mechanical Engineering Science. https://doi.org/10.1177/16878132241239802

Chen, C., Wang, C., Guo, J., Cui, P., Zheng, J., & Liu, Z. (2025). Remaining useful life prediction considering multiple uncertainty information via Bayesian BiGRU-based method. Reliability Engineering & System Safety, 264, 111431. https://doi.org/10.1016/j.ress.2025.111431

Dresia, K. (2025). Rocket Engine Control with Deep Reinforcement Learning. DLR-Forschungsbericht DLR-FB-2025-16. https://doi.org/10.57676/bwbm-p528

Dresia, K., Adler, A., Saraf, A., Hahn, R., Kurudzija, E., Traudt, T., Deeken, J., & Waxenegger-Wilfing, G. (2023). Rocket Engine Control with Neural Networks: Experimental Results of the LUMEN Turbopump Test Campaign. AIAA SciTech 2023 Forum, AIAA 2023-0137. https://doi.org/10.2514/6.2023-0137

Börner, M., Traudt, T., Dresia, K., Armbruster, W., Suslov, D., Groll, C., Hahn, R. D., Saraf, A. M., Deeken, J., Hardi, J., Oschwald, M., & Schlechtriem, S. (2024). Component Test Results and Project Progress of the DLR Liquid Upper Stage Demonstrator Engine (LUMEN). AIAA SciTech 2024 Forum, AIAA 2024-1398. https://doi.org/10.2514/6.2024-1398

DLR. (2026). LUMEN: AI-assisted reusable rocket engines. German Aerospace Center. https://www.dlr.de/en/research-and-transfer/featured-topics/reusable-space-transportation/lumen

DLR. (2026). LUMEN, Liquid Upper Stage Demonstrator Engine. German Aerospace Center. https://www.dlr.de/en/research-and-transfer/research-infrastructure/large-scale-research-facilities/lumen-en

DLR Institute of Space Propulsion. (2026). Test cell P8.3. German Aerospace Center. https://www.dlr.de/en/ra/research-transfer/research-and-test-infrastructure/test-benches-for-space-propulsion/research-and-technology-test-bench-p8-1/test-cell-p8.3

Apoye este libro de la Biblioteca de IA

Si este libro le ha resultado útil, puede apoyar las próximas traducciones, las ediciones públicas y los proyectos abiertos de conocimiento sobre inteligencia artificial a través de PayPal.

PayPal

https://paypal.me/kimkjkj

Escanee el código QR o abra el enlace anterior.

Kim Kyung-jin, abogado