

Ấn bản Thư viện AI tiếng Việt

# Trí tuệ nhân tạo cho thiết kế vật liệu mới và kỹ thuật động cơ đẩy tên lửa

Thế năng AI, phòng thí  
nghiệm tự hành và học máy  
dựa trên vật lý (PIML)

Luật sư Kim Kyung-jin

# Trí tuệ nhân tạo cho thiết kế vật liệu mới và kỹ thuật động cơ đẩy tên lửa

Thế năng AI, phòng thí nghiệm tự hành và học máy dựa trên vật lý (PIML)

Luật sư Kim Kyung-jin

Bắt đầu từ thế năng học máy học chuyển động của từng nguyên tử, cuốn sách này được sắp xếp thành mười chương: mô hình sinh tạo xác định trước tính chất mong muốn rồi tạo ngược lại cấu trúc tinh thể; phòng thí nghiệm tự hành, nơi robot tự lặp lại thí nghiệm; hợp kim entropy cao và vật liệu composite gồm để bảo vệ những vị trí nóng của tên lửa; giám sát khuyết tật và bảo đảm tuổi thọ mỗi trong in 3D kim loại; học máy dựa trên vật lý, đưa các định luật vật lý vào hàm mất mát để tính toán quá trình cháy; thiết kế kênh làm mát, chẩn đoán thời gian thực động cơ tên lửa nhiên liệu lỏng và điều khiển bằng học tăng cường. Sách được viết không dùng công thức để phục vụ độc giả phổ thông yêu thích khoa học công nghệ, mỗi chương đều kèm tài liệu căn cứ đã được kiểm chứng, và đây là bản đã được rà soát, chỉnh sửa dựa trên phê bình theo kiểu giáo sư hướng dẫn của hai người phản biện trí tuệ nhân tạo.

## Kho lưu trữ tài liệu nghiên cứu công khai

Cuốn sách này được biên soạn và sản xuất bằng trí tuệ nhân tạo. Nội dung được thu thập và tổng hợp từ tài liệu bằng các ngôn ngữ chính trên thế giới. NotebookLM không giới hạn ngôn ngữ đặt câu hỏi, bất kể tài liệu được viết bằng ngôn ngữ nào. Độc giả có thể đặt câu hỏi bằng ngôn ngữ của mình trong chính kho lưu trữ tài liệu nghiên cứu đã được dùng để chuẩn bị cuốn sách này. Điều đó vẫn đúng ngay cả khi một số tài liệu được viết bằng tiếng Hàn, tiếng Anh hoặc ngôn ngữ khác. Nếu mở phần cài đặt trong NotebookLM và đặt ngôn ngữ đầu ra thành ngôn ngữ mong muốn, câu trả lời và kết quả được tạo sẽ xuất hiện bằng ngôn ngữ đó.

Kho lưu trữ NotebookLM công khai được dùng khi chuẩn bị cuốn sách này (AI for Materials Science and Rocket Engine Engineering, 129 tài liệu):

<https://notebook.google.com/notebook/ac39e3ce-77a9-4e78-9aaf-b52c45c2934d>

Bản gốc tiếng Hàn của cuốn sách này nằm trong AI Library của kimkj.com:

<https://kimkj.com/ai-library/?mod=document&uid=12453>

# Mục lục

Nhấn vào tiêu đề từng chương để chuyển đến phần nội dung tương ứng. Nếu muốn đọc từ đầu, hãy chọn Chương 1.

Chương 1. Nền tảng của tin học vật liệu và thế năng liên nguyên tử học máy (MLIP) · 4

Chương 2. Dự đoán cấu trúc tinh thể (CSP) và thiết kế ngược vật liệu dựa trên AI tạo sinh và mô hình khuếch tán (Diffusion Models) · 21

Chương 3. Phòng thí nghiệm tự hành và hệ thống đa tác nhân khoa học · 38

Chương 4. Dự đoán nhất quán vật lý và tối ưu hóa thành phần của hợp kim entropy cao (HEA) và siêu hợp kim · 55

Chương 5. Đánh giá kiểm soát vi cấu trúc và tuổi thọ chịu gia nhiệt của vật liệu composite dùng cho hàng không vũ trụ (CMC) và gốm siêu chịu nhiệt (UHTC) · 72

Chương 6. Mô hình hóa ứng suất nhiệt trong quy trình chế tạo bồi đắp kim loại (LPBF/DED) và nhận biết khuyết tật tại chỗ (In-situ) · 90

Chương 7. Hậu xử lý hợp kim hàng không vũ trụ in 3D và bảo đảm tuổi thọ mỗi chu kỳ cao · 107

Chương 8. Mô hình hóa dòng khí phản ứng và hóa học cháy bằng học máy dựa trên vật lý (PIML/PINNs) và học toán tử (FNO) · 124

Chương 9. Tối ưu hóa thiết kế ngược topo của kênh làm mát tái sinh, lỗ làm mát màng và inducer · 142

Chương 10. Chẩn đoán lỗi thời gian thực, bản sao số và điều khiển thích ứng thời gian thực cho lực đẩy dựa trên học tăng cường trong động cơ tên lửa nhiên liệu lỏng (LRE) · 157

# Chương 1. Nền tảng của tin học vật liệu và thế năng liên nguyên tử học máy (MLIP)

## Mở đầu

Mọi đồ vật chúng ta dùng đều được làm từ vật liệu. Thìa được làm bằng kim loại, còn cốc được làm bằng thủy tinh. Vòi phun của động cơ tên lửa cũng được làm bằng kim loại và gốm đặc biệt. Khi vật liệu thay đổi, tính chất của đồ vật cũng thay đổi.

Tính chất của vật liệu bắt đầu từ những hạt nhỏ không nhìn thấy bằng mắt. Những hạt đó là nguyên tử. Nguyên tử là đơn vị cơ bản tạo nên vật chất. Tùy theo các nguyên tử liên kết và chuyển động ra sao, kim loại có thể trở nên cứng hơn hoặc mềm hơn.

Người chế tạo tên lửa muốn biết trước chuyển động của các nguyên tử này. Bên trong động cơ tên lửa, nhiệt độ tăng đến hàng nghìn độ. Áp suất cũng rất cao. Cần biết trước kim loại có thể chịu được trong môi trường như vậy mà không nóng chảy hay không. Nếu chế tạo linh kiện thật rồi thử nghiệm, sẽ tốn rất nhiều tiền và thời gian. Vì vậy các nhà khoa học cho nguyên tử chuyển động trong máy tính. Đó là mô phỏng nguyên tử.

Muốn cho nguyên tử chuyển động bằng máy tính, cần có quy tắc. Vì phải biết từng nguyên tử nhận lực theo hướng nào và mạnh bao nhiêu. Trong một thời gian dài, cách tạo quy tắc chính xác và cách tạo quy tắc nhanh là hai con đường khác nhau. Cách chính xác thì chậm, còn cách nhanh thì thiếu chính xác. Con đường mới phá bỏ bức tường giữa hai cách này là thế năng liên nguyên tử học máy (Machine Learning Interatomic Potentials, MLIP). Từ đây trở đi, công nghệ này được gọi là MLIP.

Chương này giải thích từng bước MLIP là gì và đã phát triển như thế nào. Trước hết, chúng ta xem xét hai phương pháp lâu đời trong tính toán nguyên tử. Tiếp đó, chúng ta xem học máy đã nối hai phương pháp ấy ra sao. Chúng ta lần lượt đi từ phương pháp ban đầu đến phương pháp mới nhất. Cuối cùng, chương này chỉ ra cần cần trọng điều gì khi dùng công nghệ này trong nghiên cứu vật liệu tên lửa và vật liệu mới. Chương này không dùng công thức. Nội dung được giải thích bằng những lời có thể gọi ra hình ảnh.

Thuật ngữ tin học vật liệu cũng xuất hiện lần đầu ở đây. Thuật ngữ này có nghĩa là materials informatics. Đây là ngành học thu thập nhiều dữ liệu về vật liệu, phân tích bằng máy tính và tìm vật liệu mới. Trong lĩnh vực đó, MLIP là công cụ cốt lõi đảm nhiệm tính toán ở cấp độ nguyên tử.

Tôi nói trước vì sao câu chuyện này gắn với tên lửa. Phát triển tên lửa khó tách khỏi việc tìm vật liệu mới. Nếu có kim loại nhẹ hơn và chịu được nhiệt độ nóng hơn, hiệu năng tên lửa sẽ tăng. Nếu tạo từng vật liệu như vậy trong phòng thí nghiệm rồi thử nghiệm, sẽ mất nhiều năm. Nếu sàng lọc ứng viên trước bằng tính toán nguyên tử trong máy tính, thời gian đó sẽ giảm. MLIP là chìa khóa làm cho phép tính ấy vừa nhanh vừa chính xác. Sau khi đọc chương này, bạn sẽ nắm được khung lớn về cách vật liệu mới ra đời trong máy tính.

## 1. Mở đầu và bối cảnh: sự kết hợp giữa lý thuyết phiếm hàm mật độ (DFT) và học máy

### Nội dung của mục này

Muốn xử lý nguyên tử bằng máy tính, cần biết năng lượng và lực giữa các nguyên tử. Có hai cách chính để tìm các giá trị này. Một là cách chính xác, tính cả chuyển động của electron. Cách kia là cách nhanh, dùng công thức đã định sẵn. Mục này giải thích mỗi cách làm tốt điều gì và không làm được điều gì.

Sau đó, mục này chỉ ra vấn đề cũ giữa hai cách ấy. Cách chính xác thì quá chậm, còn cách nhanh thì thường sai. Để giải quyết vấn đề này, học máy đã xuất hiện. Mục này trước hết phác họa bức tranh lớn đó.

Một vài từ trong mục này sẽ tiếp tục được dùng ở phần sau. Đó là năng lượng, lực và bề mặt thế năng. Nếu nắm được ý nghĩa của chúng bây giờ, mục tiếp theo sẽ dễ hiểu hơn. Không có công thức khó. Chúng ta chỉ giải thích bằng lời có thể hình dung như tranh.

Hai giá trị cốt lõi trong tính toán xử lý nguyên tử là năng lượng và lực. Năng lượng cho biết một cách sắp xếp nguyên tử ổn định đến đâu. Năng lượng càng thấp, cách sắp xếp đó càng ở trạng thái dễ chịu. Lực cho biết mỗi nguyên tử bị đẩy theo hướng nào. Nếu biết lực, ta có thể tính nguyên tử sẽ di chuyển đến đâu ở khoảnh khắc tiếp theo.

Thực ra năng lượng và lực không phải là hai giá trị tách rời. Lực xuất phát từ việc năng lượng thay đổi dốc đến mức nào theo vị trí. Nếu ví năng lượng là độ cao thấp của mặt đất, thì lực là độ dốc của sườn. Quả bóng lăn xuống phía có sườn dốc hơn. Nguyên tử cũng bị đẩy về phía năng lượng thấp hơn. Vì vậy, nếu biết bản đồ năng lượng, ta lấy được lực từ độ dốc của bản đồ đó. Mối quan hệ này sẽ lại trở nên quan trọng ở phần sau.

Lý thuyết phiếm hàm mật độ (Density Functional Theory, DFT) là phương pháp tính phân bố electron để tìm năng lượng và lực. Từ đây trở đi, phương pháp này được gọi là DFT. Electron là hạt điện tích rất nhỏ quay quanh nguyên tử. Electron tụ lại ở đâu và bao nhiêu quyết định liên kết hóa học. DFT tính phân bố electron này theo nguyên lý cơ học lượng tử. Cơ học lượng tử là các định luật vật lý của thế giới rất nhỏ. Vì vậy DFT mô tả liên kết hóa học một cách chính xác.

Giá trị do DFT đưa ra cũng không phải là chân lý hoàn hảo. DFT xử lý lực đẩy và lực hút giữa các electron bằng công thức gần đúng. Kết quả thay đổi đôi chút tùy theo cách chọn công thức gần đúng này. Vì vậy giá trị DFT không hẳn là giá trị thật, mà là giá trị chuẩn thu được trong một thiết lập nhất định. Học máy học các giá trị chuẩn này. Khi phần sau gọi DFT là đáp án đúng, cũng cần hiểu theo nghĩa đó.

Vấn đề nằm ở tốc độ. Khi số nguyên tử tăng, khối lượng tính toán của DFT tăng rất nhanh. Nếu số nguyên tử tăng gấp đôi, thời gian tính toán tăng hơn nhiều so với gấp đôi. Ngay cả khi dùng siêu máy tính lớn, số nguyên tử thường xử lý được chỉ ở mức vài trăm. Thời gian có thể tính cũng chỉ dừng lại ở khoảnh khắc rất ngắn. Cách này không phù hợp với việc quan sát một khối lớn trong thời gian dài như vật liệu tên lửa.

Ở phía ngược lại là trường lực kinh nghiệm (Empirical Force Field). Phương pháp này tính lực giữa các nguyên tử bằng công thức đã định sẵn. Nó không tính từng electron. Thay vào đó, nó xác định lực chỉ bằng vài giá trị như khoảng cách giữa các nguyên tử. Vì phép tính nhẹ, nó có thể xử lý cả hàng triệu nguyên tử. Thời gian quan sát cũng dài hơn nhiều.

Trường lực kinh nghiệm nhanh, nhưng có điểm yếu. Trường lực giả định liên kết cố định không mô tả tốt khoảnh khắc liên kết bị đứt và liên kết mới hình thành. Các trường lực phản ứng hoặc trường lực dành cho kim loại xuất hiện sau đó có thể xử lý phần nào sự thay đổi liên kết và khuyết tật trong phạm vi hẹp. Dù vậy, khó giữ độ chính xác trong điều kiện rộng. Chuyển pha, khi nguyên tử thay đổi vị trí lớn, cũng rất khó xử lý. Trong môi trường nóng và khắc nghiệt như động cơ tên lửa, các sự kiện như vậy xảy ra thường xuyên. Vì thế chỉ dùng trường lực giản lược sẽ tạo ra sai số lớn.

Có thể ví sự khác nhau giữa hai phương pháp này với đời sống thường ngày. DFT giống như chấm kỹ bài tập của một học sinh. Chính xác, nhưng nếu phải chấm bài tập của toàn trường thì mất rất nhiều thời gian. Trường lực kinh nghiệm giống như lướt nhanh qua đáp án để xem đúng hay không. Nhanh, nhưng gặp bài khó thì mắc lỗi. Trong một thời gian dài, các nhà khoa học phải chọn một trong hai: chính xác hoặc nhanh.

Thế năng liên nguyên tử học máy là nỗ lực phá bỏ bức tường lựa chọn ấy. Thế năng giống như một bản đồ cho biết mỗi cách sắp xếp nguyên tử có năng lượng bao nhiêu. Bản đồ này được gọi là bề mặt thế năng. MLIP học các giá trị chuẩn do DFT tính ra bằng học máy. Học máy là phương pháp tính toán tự tìm ra quy tắc từ nhiều ví dụ. Khi đã học đủ ví dụ, nó dự đoán nhanh năng lượng và lực của cách sắp xếp nguyên tử mới.

Ta hãy giải thích thêm một chút về cách MLIP học. Trước hết, dùng DFT để tính năng lượng và lực của nhiều cách sắp xếp nguyên tử. Kết quả tính toán này giống như một tập bài có sẵn đáp án. Mô hình học máy nhìn vào tập bài đó và học quy tắc. Nó tự nắm cảm giác vì sao một cách sắp xếp lại có năng lượng như vậy. Khi huấn luyện xong, nó cũng giải nhanh được cách sắp xếp mới chưa có trong tập bài.

Mục tiêu của phương pháp này rất rõ. Đó là tạo ra phép tính chính xác như DFT và nhanh như trường lực kinh nghiệm. Nếu đạt được mục tiêu này, ta có thể thử nghiệm vật liệu tên lửa trong máy tính ở quy mô lớn và trong thời gian dài. Ta có thể nhìn ở cấp độ nguyên tử xem kim loại dao động thế nào ở nhiệt độ cao, vết nứt lan ra sao. Ta có thể sàng lọc vật liệu ứng viên trước khi thử nghiệm thật. Phần còn lại của chương này là câu chuyện về nhiều công nghệ nhằm đạt mục tiêu đó.

Có một điểm cần nói trước. MLIP không thay thế hoàn toàn DFT. Vì nó chỉ học được khi có giá trị chuẩn do DFT tạo ra. Hai phương pháp này không cạnh tranh, mà đi thành một cặp. DFT tạo giá trị chuẩn, còn MLIP lan truyền nhanh các giá trị ấy. Sự phối hợp này làm cho các phép tính lớn trở nên khả thi.

## 2. Các mô hình học máy cổ điển dựa trên bộ mô tả

### Nội dung của mục này

Muốn học năng lượng nguyên tử bằng học máy, trước hết phải biến hình dạng xung quanh nguyên tử thành con số. Vì máy tính xử lý con số, không xử lý hình vẽ. Danh sách con số này được gọi là bộ mô tả (Descriptor). Mục này xem xét ba cách dùng bộ mô tả trong MLIP thời kỳ đầu.

Ba cách này có đặc điểm khác nhau. Một cách dùng mạng nơ-ron, một cách so sánh hai môi trường giống nhau đến đâu, và một cách học cả bộ mô tả bằng học sâu. Cả ba đều đạt kết quả tốt, nhưng cũng bộc lộ giới hạn. Chính các giới hạn đó gọi ra thế hệ công nghệ tiếp theo.

Một bộ mô tả tốt phải tuân theo một quy tắc. Cùng một vật chất luôn phải được biểu diễn bằng cùng một con số. Xoay cái bàn đi thì cái bàn vẫn là cái bàn. Cấu trúc nguyên tử cũng vậy, nếu xoay hoặc dịch chuyển toàn bộ, vị trí sẽ thay đổi. Nhưng vật chất vẫn không đổi. Bộ mô tả không được dao động trước những thay đổi như vậy. Tính chất này được gọi là tính bất biến (Invariance).

Nếu không giữ được tính bất biến, vấn đề sẽ xảy ra. Mô hình hiểu nhầm cùng một vật chất thành nhiều vật chất khác nhau. Khi đó nó phải học riêng nhiều lần cùng một thứ. Việc học chậm đi và dự đoán trở nên bất ổn. Nghiên cứu MLIP thời kỳ đầu tập trung đưa tính bất biến này vào bộ mô tả bằng thiết kế thủ công.

### 2.1 Mạng nơ-ron học bằng cách chia năng lượng cho từng nguyên tử (BPNN)

Behler và Parrinello đã đưa ra một ý tưởng quan trọng vào năm 2007. Đó là chia tổng năng lượng thành tổng năng lượng theo từng nguyên tử. Cách này được gọi là phân rã năng lượng nguyên tử. Mỗi nguyên tử chỉ nhìn xung quanh mình và đưa ra một giá trị năng lượng nhỏ. Cộng giá trị của tất cả nguyên tử lại thì được năng lượng của toàn bộ cấu trúc. Mạng nơ-ron này được gọi là mạng nơ-ron Behler-Parrinello (Behler-Parrinello Neural Network, BPNN).

Mạng nơ-ron là cấu trúc tính toán mô phỏng liên kết giữa các tế bào thần kinh trong não người. Nó nhận đầu vào, đi qua nhiều bước và đưa ra đáp án. Khi được huấn luyện bằng nhiều ví dụ, nó học quy tắc

giữa đầu vào và đáp án. BPNN nhận hình dạng xung quanh nguyên tử làm đầu vào và đưa ra năng lượng của nguyên tử đó làm đáp án.

Lúc này, công cụ biến môi trường xung quanh nguyên tử thành con số là hàm đối xứng tâm nguyên tử (Atom-Centered Symmetry Functions, ACSF). Hàm này đặt một nguyên tử ở trung tâm. Sau đó, nó xem xét các nguyên tử lân cận trong một khoảng cách nhất định. Nó ghi khoảng cách đến các nguyên tử lân cận và góc do các nguyên tử lân cận tạo thành con số. Nó được chia thành phần chỉ ghi khoảng cách và phần ghi cả góc. Các con số được tạo ra như vậy không thay đổi dù cấu trúc bị xoay hoặc dịch chuyển.

Phương pháp này cũng có ý nghĩa trong nghiên cứu vật liệu tên lửa. Khi kim loại nóng lên, các nguyên tử dao động mạnh quanh vị trí của chúng. Khoảng cách và góc đến các nguyên tử lân cận liên tục thay đổi. ACSF bắt được các thay đổi này bằng con số. Vì vậy nó đã trở thành bước đầu tiên để học chuyển động của vật liệu theo nhiệt độ.

Tuy nhiên, giới hạn rất rõ. Khi số loại nguyên tố tăng, lượng con số cần thiết tăng nhanh. Hợp kim dùng cho tên lửa thường trộn nhiều nguyên tố. Nếu có năm nguyên tố, số tổ hợp cặp nguyên tố tăng lên nhiều. Tổ hợp ba nguyên tử dùng để xử lý góc còn tăng nhanh hơn. Vấn đề gánh nặng tính toán tăng khi số loại như vậy tăng được gọi là lời nguyền chiều.

Một giới hạn khác là con người phải thiết kế bộ mô tả bằng tay. Con người quyết định đưa vào khoảng cách và góc nào, với độ dày đặc ra sao. Nếu lựa chọn này sai, thông tin quan trọng sẽ bị bỏ lỡ. Chính việc tạo ra bộ mô tả tốt đã là một bài toán khó. Nỗ lực giảm gánh nặng này đã dẫn đến các phương pháp tiếp theo.

Dù vậy, ý nghĩa mà phương pháp này để lại rất lớn. Ý tưởng chia tổng năng lượng theo từng nguyên tử vẫn còn sống đến hôm nay. Phần lớn các mô hình mới nhất cũng đi theo nguyên lý phân rã này. Mỗi nguyên tử nhìn vào môi trường xung quanh rồi đưa ra giá trị, sau đó cộng tất cả lại. Nhờ cách này, dù có nhiều nguyên tử, phép tính vẫn có thể được chia ra để xử lý. Ý tưởng của Behler và Parrinello đã trở thành nền móng của MLIP.

## 2.2 Thế năng đo độ giống nhau giữa hai môi trường (GAP)

Bartók, Csányi và cộng sự đã chọn một con đường khác vào năm 2010. Thay vì dùng mạng nơ-ron, họ trực tiếp so sánh hai môi trường giống nhau đến đâu. Phương pháp này được gọi là thế năng xấp xỉ Gaussian (Gaussian Approximation Potential, GAP). Khi gặp một môi trường nguyên tử mới, nó so sánh với các môi trường đã học. Nếu có nhiều môi trường giống, nó tham khảo năng lượng của chúng để đưa ra đáp án.

Công cụ cốt lõi của phép so sánh này là chồng lấp trơn các vị trí nguyên tử (Smooth Overlap of Atomic Positions, SOAP). SOAP mô tả vùng xung quanh nguyên tử không phải như những điểm cứng, mà như những đám mây mềm. Có thể hình dung rằng quanh mỗi nguyên tử lân cận ta đặt một quả cầu mờ. Nếu các đám mây của hai môi trường nguyên tử chồng lấp nhiều, chúng được xem là những môi trường giống nhau. Cách này được thiết kế để không bị dao động ngay cả khi cấu trúc quay.

Có lý do để mô tả bằng những đám mây mềm. Nguyên tử luôn dao động từng chút một. Nếu chỉ xem vị trí là những điểm cứng, giá trị sẽ nhảy mạnh chỉ vì một dao động nhỏ. Nếu mô tả bằng đám mây, dao động nhỏ sẽ được xử lý mềm mại. Vì vậy SOAP tỏ ra mạnh khi xử lý khuyết tật trong tinh thể hoặc vật liệu thiếu trật tự.

Thế năng xấp xỉ Gauss đạt độ chính xác cao ngay cả với ít dữ liệu. Điều này giúp ích lớn khi tính toán DFT tốn kém. Nó được dùng nhiều trong nghiên cứu khuyết tật tinh thể, trạng thái lỏng và vật liệu vô định hình thiếu trật tự. Trong vật liệu tên lửa, khuyết tật và sự vô trật tự cũng là chủ đề quan trọng. Vì vậy phương pháp này đã trở thành một chuẩn mực được trích dẫn trong thời gian dài.

Giới hạn lộ ra khi dữ liệu lớn lên. Phương pháp này so sánh môi trường mới với tất cả môi trường đã học. Dữ liệu huấn luyện càng nhiều, đối tượng cần so sánh càng tăng. Khi đó gánh nặng tính toán và lưu trữ cũng tăng. Với những dữ liệu lớn vượt quá hàng trăm nghìn mẫu như hiện nay, lượng tính toán là rất lớn.

Vì vậy phương pháp này đã nhường chỗ cho các cấu trúc khác trong những bài toán quy mô lớn. Để học nhanh từ dữ liệu lớn, mạng nơ-ron học sâu phù hợp hơn. Lý do là học sâu phát huy sức mạnh khi dữ liệu nhiều. Tuy vậy, chuẩn độ chính xác mà thế năng xấp xỉ Gauss đặt ra vẫn còn giá trị. Nó thường được dùng làm đối tượng so sánh khi đo hiệu năng của mô hình mới.

Ý tưởng mà SOAP để lại vẫn tiếp tục. Đó là cách nhìn vùng xung quanh nguyên tử như một mật độ mềm. Ý tưởng này đã thấm vào nhiều mô hình ra đời sau đó. Khi xử lý bằng phân bố mềm thay vì điểm cứng, phép tính ổn định hơn. Lý do là giá trị không nhảy vì những dao động nhỏ. Trong tính toán nhiệt độ cao, nơi nguyên tử dao động mạnh như vật liệu tên lửa, sự ổn định này rất hữu ích.

## 2.3 Học sâu tự học cả bộ mô tả (DeePMD)

Zhang, Car và các cộng sự đã đưa học sâu vào tính toán nguyên tử một cách nghiêm túc vào năm 2018. Phương pháp này được gọi là động lực học phân tử thế năng sâu (Deep Potential Molecular Dynamics, DeePMD). Động lực học phân tử là mô phỏng theo dõi chuyển động của nguyên tử theo thời gian. DeePMD dùng học sâu để tính lực dùng trong mô phỏng này.

Các phương pháp trước đó để con người thiết kế bộ mô tả. DeePMD đã giảm mạnh gánh nặng này. Mạng nơ-ron tự biến đổi tọa độ xung quanh nguyên tử thành dạng tốt. Con người không phải chọn từng khoảng cách và góc. Trong khi đó, mô hình vẫn giữ được tính chất không dao động trước phép quay và phép tịnh tiến.

Ở đây cũng cần giải thích thêm một chút về từ học sâu. Học sâu chỉ mạng nơ-ron được xếp thành nhiều lớp. Khi xếp các lớp đủ sâu, mô hình có thể học cả những quy tắc phức tạp. Mạng nơ-ron nông chỉ học các quy tắc đơn giản. Mạng nơ-ron sâu nắm bắt được cả quan hệ chằng chịt giữa các nguyên tử. Đối lại, việc học cần nhiều dữ liệu và tính toán. Từ sâu xuất phát từ việc các lớp được xếp sâu.

Sức mạnh của phương pháp này nằm ở quy mô lớn. Với học sâu, gánh nặng tính toán không tăng quá mạnh dù số nguyên tử nhiều. Vì vậy DeePMD có thể xử lý chuyển động của hàng triệu nguyên tử. Nó đã được dùng cho nhiều vật liệu như nước, kim loại và chất bán dẫn. Nó phát huy sức mạnh trong việc quan sát các hệ lớn trong thời gian dài.

DeePMD cũng có ý nghĩa đối với nghiên cứu vật liệu tên lửa. Muốn thấy một khối kim loại thay đổi ra sao ở nhiệt độ cao, cần có nhiều nguyên tử. Nếu chỉ nhìn một mảnh nhỏ, ta bỏ lỡ hành vi của toàn hệ. DeePMD có thể xử lý mảnh lớn, nên phù hợp với loại nghiên cứu này.

Tuy vậy, giới hạn chung của các phương pháp bộ mô tả ban đầu vẫn còn. Những phương pháp này chỉ nhìn riêng từng láng giềng trong một khoảng cách đã định. Muốn chứa ảnh hưởng của nguyên tử ở xa, phải tăng khoảng cách quan sát. Khi đó số láng giềng tăng mạnh và gánh nặng tính toán lớn lên. Về mặt cấu trúc, rất khó đưa tương tác tầm xa vào.

Một cánh cửa khác mà DeePMD mở ra là sự kết hợp với công cụ tính toán lớn. Phương pháp này đã dẫn tới các nghiên cứu xử lý hàng chục tỷ nguyên tử trên siêu máy tính lớn. Nó cũng được dùng để tính các hợp kim pha trộn nhiều nguyên tố ở quy mô lớn. Nghiên cứu hợp kim nhiệt độ cao dùng cho tên lửa cần loại tính toán lớn như vậy. Lý do là những hành vi bị bỏ lỡ khi chỉ nhìn mảnh nhỏ có thể được thấy trong hệ lớn.

Giới hạn này đã gọi ra một ý tưởng mới. Đó là vượt khỏi cách đặt một nguyên tử ở trung tâm rồi nhìn xung quanh nó. Nếu để các nguyên tử trao đổi thông tin với nhau thì sao. Khi thông tin được trao đổi nhiều lần, ảnh hưởng của nguyên tử ở xa cũng được truyền tiếp. Ý tưởng này dẫn tới mạng nơ-ron đồ thị ở phần sau.

### 3. Mạng nơ-ron đồ thị (GNN) đẳng biến (Equivariant) hiện đại

Phần này bàn về phương pháp đã trở thành dòng chính của MLIP hiện nay. Đó là mạng nơ-ron đồ thị (Graph Neural Network, GNN). Phương pháp này xem nguyên tử là điểm và quan hệ láng giềng giữa các nguyên tử là đường. Ở đây đường không phải chính liên kết hóa học. Nó có nghĩa là các láng giềng nằm trong một khoảng cách nhất định được nối với nhau. Đường này cũng có thể thay đổi trong quá trình tính toán. Nó giống như bản đồ tàu điện ngầm vẽ nhà ga thành điểm và kết nối giữa các ga thành đường. Mô hình tính năng lượng bằng cách trao đổi thông tin dọc theo các điểm và đường.

Từ khóa cốt lõi của phần này là tính đẳng biến (Equivariance). Đây là tính chất đi theo cặp với tính bất biến ở phần trước. Trước hết, tôi giải thích tính đẳng biến là gì và vì sao nó cần cho tính toán nguyên tử. Sau đó, tôi lần lượt xem xét các mô hình tiêu biểu chứa tính chất này, từ giai đoạn đầu đến hiện đại. Cuối cùng, tôi bàn đến cả các mô hình lớn đã học trước dữ liệu rộng.

Mạng nơ-ron đồ thị tính toán bằng cách trao đổi thông tin. Mỗi nguyên tử gửi thông điệp cho các láng giềng của nó. Nó gom các thông điệp đến từ láng giềng rồi cập nhật trạng thái của mình. Quá trình này được lặp lại nhiều lần. Trao đổi một lần thì biết nguyên tử ngay bên cạnh. Trao đổi hai lần thì biết đến láng giềng của láng giềng. Càng lặp lại, ảnh hưởng của nguyên tử xa hơn càng được truyền tới. Cách này được gọi là truyền thông điệp (Message Passing).

Cấu trúc này giải quyết phần nào giới hạn của phần trước. Không cần miễn cưỡng tăng khoảng cách quan sát. Lý do là khi trao đổi thông điệp nhiều lần, thông tin vượt ra ngoài láng giềng sẽ được truyền tới. Tuy nhiên, nếu số lớp đã định thì phạm vi thông tin chạm tới cũng đã định. Những hiệu ứng như tĩnh điện rất xa hoặc lực phân tán rất xa khó được chứa đầy đủ chỉ bằng cách này. Các hiệu ứng đó phải được cộng thêm bằng một hạng tầm xa được tạo riêng. Dù vậy, mạng nơ-ron đồ thị là một bước tiến lớn để chứa hiệu quả các tương tác trên phạm vi rộng.

#### 3.1 Tính đối xứng vẫn đúng dù quay hay lật

Trong thế giới nguyên tử có những đối xứng được bảo toàn. Dù tịnh tiến hay quay cả cấu trúc, các định luật vật lý vẫn như cũ. Tập hợp cả tịnh tiến, quay và lật được gọi là  $E(3)$ . Riêng tập hợp các phép quay được gọi là  $SO(3)$ . Tên gọi khó, nhưng ý nghĩa thì dễ. Đó là tập hợp của tịnh tiến, quay và lật như trong gương. Quay và lật là hai chuyển động khác nhau. Phép quay xoay hình dạng như cũ, còn phép lật tạo ra ảnh gương đối trái phải.

Tính bất biến đã nói ở trên là tính chất mà giá trị không đổi trước những chuyển động như vậy. Năng lượng là một đại lượng bất biến. Dù quay cấu trúc, năng lượng của cấu trúc đó phải như nhau. Nhưng không phải giá trị nào cũng bất biến. Lực là giá trị có hướng. Khi quay cấu trúc, hướng của lực cũng phải quay theo.

Chính tính chất này là tính đẳng biến. Tính đẳng biến là tính chất mà khi đầu vào quay, đầu ra cũng quay theo cùng cách. Lực là đại lượng đẳng biến. Nếu quay cấu trúc sang phải, lực tác dụng lên từng nguyên tử cũng quay sang phải. Tính toán nguyên tử chỉ phù hợp với định luật vật lý khi giữ được cả tính bất biến và tính đẳng biến.

Có thể tách đại lượng bất biến và đại lượng đẳng biến trong một câu. Nếu quay mà giá trị vẫn giữ nguyên thì đó là đại lượng bất biến. Nếu quay mà giá trị cũng quay theo thì đó là đại lượng đẳng biến.

Năng lượng là một con số không có hướng, nên là đại lượng bất biến. Các giá trị như nhiệt độ hoặc mật độ cũng không có hướng, nên là đại lượng bất biến. Những giá trị có hướng như lực hoặc vận tốc là đại lượng đẳng biến. Mô hình phải biết mỗi giá trị thuộc bên nào để xử lý.

Ở đây tôi nối tiếp câu chuyện đã để lại từ phần trước. Tôi đã nói lực là độ dốc của bản đồ năng lượng. Một mô hình tốt không khớp lực riêng rẽ. Trước hết nó tính năng lượng, rồi lấy lực từ độ dốc của bản đồ năng lượng đó. Có lý do để giữ thứ tự này. Nếu xuất lực tùy tiện, có thể sinh ra lực không bảo toàn năng lượng. Nếu tính lâu bằng lực như vậy, năng lượng vốn không có sẽ tự sinh ra hoặc biến mất. Khi đó mô phỏng sẽ dần phình lên hoặc sụp xuống. Nếu lấy lực từ năng lượng bằng độ dốc, sai lệch như vậy sẽ không xảy ra. Trong vật lý, lực sinh ra từ năng lượng như vậy được gọi là lực bảo toàn.

Nếu đưa hai tính chất này vào mô hình ngay từ đầu, ta có lợi ích lớn. Mô hình không phải học riêng cùng một vật chất ở nhiều hướng khác nhau. Quy tắc đã học một lần tự động đúng cho mọi hướng. Vì vậy mô hình học ổn định ngay cả với ít dữ liệu. Vì tính toán DFT đắt đỏ, việc tiết kiệm dữ liệu này có sức nặng lớn.

Trong vật liệu tên lửa, có nhiều sự kiện mà hướng đóng vai trò quan trọng. Vết nứt lan theo một hướng nhất định. Biến dạng trượt dọc theo mặt tinh thể cũng có hướng. Phản ứng oxy hóa cũng diễn ra khác nhau tùy hướng của bề mặt. Nếu không xử lý chính xác hướng của lực, mô hình sẽ mô tả sai các sự kiện này. Mô hình giữ tính đẳng biến có lợi thế ở điểm đó.

Hãy chuyển tính đối xứng này thành một hình ảnh dễ hiểu hơn. Giả sử ta chụp nhiều ảnh của cùng một cấu trúc khối đồ chơi. Ảnh chụp từ phía trước và ảnh chụp từ bên cạnh trông khác nhau. Nhưng bản thân cấu trúc khối vẫn là một. Con người biết điều này ngay. Nhưng với máy tính, hai bức ảnh là hai đầu vào khác nhau. Mô hình không có tính đối xứng có thể xem hai bức ảnh là hai vật liệu khác nhau. Mô hình có tính đối xứng xem hai bức ảnh là cùng một vật liệu. Khác biệt này quyết định lớn đến hiệu quả học.

Tóm lại, tính đối xứng không phải trang trí mà là khung xương. Mô hình giữ tính đối xứng tiết kiệm dữ liệu và làm dự đoán ổn định. Mô hình vi phạm tính đối xứng có thể cho ra câu trả lời kỳ lạ về mặt vật lý. Vì vậy nghiên cứu MLIP hiện đại tập trung vào cách đưa tốt tính chất này vào mô hình.

### 3.2 Từ cách chỉ nhìn khoảng cách sang cách trao đổi cả hướng

SchNet là một mạng nơ-ron đồ thị giai đoạn đầu ra đời năm 2018. Mô hình này chỉ trao đổi khoảng cách giữa các nguyên tử dưới dạng thông điệp. Khoảng cách là giá trị không có hướng. Vì vậy SchNet là mạng nơ-ron đồ thị bất biến chỉ xử lý thông tin bất biến. Cấu trúc gọn gàng và tính toán nhanh.

Cách chỉ dùng khoảng cách có điểm yếu. Nó khó chứa trực tiếp các liên kết mà hướng đóng vai trò quan trọng. Dù khoảng cách giống nhau, nếu góc khác nhau thì liên kết cũng khác. Nếu bỏ lỡ khác biệt này, sai số sẽ sinh ra. DimeNet bổ sung điểm này bằng cách đưa cả góc giữa các nguyên tử vào thông điệp. Nỗ lực đưa thông tin hướng vào mô hình đã tiếp tục như vậy.

Batzner, Kozinsky và các cộng sự đã tiến một bước lớn vào năm 2022. Đó là mô hình NequIP. Mô hình này vượt khỏi khoảng cách và góc, trao đổi chính hướng dưới dạng thông điệp. Nó truyền và trộn nguyên vẹn các giá trị có hướng. Đây là truyền thông điệp đẳng biến. Nó chứa hướng của lực và hình dạng 3 chiều xung quanh nguyên tử một cách trung thực hơn nhiều.

Ưu điểm lớn của NequIP là hiệu quả dữ liệu. Vì chứa thông tin hướng theo đúng nguyên lý, nó học tốt ngay cả với ít ví dụ. Bài báo công khai cho biết mô hình đạt độ chính xác cao với dữ liệu huấn luyện ít hơn nhiều so với các mô hình trước. Trong nghiên cứu mà từng phép tính DFT đều đắt đỏ, sự tiết kiệm này là lợi ích thực tế.

Cần nhắc lại vì sao hiệu quả dữ liệu quan trọng. Một phép tính DFT cần nhiều thời gian máy tính. Nếu huấn luyện cần hàng chục nghìn phép tính, chi phí tính toán sẽ phình to. NequIP đạt độ chính xác tương tự với số mẫu ít hơn nhiều. Khi đó, với cùng ngân sách, ta có thể xử lý vật liệu khó hơn. Vật liệu mới dùng cho tên lửa thường thiếu dữ liệu. Mô hình có hiệu quả dữ liệu cao rất hữu dụng trong hoàn cảnh như vậy.

Bước phát triển này có ý nghĩa trực tiếp đối với tính toán vật liệu tên lửa. Vết nứt, khuyết tật và chuyển pha đều gắn với hướng. Mô hình đẳng biến mô tả chính xác hơn các sự kiện có hướng như vậy. Ngay cả với ít dữ liệu DFT, nó vẫn có thể học hành vi trong môi trường cực hạn. Vì vậy sau NequIP, phương pháp đẳng biến đã gần trở thành chuẩn trong lĩnh vực này.

Cùng thời kỳ với NequIP, các nghiên cứu nhằm nâng hiệu quả cũng tiếp tục. Đó là các thử nghiệm xử lý thông tin hướng nhưng làm cho tính toán nhẹ hơn. Một nghiên cứu gần đây đã đưa ra mô hình đẳng biến có hiệu quả cao hơn và được đăng trên tạp chí vào năm 2025. Hướng đi này là giảm gánh nặng tính toán trong khi vẫn giữ độ chính xác. Khi những nỗ lực như vậy tích lũy, phương pháp đẳng biến trở nên đủ thực dụng cho nghiên cứu thực tế.

Tuy nhiên, nó có cái giá phải trả. Muốn trộn thông tin về hướng theo đúng nguyên lý thì cần các phép toán nặng. Phép toán này được gọi là tích tensor. Đổi lại độ chính xác, khối lượng tính toán và gánh nặng bộ nhớ lớn. Gánh nặng này về sau lại nổi lên thành vấn đề. Câu chuyện này sẽ được nói tiếp ở mục 5.

### 3.3 Cách đưa quan hệ của nhiều nguyên tử vào cùng lúc(MACE)

Năm 2022, Batatia, Csányi và các cộng sự đã đưa ra MACE. Mô hình này kết hợp mạng nơ-ron đồ thị đồng biến với một lý thuyết vật lý cũ gọi là khai triển đa thể. Đa thể có nghĩa là hiệu ứng do nhiều nguyên tử cùng tạo ra. Trong vật liệu, quan hệ giữa một nguyên tử và một nguyên tử khác không phải là điều duy nhất quan trọng. Cách sắp xếp do ba nguyên tử, bốn nguyên tử cùng tạo ra cũng làm thay đổi tính chất.

Các mạng nơ-ron đồ thị trước đó trao đổi thông điệp nhiều lần để đưa những hiệu ứng đa thể này vào. Khi số lần trao đổi tăng lên, gánh nặng tính toán lớn hơn và việc học chậm lại. MACE tìm một con đường khác. Nó nhai các thông tin láng giềng đã gom được với nhau để tạo hiệu ứng của nhiều nguyên tử cùng lúc. Vì vậy, chỉ với ít lần truyền thông điệp, nó vẫn đưa vào được cả hiệu ứng của bốn nguyên tử trở lên.

Lợi ích của thiết kế này nằm ở chỗ nắm được cả độ chính xác và tốc độ. Khi giảm số lần truyền thông điệp, tính toán nhẹ hơn. Đồng thời, vì không bỏ lỡ hiệu ứng đa thể nên độ chính xác được giữ vững. Ưu điểm này càng lớn trong tinh thể lớn hoặc trong động lực học phân tử kéo dài. Vì vậy, MACE được dùng như một mô hình chuẩn mạnh trong nhiều nghiên cứu.

MACE cũng được dùng trong nghiên cứu vật liệu cho tên lửa. Gần đây, một nhóm nghiên cứu đã xử lý hợp kim entropy cao trộn nhôm, titan, crom, molybdenum và tungsten. Hợp kim này là vật liệu trộn nhiều nguyên tố với lượng gần tương đương nhau. Nhóm nghiên cứu đã dùng MACE để tạo thế năng liên nguyên tử của hợp kim này. Sau đó họ xem xét độ ổn định pha theo nhiệt độ. Nghiên cứu này được đăng trên một tạp chí vật lý vào năm 2026. Đây là một ví dụ dùng MACE trong thiết kế hợp kim cho nhiệt độ cao. Tuy nhiên, điều nghiên cứu này cho thấy là phép tính độ ổn định pha. Nghiên cứu riêng lẻ này chưa kiểm chứng đến mức nứt hoặc tuổi thọ của linh kiện thực tế.

Ta cũng cần nhìn vào giới hạn. MACE cũng dùng tích tensor khi trộn thông tin về hướng. Vì vậy, nếu nâng độ tinh vi của hướng, tức mô men động lượng, gánh nặng tính toán tăng rất nhanh. Trong các hệ rất lớn hoặc mô phỏng trong thời gian rất dài, gánh nặng này vẫn là vấn đề. Các nghiên cứu nhằm giảm điểm này sẽ được nói tiếp ở phần sau.

Có thể chuyển hiệu ứng đa thể sang một ví dụ dễ hiểu. Nếu chỉ nhìn hai học sinh trong lớp, ta chỉ biết họ có thân nhau hay không. Khi ba học sinh tụ lại, bầu không khí thay đổi. Khi bốn học sinh tụ lại, một dòng chảy khác lại xuất hiện. Nguyên tử cũng vậy. Chỉ quan hệ giữa hai nguyên tử không thể chứa hết tính chất của vật liệu. Cách sắp xếp do nhiều nguyên tử cùng tạo ra làm thay đổi tính chất. MACE đưa quan hệ của nhiều nguyên tử này vào bằng ít bước.

Một nhiệm vụ khác là các tình huống nằm ngoài phạm vi đã học. Dù MACE chính xác đến đâu, nó vẫn có thể sai trong những cách sắp xếp lạ chưa có trong dữ liệu học. Môi trường tên lửa luôn cực hạn và có nhiều sự kiện lạ. Vì vậy, mô hình càng chính xác càng cần một thiết bị tự báo khi nó gặp điều mình không biết. Thiết bị đó là học chủ động ở mục 4.

### 3.4 Mô hình lớn học trước trên phạm vi vật liệu rộng(CHGNet, MatterSim)

Các mô hình trước thường được huấn luyện nhằm vào một vật liệu hoặc một phạm vi hẹp. Khi gặp vật liệu mới, ta phải huấn luyện lại bằng dữ liệu phù hợp với vật liệu đó. Dòng nghiên cứu muốn giảm gánh nặng này là thế năng nền tảng. Thế năng nền tảng là mô hình lớn đã học trước từ dữ liệu vật liệu rất rộng. Nó giống một đầu bếp đã quen với nhiều món ăn nên cũng ứng phó khá tốt với nguyên liệu mới.

CHGNet là một thế năng nền tảng tiêu biểu xuất hiện vào năm 2023. Mô hình này được huấn luyện bằng một cơ sở dữ liệu tính toán lớn gọi là Materials Project. Materials Project là nguồn dữ liệu công khai tập hợp kết quả tính toán DFT của rất nhiều vật liệu. CHGNet còn học thêm cả manh mối về điện tích và trạng thái từ. Vì vậy, nó theo dõi tốt sự di chuyển của lithium trong vật liệu pin hoặc sự thay đổi pha của oxide kim loại chuyển tiếp.

Có lý do khiến CHGNet học điện tích và trạng thái từ. Trong vật liệu pin, lithium đi vào và đi ra, kéo theo sự trao đổi điện tích. Trong oxide kim loại chuyển tiếp, trạng thái từ của nguyên tử làm thay đổi tính chất. Nếu bỏ lỡ những thay đổi này, ta sẽ vẽ sai hành vi của vật liệu. CHGNet đưa các manh mối này vào biểu diễn môi trường nguyên tử, nhờ đó bao quát phạm vi vật liệu rộng hơn. Tuy nhiên, đây không phải là mô hình điện hóa theo dõi tỉ mỉ điện tích như một đại lượng được bảo toàn. Cách hiểu đúng hơn là nó đưa các đầu mối về trạng thái từ vào biểu diễn đã học. Trong vật liệu cho tên lửa, oxide và kim loại chuyển tiếp cũng thường được dùng, nên năng lực này có ích.

MatterSim là một thế năng nền tảng khác được công bố vào năm 2024. Mô hình này nhằm vào phạm vi nhiệt độ và áp suất rất rộng. Bài báo công khai cho biết nó xử lý từ 0 kelvin đến 5000 kelvin, và từ 0 gigapascal đến 1000 gigapascal. Kelvin là đơn vị nhiệt độ tuyệt đối, còn gigapascal là đơn vị áp suất. Điều kiện nóng của động cơ tên lửa cũng nằm trong phạm vi này. Tuy nhiên, phạm vi học có giao nhau và việc mô hình thật sự dự đoán đúng là hai chuyện khác nhau. 1000 gigapascal là giá trị cao hơn nhiều so với áp suất buồng đốt tên lửa. Những tình huống riêng của tên lửa như khí cháy hoặc bề mặt oxy hóa cần được kiểm chứng riêng. Vì vậy, mô hình này được quan tâm như điểm khởi đầu cho nghiên cứu vật liệu cực hạn. Bài báo công khai cho biết mô hình này đã tiết kiệm được nhiều dữ liệu. Nhờ đã học trước các điều kiện rộng, nó phản ứng nhanh với vật liệu mới.

Gần đây cũng đã xuất hiện một thế năng nền tảng mới nâng cả độ chính xác và tốc độ. Đó là mô hình GRACE. Mô hình này được xây dựng bằng khung khai triển cụm nguyên tử đồ thị. Nhóm nghiên cứu đã huấn luyện mô hình này bằng một phần của bộ dữ liệu vật liệu thuộc loại rất lớn. Nghiên cứu công khai cho biết mô hình này đã dựng lên một đường cân bằng mới giữa độ chính xác và hiệu suất. Nghiên cứu này được đăng trên một tạp chí tính toán vật liệu vào năm 2026. Nó cho thấy rằng nếu tinh chỉnh tốt mô hình lớn, ta có thể giữ độ chính xác trong khi điều chỉnh nó cho phù hợp với một nhiệm vụ cụ thể.

Thế năng nền tảng đang thay đổi cách nghiên cứu. Trước đây, với mỗi vật liệu, người ta tạo mô hình từ đầu. Giờ đây, người ta dùng mô hình lớn trước rồi chỉ chỉnh phần cần thiết. Cách chỉ sửa một chút rồi

dùng như vậy được gọi là tinh chỉnh. Khi gặp hợp kim mới, ta cũng có thể phản ứng nhanh bằng ít dữ liệu.

Nỗ lực so sánh công bằng các mô hình lớn này cũng đang tăng lên. Khi nhiều thế năng nền tảng xuất hiện, việc phán đoán mô hình nào tốt hơn trở nên khó khăn. Vì vậy, đã xuất hiện các sân đánh giá công khai đo hiệu năng trong cùng điều kiện. Đó là các nền tảng benchmark do nhiều nhóm nghiên cứu tạo ra. Những sân này so sánh đồng thời độ chính xác, tốc độ và độ ổn định. Người dùng có thể chọn mô hình phù hợp với vấn đề của mình.

Dù vậy, vẫn có điểm cần thận trọng. Thế năng nền tảng không tự động đúng trong mọi tình huống. Sai số có thể tăng trong các hợp kim hoặc phản ứng bề mặt hiếm gặp trong dữ liệu học. Vì vậy, mô hình tốt càng phải đi kèm kiểm chứng. Cấu trúc lạ cần được tính lại bằng DFT để xác nhận. Quy trình xác nhận này là chủ đề của mục tiếp theo.

## 4. Vòng lặp học chủ động(Active Learning), đánh giá sai số và độ bất định

Dù mô hình tốt đến đâu, nó vẫn có thể sai trong tình huống chưa học. Vấn đề là khi mô hình sai nhưng vẫn đưa ra câu trả lời với vẻ tự tin. Khi đó, con người khó nhận ra sai số. Mục này bàn về cách làm cho mô hình tự báo thời điểm nó không biết.

Cách đó là học chủ động(Active Learning). Học chủ động là quy trình lặp trong đó mô hình chọn các cấu trúc mà nó không tự tin rồi tính lại chính xác. Ta sẽ xem vòng lặp này chạy ra sao và đo sự thiếu tự tin bằng con số như thế nào. Ta cũng sẽ chỉ ra vì sao quy trình này là thiết bị an toàn trong môi trường cực hạn và xa lạ như vật liệu tên lửa.

Có thể chuyển nội suy và ngoại suy sang ví dụ đời sống. Nếu biết nhiệt độ của nhiều năm trước, ta có thể ước lượng nhiệt độ hôm nay. Ước lượng trong phạm vi đã trải qua là nội suy. Nhưng nếu một đợt nóng chưa từng có xảy ra, ước lượng sẽ chệch đi. Đi ra ngoài phạm vi chưa từng trải qua là ngoại suy. Mô hình cũng vậy, nó khớp tốt trong phạm vi đã trải qua và dao động ở bên ngoài.

Nơi mô hình dự đoán an toàn là bên trong dữ liệu học. Vùng này được gọi là vùng nội suy. Bên trong phạm vi được dữ liệu học bao quanh, dự đoán đáng tin. Nhưng khi đi ra ngoài phạm vi đó, tình hình khác đi. Phần bên ngoài này được gọi là vùng ngoại suy. Trong vùng ngoại suy, mô hình có thể đưa ra câu trả lời trông có vẻ hợp lý nhưng sai.

Sai số trong vùng ngoại suy rất nguy hiểm. Trong khi mô phỏng, các nguyên tử có thể chồng lên nhau hoặc cấu trúc có thể sụp đổ theo cách kỳ lạ. Nếu điều này xảy ra lặng lẽ, toàn bộ kết quả sẽ không còn đáng tin. Môi trường tên lửa luôn cực hạn nên ngoại suy xảy ra thường xuyên. Nhiệt độ cao, va chạm mạnh, oxy hóa và vết nứt là những sự kiện hiếm trong dữ liệu học. Vì vậy, nhất định cần có thiết bị bắt được những khoảnh khắc này.

Vòng lặp học chủ động xoay qua bốn bước. Trước hết, ta huấn luyện mô hình bằng một bộ dữ liệu DFT nhỏ. Sau đó, dùng mô hình này để tính nhanh chuyển động nguyên tử. Trong khi tính, nếu xuất hiện cấu trúc khiến mô hình không tự tin, ta đánh dấu lại. Chỉ chọn cấu trúc đó và tính lại chính xác bằng DFT. Đưa kết quả mới vào dữ liệu rồi sửa mô hình. Nếu lặp lại các bước này, mô hình dần học được phạm vi tình huống rộng hơn. Nó giống việc đi trên con đường lạ và kiểm tra lại bản đồ ở mỗi ngã rẽ.

Cốt lõi của vòng lặp này là đo sự thiếu tự tin bằng con số. Con số này được gọi là độ bất định(Uncertainty). Có nhiều cách đo độ bất định. Các tiểu mục tiếp theo sẽ xem xét những phương pháp tiêu biểu.

## 4.1 Đo bằng con số thời điểm mô hình không biết(UQ)

Cách đầu tiên để đo độ bất định là dùng độ ngoại suy. Mô hình gọi là thế năng tensor mô men(Moment Tensor Potential, MTP) dùng cách này. Mô hình này đo bằng con số xem cách sắp xếp nguyên tử hiện tại khác dữ liệu học đến mức nào. Con số này được gọi là độ ngoại suy. Nếu độ ngoại suy thấp, đó là cách sắp xếp quen thuộc và dự đoán an toàn. Nếu độ ngoại suy cao, đó là cách sắp xếp lạ nên cần tính lại.

Cách thứ hai là so sánh câu trả lời của nhiều mô hình. Ta tạo nhiều mô hình được huấn luyện hơi khác nhau từ cùng dữ liệu. Tập hợp này được gọi là ensemble. Với cấu trúc mới, nếu câu trả lời của các mô hình này giống nhau, nhìn chung có thể tin được. Nếu câu trả lời khác nhau nhiều, cấu trúc đó là tín hiệu nguy hiểm. Tuy nhiên, các mô hình học từ cùng dữ liệu cũng có thể cùng sai trước đầu vào lạ. Vì vậy, sự nhất trí của ensemble chỉ là chỉ báo tham khảo, không phải bảo đảm đầy đủ. Nó giống việc hỏi cùng một câu cho nhiều người, nếu câu trả lời phân tán thì xem đó là vấn đề khó.

Cách thứ ba là để chính mô hình đưa ra khoảng sai số. Nghiên cứu gần đây đã tiến triển theo hướng này. Một nhóm nghiên cứu đã gắn phương pháp gọi là học sâu dựa trên bằng chứng vào thế năng liên nguyên tử. Phương pháp này đưa ra giá trị dự đoán cùng với mức độ có thể tin giá trị đó. Vì không phải chạy riêng nhiều mô hình nên tính toán nhẹ hơn. Nghiên cứu này được đăng trên tạp chí vào năm 2025.

Cũng có nghiên cứu tìm cách đo độ bất định sâu hơn vào học chủ động. Một phương pháp khiến mô phỏng tự đi tìm những nơi bất định. Nó dẫn chuyển động nguyên tử về phía các cách sắp xếp mà mô hình không tự tin. Khi đó, ta thu thập nhanh hơn các cấu trúc có giá trị học tập lớn. Đây là cách lấp đầy hiệu quả các tình huống xa lạ.

Quy trình này là khung xương giữ an toàn trong nghiên cứu vật liệu tên lửa. Trong môi trường cực hạn, các cách sắp xếp không dự đoán được luôn xuất hiện. Nếu không có thiết bị độ bất định, ta có nguy cơ tin kết quả sai là đúng. Nếu có thiết bị độ bất định, phép tính dừng lại ở thời điểm nguy hiểm để kiểm tra. Nhờ đó, ta có thể tin kết quả hơn.

Các cách xử lý độ bất định vẫn đang tiếp tục xuất hiện. Gần đây có nghiên cứu xử lý cả trường hợp mô hình được thiết kế sai. Nếu mô hình sai ngay từ giả định cơ bản, độ bất định cũng sẽ sai. Nghiên cứu chỉ ra vấn đề này đã được đăng trên tạp chí vào năm 2025. Điều đó có nghĩa là đo độ bất định là việc khó. Dù vậy, tiến bộ theo hướng này chống đỡ sự an toàn của tính toán vật liệu cực hạn.

Một lợi ích khác của quy trình lặp này là tiết kiệm dữ liệu. Tính toán DFT đắt đỏ. Vì vậy, nếu tính bừa bất kỳ cấu trúc nào, lãng phí sẽ lớn. Học chủ động chỉ chọn các cấu trúc có giá trị học tập lớn để tính. Với cùng ngân sách, ta thu thập được dữ liệu hữu ích hơn. Nhờ đó, ta có được mô hình tốt nhanh hơn.

Có thể ví cách này với việc ôn thi. Nếu giải lại bài đã biết, thời gian sẽ uổng phí. Nếu chỉ chọn bài sai để giải, năng lực tăng nhanh. Học chủ động cũng chỉ chọn và lấp đầy những phần mô hình còn yếu. Đây là cách đạt tiến bộ lớn bằng ít tính toán. Trong nghiên cứu vật liệu tên lửa, ngân sách tính toán luôn eo hẹp. Vì vậy, sự tiết kiệm này trở thành sức mạnh trong phát triển thực tế.

Tuy nhiên, bản thân việc đo chính xác độ bất định là một bài toán khó. Có phương pháp đòi hỏi chi phí tính toán lớn, có phương pháp lại kém chính xác. Nếu mô hình ngay từ đầu đã được thiết kế sai, ngay cả độ bất định cũng có thể cho kết quả sai. Gần đây, các nghiên cứu xử lý vấn đề này vẫn tiếp tục. Định lượng độ bất định chưa phải là một kỹ thuật đã hoàn tất, mà là một lĩnh vực đang phát triển.

## 5. Mạng nơ-ron hình học loại bỏ tích tensor: hiệu quả tính toán và độ trơn của đường cong

Mô hình hiệp biến chính xác, nhưng mang gánh nặng tính toán lớn. Một nguyên nhân lớn của gánh nặng đó là tích tensor đã nói ở trên. Mục này bàn về lý do vì sao tích tensor gây nhiều tải tính toán, và các nghiên cứu mới đang cố giảm hoặc loại bỏ nó là gì.

Ở đây còn có một yếu tố quan trọng khác là độ trơn của đường cong. Kết quả của tính toán nguyên tử là bề mặt thế năng, tức bản đồ năng lượng. Nếu bề mặt này gồ ghề, phép tính lực sẽ bị giật và mô phỏng sẽ đổ vỡ. Mục này giải thích quá trình các phương pháp mới cố nắm bắt cả tốc độ lẫn độ trơn.

Tích tensor là một phép toán toán học trộn các giá trị có hướng theo đúng nguyên lý. Mô hình hiệp biến dùng phép toán này để xử lý chính xác thông tin hướng. Vấn đề là khi nâng độ chính xác của hướng, gánh nặng của phép toán này tăng rất nhanh. Thời gian tính toán và bộ nhớ cùng phình ra. Trong mô phỏng hệ lớn hoặc thời gian dài, gánh nặng này trở thành một bức tường.

Hãy hình dung tích tensor bằng lời dễ hơn một chút. Khi trộn thông tin có hướng, quy tắc rất khắt khe. Không thể chỉ cộng hoặc nhân hai giá trị hướng với nhau. Phải ghép cặp và trộn chúng sao cho phù hợp với định luật vật lý. Tích tensor thực hiện chính xác việc ghép cặp khắt khe này. Đổi lại cho độ chính xác, số cặp tăng nhanh theo độ chính xác của hướng. Vì vậy, chỉ cần tăng độ chính xác một chút, gánh nặng tính toán đã lớn lên.

Các nghiên cứu nhằm vượt qua bức tường này xuất hiện theo nhiều hướng. Một hướng là tạo ra chương trình đặc biệt để tính tích tensor nhanh hơn. Một nghiên cứu gần đây đã làm phép tính động lực học phân tử nhanh hơn nhiều theo cách này. Nghiên cứu này được đăng trên tạp chí khoa học vào năm 2026. Đây là con đường giữ lại bản thân tích tensor nhưng tối ưu hóa phép tính.

### 5.1 Thiết kế mới hoàn toàn không dùng tích tensor

Một hướng khác là hoàn toàn không dùng tích tensor. Nghiên cứu có tên Geodite đã chọn con đường này. Tiêu đề của nghiên cứu này là thế năng liên nguyên tử hiệp biến không dùng tích tensor. Nó thay thế phép toán nặng dùng để trộn hướng bằng một cách khác. Dù vậy, nó vẫn giữ đối xứng quay theo đúng nguyên lý.

Cần điểm qua ngắn gọn nguyên lý mà phương pháp này dùng để thay thế tích tensor. Thay cho phép ghép cặp nặng, nó dựng hệ tọa độ cục bộ. Tức là đặt một khung hướng chuẩn quanh mỗi nguyên tử. Nếu xử lý thông tin trên khung này, ta có thể giữ hướng bằng phép tính nhẹ. Đối xứng vẫn được giữ, chỉ giảm gánh nặng.

Nghiên cứu này nêu hai kết quả. Một là tốc độ. Bản tóm tắt công khai cho biết nó nhanh hơn ba đến năm lần so với các mô hình có độ chính xác tương tự. Hai là phạm vi độ chính xác. Nó đạt độ chính xác có thể so với các mô hình đi trước trong dự đoán độ bền vật liệu, dẫn nhiệt, dao động mạng tinh thể và động lực học phân tử ở thang nano giây. Có thể xem là đã tăng tốc mạnh mà vẫn giữ độ chính xác. Nghiên cứu này là nghiên cứu giai đoạn đầu chưa qua bình duyệt. Vì vậy, nên xem nó là một công nghệ ứng viên nhiều triển vọng, không phải tiêu chuẩn đã được xác lập.

Loại bỏ tích tensor không chỉ cải thiện tốc độ. Nó cũng giúp bề mặt trơn hơn. Nếu cắt giảm độ chính xác hướng trong tích tensor, trên bề mặt có thể xuất hiện các gợn sóng nhỏ. Những gợn này làm phép tính lực bị giật. Thiết kế không dùng tích tensor được tạo ra để tránh các gợn như vậy. Đường cong vẫn mềm mại cả ở đoạn liên kết hình thành và đứt gãy.

Độ trơn này quan trọng trong tính toán vật liệu tên lửa. Ở nhiệt độ cao, nguyên tử dao động mạnh và liên kết thường xuyên bị đứt. Nếu đường cong thô, lực sẽ bị giật ở những đoạn này và mô phỏng sẽ sụp

đổ. Đường cong trơn giúp phép tính ổn định ngay cả trong điều kiện cực hạn. Vì vậy, độ trơn là một tính chất quan trọng ngang với độ chính xác.

Cũng có nghiên cứu khác có mục tiêu tương tự. Một nghiên cứu trải thông tin hướng lên lưới cầu rồi trộn bằng mạng nơ-ron nhẹ. Cách này cũng nhanh hơn tích tensor mà vẫn giữ năng lực biểu diễn. Cũng đã xuất hiện nghiên cứu thay đổi mềm khoảng cách cắt theo tình huống. Tất cả các nghiên cứu này đều ở giai đoạn đầu. Nhiều nhóm đang thử những con đường khác nhau hướng đến cùng một mục tiêu.

Hãy xem thêm vì sao độ trơn lại quan trọng đến thế. Lực xuất phát từ độ dốc của bề mặt năng lượng. Nếu bề mặt trơn, độ dốc cũng thay đổi trơn. Nếu bề mặt có gợn nhỏ, độ dốc sẽ đột ngột giật. Lực bị giật đẩy nguyên tử theo hướng sai và phá hỏng mô phỏng. Vì vậy, cần bề mặt trơn không kém gì năng lượng chính xác. Dù sai số nhìn thấy bên ngoài nhỏ, nếu bề mặt thô thì vấn đề sẽ lớn dần trong phép tính dài.

Ý nghĩa của dòng nghiên cứu này rất rõ. Đó là giữ độ chính xác của mô hình hiệp biến trong khi giảm gánh nặng tính toán. Khi phép tính nhẹ hơn, ta có thể quan sát hệ lớn hơn trong thời gian dài hơn. Điều này giúp ngay cho việc theo dõi lâu các vật liệu lớn như linh kiện tên lửa. Tuy nhiên, bằng chứng thực nghiệm vẫn đang được tích lũy, nên phải xem cùng với kiểm chứng.

## 6. Mô hình hóa mật độ electron và cắt tỉa có cấu trúc để chạy trên thiết bị nhẹ

Mô hình lớn chính xác, nhưng nặng và đòi hỏi thiết bị tính toán lớn. Mục này bàn về cách làm mô hình lớn nhẹ hơn mà vẫn giữ định luật vật lý. Ta sẽ xem hai ý tưởng. Một là cắt tỉa, tức cắt bỏ các đường tính toán ít quan trọng hơn. Hai là đưa manh mối về phân bố electron vào mô hình. Trọng tâm của mục này là làm cho nhẹ, nhưng không cắt tùy tiện. Tính toán nguyên tử có nhiều tính chất vật lý phải giữ. Đó là đối xứng, hướng của lực, và sự nhất quán giữa năng lượng và lực. Mục này giải thích cách thu nhỏ mô hình trong khi vẫn giữ các tính chất đó.

Cắt tỉa (Pruning) là phương pháp giảm các đường ít quan trọng hơn trong mạng nơ-ron. Các đường mang thông tin bên trong mạng nơ-ron được gọi là kênh. Không phải mọi kênh đều quan trọng như nhau. Nếu giảm các kênh ít dùng, phép tính sẽ nhẹ hơn. Nó giống như bỏ bớt đồ ít dùng khỏi túi để chiếc túi nhẹ đi.

Vấn đề là nếu cắt bất kỳ kênh nào, định luật vật lý sẽ lung lay. Trong mô hình hiệp biến, kênh gắn với hướng. Nếu cắt rời rạc các kênh thuộc cùng một nhóm hướng, hướng của lực sẽ trở nên kỳ lạ. Ở đây cần sửa một hiểu lầm. Nếu mô hình trước hết xuất ra năng lượng rồi lấy lực bằng độ dốc của năng lượng đó, thì ngay cả khi giảm kênh, quan hệ giữa năng lượng và lực vẫn được giữ. Cắt tỉa không lập tức phá vỡ quan hệ này. Thay vào đó, độ chính xác trong xử lý hướng và độ chính xác dự đoán có thể xấu đi. Vì vậy, phải cắt tỉa thận trọng.

### 6.1 Cắt tỉa có cấu trúc giữ tính nhất quán vật lý

Cắt tỉa có cấu trúc (Structured Pruning) là phương pháp cắt kênh theo đơn vị nhóm hướng. Không cắt rời từng kênh, mà xử lý cả nhóm. Làm vậy thì cân bằng hướng mới không bị phá vỡ. Tỷ lệ giữa kênh vô hướng và kênh vector cũng được quản lý để giữ nguyên. Vô hướng là giá trị không có hướng, còn vector là giá trị có hướng. Phải giữ tỷ lệ này thì đối xứng và hướng của lực mới nguyên vẹn.

Hãy hình dung vì sao phải cắt theo cả nhóm. Kim la bàn chỉ có ý nghĩa khi nó chỉ cùng lúc các hướng đông, tây, nam, bắc. Nếu chỉ để lại nam bắc và cắt bỏ đông tây, hướng sẽ méo đi. Kênh hướng của mô hình hiệp biến cũng như vậy. Các giá trị trong cùng một nhóm phải đi cùng nhau thì mới chỉ hướng đúng. Nếu chỉ cắt rời một phần trong đó, hướng của lực sẽ lệch. Vì vậy, khi cắt, bắt buộc phải lấy nhóm làm đơn vị.

Việc cắt kênh nào được quyết định theo mức độ quan trọng. Ta đo mỗi kênh đóng góp bao nhiêu vào kết quả. Các kênh đóng góp nhỏ sẽ được giảm trước. Khi đó cũng lấy nhóm hướng làm đơn vị. Ta xem và

đánh giá các giá trị trong một nhóm cùng nhau. Làm như vậy sẽ giữ bộ khung mang thông tin và chỉ bỏ phần thừa.

Gần đây, một nhóm nghiên cứu đã áp dụng phương pháp này cho thế năng nền tảng. Tiêu đề của nghiên cứu này là mô hình nguyên tử hiệp biến SO3 đặc được tạo bằng cắt tỉa có cấu trúc. Bản tóm tắt công khai cho biết họ đã giảm số tham số nhiều lần. Tham số là số lượng giá trị mà mô hình học. Khối lượng tính toán cần cho tiền huấn luyện cũng giảm nhiều lần. Đồng thời, sau tinh chỉnh, sai số năng lượng và lực lại giảm. Nghiên cứu này là nghiên cứu giai đoạn đầu được công bố vào năm 2026. Vì chưa qua bình duyệt, không nên xem đây là kết quả đã được xác lập.

Mục tiêu mà nghiên cứu này nhắm tới là chạy tính toán chính xác ngay cả trên thiết bị nhẹ. Mô hình lớn thường đòi hỏi nhiều thiết bị hiệu năng cao. Nếu thu nhỏ mô hình đúng cách, thiết bị nhỏ hơn cũng có thể thực hiện phép tính tương tự. Khi ngưỡng tham gia nghiên cứu thấp hơn, nhiều nhà nghiên cứu hơn có thể tham gia tính toán vật liệu. Tuy nhiên, thiết bị nào có thể đáp ứng đến mức nào vẫn cần được kiểm tra thêm.

Trong nghiên cứu vật liệu tên lửa, hướng này cũng có ý nghĩa. Để mô phỏng lâu môi trường cực hạn, phải lặp lại phép tính nhiều lần. Nếu mô hình nhẹ, cùng một thiết bị có thể chạy nhiều thí nghiệm hơn. Điều này giúp sàng lọc nhanh vật liệu ứng viên. Trọng tâm là đạt tốc độ trong khi vẫn giữ độ chính xác.

Cũng có phương pháp khác có mục tiêu giống cắt tỉa. Đó là chuyển tri thức của mô hình lớn sang mô hình nhỏ. Phương pháp này được gọi là chưng cất tri thức. Mô hình lớn đã học tốt trở thành thầy của mô hình nhỏ. Mô hình nhỏ học bằng cách bắt chước câu trả lời của thầy. Khi đó, dù có thân hình nhỏ, nó vẫn đạt năng lực gần với mô hình lớn. Cắt tỉa và chưng cất tri thức đều là con đường chuyển mô hình lớn sang dạng nhẹ.

Điểm cần chú ý là không được bỏ qua kiểm chứng vật lý. Mô hình nhẹ có được nhờ cắt tỉa cũng bắt buộc phải được kiểm tra. Cần kiểm tra cùng lúc tính nhất quán giữa năng lượng và lực, hướng của lực, đối xứng và chỉ báo độ bất định. Khi các tính chất này được giữ, mô hình nhẹ mới đáng tin. Mô hình chỉ nhanh thì nguy hiểm trong tính toán cực hạn.

## 6.2 Phương pháp thêm mạnh mỗi mật độ electron

Khi thu nhỏ mô hình, năng lực biểu diễn cũng có nguy cơ giảm theo. Một ý tưởng nhằm bù đắp mất mát này là đưa mật độ electron vào. Mật độ electron là bản đồ cho biết electron phân bố trong không gian nhiều đến mức nào. Vốn dĩ đây là đại lượng vật lý cơ bản mà DFT xử lý. Ý tưởng là đưa trực tiếp mạnh mỗi này vào trong mô hình để bổ sung năng lực biểu diễn còn thiếu. Cách trộn các thông tin khác nhau như vậy được gọi là hybrid.

Ở đây cần điểm qua ngắn gọn các loại liên kết. Liên kết ion là liên kết hình thành bằng cách cho và nhận electron. Muối được tạo nên bằng loại liên kết này. Liên kết kim loại là liên kết trong đó electron chảy tự do giữa nhiều nguyên tử. Nhờ liên kết này, sắt dẫn điện tốt. Liên kết cộng hóa trị là liên kết trong đó các nguyên tử chia sẻ electron với nhau. Như vậy, mỗi loại liên kết có hình dạng electron khác nhau.

Thông tin mật độ electron giúp phân biệt loại liên kết. Trong vật liệu có nhiều loại liên kết. Liên kết ion, liên kết kim loại và liên kết cộng hóa trị khác nhau. Nếu nhìn electron tập trung ở đâu và nhiều bao nhiêu, ta có thể biết sự khác biệt này. Khi đưa mạnh mỗi này vào mô hình, mô hình có thể nắm bắt tính chất liên kết tốt hơn.

Cách này cũng được kỳ vọng là hữu ích trong việc khôi phục các hiệu ứng tác động xa. Ảnh hưởng của electron lan ra vượt quá các láng giềng gần. Khi thu nhỏ mô hình cho nhẹ, ta dễ bỏ sót các hiệu ứng xa này. Kỳ vọng là nếu dùng tốt mạnh mỗi mật độ electron, mô hình có thể chứa phần nào ảnh hưởng xa đó. Tuy nhiên, việc đưa mật độ electron vào không tự động giải quyết tĩnh điện và lực phân tán ở

khoảng cách xa. Việc đưa loại mật độ nào vào, theo cách nào, và kiểm chứng ra sao vẫn đang được nghiên cứu thêm. Vì vậy, nên hiểu hướng này là một đề xuất, không phải lời giải đã xác lập.

Vật liệu tên lửa thường có nhiều loại liên kết trộn phức tạp. Kim loại và gốm được dùng cùng nhau, và màng oxide hình thành. Để xử lý các vật liệu như vậy, phải phân biệt chính xác tính chất của liên kết. Cách đưa mật độ electron vào có thể hỗ trợ sự phân biệt này và góp sức cho tính toán vật liệu cực hạn.

Khả năng tính toán quy mô lớn cũng mở ra cùng lúc. Gần đây, một nhóm nghiên cứu đã mô phỏng hợp kim trộn nhiều nguyên tố ở quy mô một triệu nguyên tử. Họ xử lý hệ lớn bằng cách dùng thế năng liên nguyên tử cho nhiều thành phần. Nghiên cứu này là nghiên cứu giai đoạn đầu được công bố vào năm 2026. Hợp kim entropy cao là ứng viên quan trọng cho vật liệu nhiệt độ cao dùng trong tên lửa. Càng xử lý được hệ lớn, ta càng có thể thấy hành vi gần với linh kiện thật.

Hãy nhìn câu chuyện của mục này qua thực tế phát triển tên lửa. Không phải công ty và viện nghiên cứu tên lửa nào cũng có tài nguyên tính toán dồi dào. Nếu mô hình lớn chỉ chạy được trên nhiều thiết bị, ngưỡng tham gia sẽ cao. Khi mô hình nhẹ hơn, nhiều nhóm hơn có thể bắt tay vào tính toán nguyên tử. Họ sàng lọc nhanh vật liệu ứng viên và giảm số đối tượng phải thử nghiệm thật. Vòng lặp qua lại giữa tính toán và thí nghiệm này rút ngắn thời gian phát triển.

Hai ý tưởng trong mục này hướng đến cùng một mục tiêu. Đó là làm phép tính nhẹ hơn trong khi vẫn giữ độ chính xác. Cắt tỉa bỏ phần thừa, còn mạnh mỗi mật độ electron lấp phần năng lực biểu diễn còn thiếu. Nếu dùng cả hai cùng nhau, ta có thể tạo mô hình đủ dùng ngay cả trên thiết bị nhỏ. Hướng này vẫn đang phát triển và đang ở giai đoạn tích lũy kiểm chứng.

## Tổng kết

Chương này là câu chuyện về các quy tắc dùng máy tính để xử lý nguyên tử. Quy tắc đó chính là thế năng liên nguyên tử. Từ lâu, các nhà khoa học đã phải chọn một trong hai con đường: phương pháp chính xác hoặc phương pháp nhanh. DFT chính xác nhưng chậm, còn trường lực kinh nghiệm nhanh nhưng thiếu chính xác. Thế năng liên nguyên tử học máy là công cụ nhằm phá bỏ bức tường giữa hai con đường đó.

Ta hãy nhìn lại ngắn gọn hành trình này. Ban đầu, con người biến môi trường xung quanh nguyên tử thành các con số. Sau đó, mô hình tự học các con số ấy. Tiếp nữa, người ta để các nguyên tử trao đổi thông tin với nhau. Giờ đây, đã xuất hiện cả những mô hình lớn được học trước trên phạm vi vật liệu rộng. Qua mỗi bước, bàn tay con người bớt đi, còn sức mạnh của dữ liệu tăng lên. Dòng chảy này là mạch phát triển lớn của MLIP.

MLIP giai đoạn đầu biến hình dạng xung quanh nguyên tử thành các con số do con người thiết kế. Cách dùng bộ mô tả này có BPNN, thế năng xấp xỉ Gaussian và DeePMD. Chúng đạt được thành quả lớn, nhưng cũng bộc lộ giới hạn. Khi số nguyên tố tăng lên, gánh nặng tính toán lớn hơn và khó đưa vào các hiệu ứng ở khoảng cách xa.

Thế hệ tiếp theo là mạng nơ-ron đồ thị đồng biến. Phương pháp này xem nguyên tử là điểm, quan hệ láng giềng là đường nối, rồi cho chúng trao đổi thông tin. Nó đưa các đối xứng phù hợp với phép quay và phép tịnh tiến vào trong mô hình, nhờ đó tiết kiệm dữ liệu. Từ SchNet đến NequIP rồi MACE, độ chính xác và hiệu quả cùng tăng lên. Các thế năng nền tảng như CHGNet, MatterSim và GRACE đã trở thành những mô hình lớn được học trước trên phạm vi vật liệu rộng.

Ngay cả mô hình chính xác cũng có điểm yếu. Trong tình huống lạ mà nó chưa học, mô hình có thể sai. Vì vậy cần học chủ động và định lượng bất định, để mô hình tự báo khi gặp điều mình không biết. Môi trường tên lửa luôn cực hạn và có nhiều sự kiện xa lạ. Phải có thiết bị an toàn này thì ta mới có thể tin vào tính toán vật liệu cực hạn.

Các nghiên cứu gần đây theo đuổi đồng thời tốc độ, độ trơn và sự gọn nhẹ. Chúng giảm hoặc loại bỏ tích tensor để tính nhanh hơn. Chúng thu nhỏ mô hình bằng cắt tỉa có cấu trúc và bổ sung khả năng biểu diễn bằng manh mối từ mật độ electron. Nhiều phương pháp kiểu này vẫn là nghiên cứu giai đoạn đầu, chưa qua bình duyệt. Chúng có triển vọng, nhưng còn sớm để xem là tiêu chuẩn đã được xác lập.

Có năm điểm mà độc giả phổ thông nên nhớ sau khi đọc chương này. Tính toán nguyên tử từ lâu đã có sự đánh đổi giữa chính xác và nhanh. Học máy đã bắc một cây cầu để giảm sự đánh đổi ấy. Giữ đối xứng là cốt lõi của một mô hình tốt. Dù mô hình tốt đến đâu, nó vẫn phải tự báo khi gặp điều mình không biết thì mới an toàn. Và phần lớn công nghệ mới nhất vẫn là nghiên cứu giai đoạn đầu đang được kiểm chứng.

Năm điểm này cũng đúng với các chương phía sau. Việc tạo cấu trúc tinh thể, thiết kế hợp kim và xử lý gốm đều đứng trên nền tính toán nguyên tử này. Nếu không có tính toán nguyên tử chính xác, thiết kế bên trên cũng lung lay. Vì vậy nội dung của chương này trở thành viên đá nền của cả cuốn sách.

Ta cũng cần giữ một cái nhìn cân bằng. MLIP là công cụ mạnh, nhưng không phải vạn năng. Nếu chất lượng dữ liệu học kém, dự đoán cũng kém. Nếu lơ là kiểm chứng, ta sẽ bị những sai số có vẻ hợp lý đánh lừa. Vì vậy công nghệ này không thay thế phán đoán của con người và thí nghiệm. Tính toán thu hẹp danh sách ứng viên, còn thí nghiệm đảm nhận xác nhận cuối cùng. Sự phối hợp này là con đường đưa vật liệu mới ra thế giới một cách an toàn.

Mục tiêu của tất cả công nghệ này quy tụ về một điểm. Đó là thử nghiệm vật liệu tên lửa và vật liệu mới trong máy tính ở quy mô lớn và thời gian dài, một cách chính xác và an toàn. Nếu sàng lọc ứng viên trước khi chế tạo linh kiện thật, ta tiết kiệm được tiền và thời gian. Từ chương sau, câu chuyện sẽ chuyển sang việc trực tiếp tạo ra vật liệu mới trên nền tính toán nguyên tử đã được xây chắc như vậy.

## Tài liệu căn cứ liên quan

- 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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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>

## Chương 2. Dự đoán cấu trúc tinh thể (CSP) và thiết kế ngược vật liệu dựa trên AI tạo sinh và mô hình khuếch tán (Diffusion Models)

### Mở đầu

Khi làm ra một đồ vật, chúng ta thường chọn vật liệu trước. Giả sử ta làm một cái nồi. Trước hết ta chọn sắt hoặc nhôm. Sau đó ta định độ dày và hình dạng. Cuối cùng ta đặt lên lửa và kiểm tra xem nó có tính chất gì.

Cách cũ để tìm vật liệu mới cũng gần giống như vậy. Trước hết ta chọn một chất đã biết. Sau đó ta đo tính chất bằng máy tính hoặc thí nghiệm. Nếu tốt thì giữ lại, nếu xấu thì bỏ đi. Cách này đáng tin. Nhưng phải kiểm tra từng thứ một nên tốc độ chậm.

Vấn đề nằm ở chỗ thế giới vật chất quá rộng. Có nhiều loại nguyên tố. Số lượng nguyên tử cũng có thể thay đổi. Vị trí đặt nguyên tử và kích thước mạng tinh thể cũng có thể thay đổi. Khi kết hợp tất cả những thứ này, số trường hợp tăng lên như sao trên trời. Con người không thể thử từng trường hợp một cho đến hết.

Chương này kể câu chuyện đảo ngược hướng suy nghĩ. Ta không chọn vật liệu trước. Ta xác định tính chất mong muốn trước. Ta xác định rằng mình muốn một chất điện giải rắn dẫn điện tốt. Ta xác định rằng mình muốn một chất xúc tác chịu được trong vòi phun tên lửa nóng. Khi đó trí tuệ nhân tạo tự vẽ ra cấu trúc tinh thể phù hợp với tính chất đó. Cách làm này gọi là thiết kế ngược vật liệu (Inverse Design).

Có thể so sánh bằng nấu ăn để thấy vì sao thiết kế ngược quan trọng. Cách cũ là nấu món ăn bằng nguyên liệu có trong tủ lạnh. Nguyên liệu đã được định sẵn nên món ăn có thể làm cũng bị định sẵn. Thiết kế ngược là xác định trước món mình muốn ăn. Sau đó ta tìm ngược lại những nguyên liệu cần cho món đó. Nguyên liệu chưa có trong tủ lạnh cũng được tìm mới. Làm như vậy, ta không bị nhốt trong những nguyên liệu có sẵn trong tủ lạnh mà tiến gần hơn đến món ăn mình tưởng tượng.

Nhân vật chính ở đây là trí tuệ nhân tạo tạo sinh (Generative AI). Có thể hiểu dễ hơn nếu nghĩ đến AI vẽ tranh. Nó bắt đầu từ một đám điểm mờ. Nó xóa nhiều từng chút một. Cuối cùng, khuôn mặt người hoặc phong cảnh hiện ra. Thiết kế ngược vật liệu cũng có nguyên lý như vậy. Nó bắt đầu từ một đám nguyên tử mờ. Nó xóa nhiều và tạo thành tinh thể có trật tự. Cách xóa nhiều này được gọi là mô hình khuếch tán (Diffusion Model).

Trong chương này, ta sẽ lần lượt xem mô hình khuếch tán tạo ra tinh thể như thế nào. Ta sẽ xem cách nó tuân thủ quy tắc đối xứng. Ta sẽ xem kỹ thuật khớp dòng (Flow Matching) giúp tính toán nhanh hơn. Ta cũng sẽ xem những ý tưởng mới như inpainting để lấp chỗ trống và nguyên tử mirage. Ta cũng sẽ xem cách học tăng cường dẫn mô hình về phía cấu trúc tốt. Cuối cùng, ta sẽ xem các trường hợp phương pháp này đã được dùng trong thực tế cho chất điện giải rắn của pin, chất xúc tác và hợp kim hàng không vũ trụ. Ta sẽ lần lượt giải thích vì sao tên lửa và vật liệu mới cần công nghệ này.

### 1. Từ cách chọn vật liệu trước sang thiết kế ngược xác định tính chất trước

Phần này bàn về hướng tìm kiếm vật liệu đã thay đổi như thế nào. Cách cũ chọn vật liệu trước rồi kiểm tra tính chất sau. Cách mới xác định tính chất trước rồi tạo vật liệu sau. Ta sẽ xem vì sao cần chuyển đổi như vậy, và điều đó có ý nghĩa gì đối với tên lửa và vật liệu mới.

Cách cũ được gọi là sàng lọc tính chất theo chiều thuận (Forward Property Screening). Chiều thuận có nghĩa là đi từ vật liệu đến tính chất. Sàng lọc là việc lọc ra các ứng viên từ nhiều lựa chọn. Nhà nghiên cứu lấy ứng viên từ cơ sở dữ liệu tinh thể đã biết. Có hai kho tiêu biểu. Đó là Materials Project và Cơ sở dữ liệu cấu trúc tinh thể vô cơ (ICSD).

Sau khi lấy ứng viên ra, ta tính tính chất. Lúc này ta dùng phép tính vật lý gọi là lý thuyết phiếm hàm mật độ (Density Functional Theory, DFT). Phép tính này giải chuyển động của electron trong nguyên tử để tìm năng lượng và tính chất. Kết quả chính xác. Nhưng khối lượng tính toán lớn. Máy tính phải chạy lâu. Vì vậy số ứng viên có thể xem trong một ngày rất ít.

Giới hạn căn bản của cách này nằm ở chỗ nó bị nhốt trong một cái chuồng hẹp. Việc kiểm tra chỉ diễn ra với các chất đã biết. Vật liệu mới không có trong cơ sở dữ liệu thì ngay từ đầu không thể trở thành ứng viên. Dù có một vật liệu chưa biết với tính chất vượt trội, ta cũng không chạm tới được. Vì vậy cách cũ yếu ở khả năng tạo ra cái mới.

Thiết kế ngược vật liệu phá vỡ cái chuồng đó. Nó dựng lại thứ tự theo chiều ngược. Trước hết ta xác định tính chất mục tiêu. Ta xác định rằng độ dẫn ion phải cao. Ta xác định rằng vật liệu phải ổn định về nhiệt động lực học. Sau đó trí tuệ nhân tạo trực tiếp vẽ ra cấu trúc thỏa mãn các điều kiện đó. Nó cũng vẽ ra những cấu trúc chưa từng có trong cơ sở dữ liệu. Trọng tâm của tìm kiếm vật liệu chuyển từ kiểm tra sang sáng tạo.

Hãy lấy ví dụ để thấy số trường hợp lớn đến mức nào. Trong tự nhiên có khoảng 90 nguyên tố có thể dùng. Giả sử ta chỉ đặt mười nguyên tử trong một chiếc hộp. Ta chọn nguyên tố nào đặt vào từng vị trí. Vị trí của nguyên tử cũng có thể thay đổi từng chút một. Chỉ riêng tổ hợp này cũng đã thành một con số thiên văn. Số ứng viên sinh ra còn nhiều hơn cát trên bờ biển. Nếu con người tính từng cái một thì dùng cả đời cũng không xong.

Thời gian cho một phép tính cũng không nhỏ. Tính toán DFT càng mất lâu khi số nguyên tử càng tăng. Khi số nguyên tử tăng gấp đôi, phép tính tăng lên nhiều lần. Để giải một cấu trúc lớn, máy tính chạy trong nhiều ngày. Vì vậy không thể cứ tính thật nhiều ứng viên một cách bừa bãi. Phải chọn trước ứng viên thật tốt. Trí tuệ nhân tạo đem lại trí khôn để chọn ứng viên nào nên tính trước.

Vẽ một tinh thể là việc xác định đồng thời ba thứ. Trước hết là xác định mạng tinh thể (Lattice). Mạng tinh thể là hình dạng và kích thước của chiếc hộp chứa các nguyên tử. Tiếp theo là xác định vị trí nguyên tử. Đó là nguyên tử nằm ở đâu trong chiếc hộp. Cuối cùng là xác định loại nguyên tử. Đó là nguyên tố nào ngồi ở vị trí đó. Ba thứ phải khớp với nhau thì mới thành tinh thể thật. Chỉ cần một thứ lệch đi, nguyên tử sẽ va vào nhau hoặc cấu trúc sẽ sụp đổ.

Khó xử lý đồng thời cả ba thứ này. Mạng tinh thể là giá trị liên tục có chiều dài và góc. Vị trí nguyên tử cũng là giá trị liên tục. Nhưng loại nguyên tử là giá trị rời rạc. Sau clo không có nguyên tố nào giống nửa clo. Như vậy phải xử lý cùng lúc các giá trị có tính chất khác nhau. Mô hình khuếch tán làm được việc khó này. Nó khôi phục song song các giá trị liên tục và giá trị rời rạc từ nhiễu.

Ta hãy giải thích thêm một chút nguyên lý của mô hình khuếch tán. Quá trình học có hai hướng. Trước hết, ở chiều thuận, nhiễu được thêm từng chút một vào tinh thể sạch. Sau nhiều lần thêm, tinh thể biến thành một đám điểm hỗn loạn. Trí tuệ nhân tạo học cách đảo ngược quá trình này. Nó luyện tập đoán tinh thể ban đầu từ trạng thái có nhiễu. Sau khi luyện tập xong, nó tạo ra tinh thể mới từ các điểm hỗn loạn.

Sức mạnh của cách này nằm ở chỗ nó học dữ liệu tốt. Ta cho mô hình xem hàng chục nghìn tinh thể đã được biết đến trên thế giới. Mô hình học cảm giác về cách sắp xếp nào có vẻ hợp lý. Nó học khoảng cách mà các nguyên tử phải giữ với nhau. Nó học nguyên tố nào ngồi tốt ở vị trí nào. Bằng cảm giác đã học như vậy, nó đề xuất tinh thể mới. Con người không cần viết ra từng quy tắc, mô hình tự học.

Thiết kế ngược không chỉ dừng lại ở tinh thể có trật tự. Cũng có những vật liệu trong đó nguyên tử rối loạn như thủy tinh. Những vật liệu này được gọi là vô định hình (Amorphous). Một nghiên cứu công bố năm 2025 đã tạo cấu trúc vô định hình bằng mô hình khuếch tán. Nó tạo nhanh hơn phương pháp tính toán cũ hàng nghìn lần. Nghiên cứu đó tạo ra nhiều cấu trúc khác nhau bằng cách thay đổi thành phần và điều kiện quy trình. Có thể nói bàn tay của thiết kế ngược đã vươn ra ngoài tinh thể.

Mô hình thiết kế ngược cũng học các quy tắc hóa học. Mỗi nguyên tử có số liên kết có thể tạo với láng giềng. Con số này gọi là hóa trị (Valence). Nếu vi phạm hóa trị, cấu trúc tạo ra sẽ không thể tồn tại. Một nghiên cứu công bố năm 2026 đã đưa quy tắc hóa trị này vào quá trình tạo sinh. Kết quả là nó chỉ lọc ra những vật liệu thực sự có thể tồn tại. Xu hướng cấy kiến thức thường thức của vật lý và hóa học vào mô hình đang tiếp diễn.

Ý nghĩa của sự chuyển đổi này đối với tên lửa và vật liệu mới rất lớn. Tên lửa cần vật liệu chịu được môi trường cực hạn. Buồng đốt nóng. Vòi phun bị dòng khí nhanh bào mòn. Pin phải nhẹ và an toàn. Những vật liệu thỏa mãn tất cả các điều kiện như vậy không phổ biến trong tự nhiên. Vì vậy cần thiết kế ngược để vẽ ra thứ chưa có. Trí tuệ nhân tạo thay con người bước vào không gian vật liệu mà con người chưa từng đi tới.

## 2. Mô hình tạo sinh tuân thủ quy tắc đối xứng của tinh thể

Phần này bàn về quy tắc đối xứng của tinh thể. Tinh thể không phải là đồng nguyên tử bị ném lung tung. Nó là trật tự trong đó nguyên tử lặp lại theo quy tắc. Nếu trí tuệ nhân tạo không biết quy tắc này, nó sẽ vẽ ra tinh thể giả. Ta sẽ chia thành hai thế hệ mô hình để xem đối xứng được đưa vào mô hình như thế nào.

Để hiểu tinh thể, trước hết phải nắm được từ đối xứng. Đối xứng là tính chất mà hình dạng vẫn giữ nguyên dù ta thực hiện một thao tác nào đó. Khi gấp tờ giấy làm đôi mà hai bên chồng khít lên nhau, đó là đối xứng trái phải. Trong tinh thể, nhiều lớp đối xứng như vậy xếp chồng lên nhau. Dịch sang bên cũng giống nhau, xoay cũng giống nhau, soi gương cũng giống nhau.

Tập hợp và phân loại tất cả các quy tắc đối xứng như vậy được gọi là nhóm không gian (Space Group). Trong tinh thể 3 chiều có đúng 230 nhóm không gian. Một tinh thể thuộc một trong 230 nhóm này. Phân loại này phù hợp với tinh thể tuần hoàn 3 chiều lý tưởng. Vật liệu vô định hình như thủy tinh hoặc giả tinh thể nằm ngoài khung này. Nếu biết nhóm không gian, coi như ta biết toàn bộ quy tắc lặp lại và quay của tinh thể đó.

Trong nhóm không gian có những vị trí cố định mà nguyên tử có thể ngồi. Tập hợp các vị trí này được gọi là vị trí Wyckoff (Wyckoff Position). Vị trí Wyckoff không phải là một vị trí riêng lẻ, mà là một nhóm vị trí được liên kết bằng đối xứng. Khi đặt một nguyên tử vào một vị trí, quy tắc đối xứng cũng đặt nguyên tử giống vậy vào vị trí ghép cặp. Nó giống như đặt một quả bóng trước gương thì trong gương cũng xuất hiện một quả bóng. Vì vậy khi xác định một vị trí, vị trí của nhiều nguyên tử được xác định cùng lúc. Số vị trí đi kèm với nhóm đó được gọi là bội số (Multiplicity). Khi mô hình tạo sinh nhận quy tắc này, nó có thể tránh việc nguyên tử chồng lên nhau.

Sức mạnh của nhóm không gian nằm ở việc vẽ toàn bộ bằng ít thông tin. Với tinh thể có đối xứng cao, chỉ cần xác định vài nguyên tử. Phần còn lại được quy tắc đối xứng tự điền vào. Có thể hiểu dễ hơn nếu nghĩ đến tinh thể tuyết. Hình dạng vươn ra sáu hướng giống hệt nhau. Chỉ cần vẽ một hướng, năm hướng còn lại thu được bằng cách xoay. Tinh thể cũng xác định toàn bộ theo cách như vậy từ một mảnh nhỏ.

Nếu mô hình tạo sinh không biết quy tắc này, vấn đề sẽ xảy ra. Giả sử trí tuệ nhân tạo đặt từng nguyên tử một cách tự do. Chỉ cần một chút nhiễu tính toán nhỏ, đối xứng cũng bị phá vỡ. Tinh thể gần như

không có đối xứng được gọi là P1. P1 cũng là một nhóm không gian chính thức. Trên thực tế cũng có vật liệu thuộc P1. Vì vậy không thể chỉ vì là P1 mà gọi là giả. Vấn đề là khi tinh thể vốn phải có đối xứng cao lại sụp xuống P1. Khi đó phải kiểm tra riêng khoảng cách giữa các nguyên tử, năng lượng và độ ổn định.

Vì vậy các nghiên cứu mới cố đưa sẵn quy tắc đối xứng vào trong mô hình. Thay vì đặt từng nguyên tử một, chúng đi theo một thứ tự khác. Trước hết tạo ra một đơn vị cơ bản nhỏ chứa toàn bộ đối xứng. Sau đó áp dụng quy tắc đối xứng để mở rộng thành toàn bộ tinh thể. Làm như vậy, sản phẩm tạo ra có khả năng tuân thủ đối xứng cao hơn.

Một cách tốt để xử lý đối xứng là tính đồng biến (Equivariance). Đồng biến là tính chất mà nếu đầu vào xoay thì đầu ra cũng xoay theo. Có thể hiểu dễ hơn bằng ví dụ chiếc ghế tay. Khi lộn ghế tay phải, nó trở thành ghế tay trái. Bàn tay và ghế tay cùng bị lộn. Nếu mô hình tinh thể có tính đồng biến, dù xoay cấu trúc thì tính chất vẫn giữ nguyên. Đối xứng tự được giữ, không cần ép khớp một cách gượng gạo.

Khi dùng mạng nơ-ron có tính đồng biến này, việc học trở nên dễ hơn. Vì mô hình tự biết đối xứng. Không cần cho nó xem cùng một cấu trúc ở nhiều góc khác nhau. Dữ liệu ít vẫn học tốt. Vì vậy các mô hình tạo sinh tinh thể mới xem đồng biến là một công cụ quan trọng. Tính chất này cũng sẽ trở thành lõi trong mô hình khớp dòng được nói đến sau.

Một bài tổng quan về mô hình tạo sinh tinh thể công bố năm 2026 đã sắp xếp tốt xu hướng này. Bài tổng quan chỉ ra rằng năng lực tuân thủ đối xứng là điều kiện cốt lõi của một mô hình tạo sinh tinh thể tốt. Nó xem các mô hình tuân thủ đối xứng là mô hình tạo ra nhiều ứng viên ổn định hơn. Bài tổng quan này so sánh song song mô hình khuếch tán và mô hình khớp dòng.

## 2.1 Điểm xuất phát của khuếch tán tạo sinh tinh thể (CDVAE)

Nghiên cứu khuếch tán tạo sinh tinh thể bắt đầu từ một mô hình ban đầu. Tên của mô hình này là CDVAE. Ta sẽ lần lượt xem nó đã làm tốt điều gì và còn thiếu điều gì.

CDVAE là viết tắt của Crystal Diffusion Variational Autoencoder, tức bộ tự mã hóa biến phân khuếch tán tinh thể. Mô hình này được công bố năm 2022 và mở cửa cho lĩnh vực này. Nó thiết lập khung khuếch tán đầu tiên xử lý cùng lúc mạng tinh thể, vị trí nguyên tử và loại nguyên tử.

Công lớn của CDVAE là đã tính đến điều kiện biên tuần hoàn (Periodic Boundary Condition). Điều kiện biên tuần hoàn có nghĩa là tinh thể lặp lại vô tận với cùng một hình dạng. Có thể hiểu dễ hơn nếu nghĩ đến hoa văn giấy dán tường. Một mảnh nối tiếp lên xuống trái phải và phủ kín cả bức tường. Tinh thể cũng do những đơn vị nhỏ lặp lại và tạo thành một khối.

Mô hình này dịch chuyển vị trí nguyên tử từng chút một bằng phương pháp gọi là động lực học Langevin (Langevin Dynamics). Khi đó, hướng đẩy nguyên tử là gradient của xác suất mà mô hình đã học. Nó khác với lực thực giữa các nguyên tử hoặc năng lượng tự do. Vì vậy, chỉ riêng quá trình này không bảo đảm tạo ra cấu trúc ổn định. Sau khi dùng mô hình sinh để dựng khung, ta lại tinh chỉnh cấu trúc bằng DFT. Cần nhìn hai bước này tách biệt với nhau.

Tuy nhiên, CDVAE có giới hạn rõ ràng. Nó không thể cố định trực tiếp đối xứng nhóm không gian. Vì nguyên tử được dịch chuyển tự do, đối xứng dễ bị phá vỡ. Kết quả là cấu trúc được tạo ra thường sụp về trạng thái P1, tức trạng thái mất đối xứng. Với các tinh thể bắt buộc phải có đối xứng, cần kiểm chứng riêng thêm.

Một giới hạn khác là phải định trước số nguyên tử. Mô hình này dự đoán số nguyên tử trước, rồi tạo cấu trúc theo số đó. Khi số lượng bị cố định, độ linh hoạt giảm xuống. Những tình huống nguyên tử bị thiếu hoặc được thêm vào trở nên khó xử lý. Giới hạn này về sau được giải bằng phương pháp nguyên tử ảo.

Những thiếu sót này không phải vết lỗi, mà là bậc thềm. Bài toán mà CDVAE đặt ra đã trở thành mục tiêu của thế hệ mô hình tiếp theo. Làm sao giữ được đối xứng. Làm sao xử lý số nguyên tử một cách tự do. Hai câu hỏi này đã dẫn dắt các nghiên cứu sau đó. Hiểu giới hạn của mô hình ban đầu là điểm xuất phát của phát triển.

## 2.2 Mô hình khuếch tán có điều kiện biết đối xứng

Thế hệ mô hình tiếp theo cố gắng giữ các quy tắc đối xứng ngay từ đầu. Tôi sẽ tách riêng hai nghiên cứu có tên gần giống nhau để tránh nhầm lẫn. Các con số về hiệu năng chỉ được xử lý thận trọng trong phạm vi đã được kiểm chứng.

Ý tưởng cốt lõi của mô hình biết đối xứng khá dễ hiểu. Nó không đặt từng nguyên tử một cách tự do. Trước hết, nó tạo ra đơn vị bất đối xứng (Asymmetric Unit). Đơn vị bất đối xứng là nhóm nguyên tử nhỏ nhất có thể tạo ra toàn bộ tinh thể bằng quy tắc đối xứng. Sau đó, các thao tác đối xứng của nhóm không gian được áp dụng lên hạt giống nhỏ này. Qua các thao tác đó, hạt giống mở rộng thành toàn bộ tinh thể.

Ưu điểm của cách này là đối xứng tự nhiên được giữ lại. Vì hạt giống đã được mở rộng bằng quy tắc đối xứng, kết quả có sẵn đối xứng bên trong. Không cần chỉnh từng nguyên tử riêng lẻ. Có thể hiểu dễ hơn nếu nghĩ đến gấp giấy. Khi gấp một lần rồi cắt, lúc mở ra hai bên trái phải giống hệt nhau. Mô hình nhận biết đối xứng cắt tinh thể theo cách đó.

Ở đây cần phân biệt hai nghiên cứu có tên gần giống nhau. Một là SymmCD. Đây là mô hình khuếch tán sinh tinh thể bảo toàn đối xứng. Levy và cộng sự công bố tại Hội nghị quốc tế về biểu diễn học (ICLR) năm 2025. Đây là bài báo chính thức đã qua bình duyệt. Mã định danh của bài báo là arXiv:2502.03638.

Nghiên cứu còn lại là một nghiên cứu riêng xử lý vị trí Wyckoff. Ishii và cộng sự công bố bản thảo trước bình duyệt vào năm 2026. Bản thảo trước bình duyệt (preprint) là nghiên cứu giai đoạn đầu chưa qua bình duyệt. Bài báo này đề xuất phương pháp WyckoffDiff-Adaptor. Mã định danh là arXiv:2601.08115. Hai bài báo có mục tiêu tương tự, nhưng phương pháp và tác giả khác nhau. Cuốn sách này không trộn hai nghiên cứu lại với nhau, mà giới thiệu riêng từng nghiên cứu.

Giữa CDVAE và mô hình nhận biết đối xứng còn có một cây cầu trung gian. Một trong số đó là DiffCSP. Mô hình này sinh đồng thời mạng tinh thể và tọa độ phân số bằng khuếch tán. Tọa độ phân số là vị trí tương đối khi xem kích thước hộp bằng 1. DiffCSP giữ đối xứng tốt hơn CDVAE. Nó mở đường cho thế hệ mô hình sau, nơi đối xứng được đưa trực tiếp vào. Cứ như vậy, các mô hình tiến lên từng thế hệ bằng cách lấp dần những thiếu sót.

Cần thận trọng khi xử lý các con số hiệu năng. Có con số nói rằng tỷ lệ bảo toàn đối xứng vượt 98%, nhưng con số này chưa được xác định chắc chắn trong nguyên văn. Vì vậy, cuốn sách này không dùng một con số cụ thể như hiệu năng đã được xác nhận. Thay vào đó, ta hiểu theo hướng phát triển. Mô hình nhận biết đối xứng giữ đối xứng tốt hơn mô hình cũ. Nhờ vậy, nó tạo ra nhiều ứng viên gần với tinh thể thực hơn.

Mô hình nhận biết đối xứng cũng xử lý tốt sinh có điều kiện. Sinh có điều kiện (Conditional Generation) là cách đưa tính chất mong muốn làm điều kiện. Giả sử ta đặt trước một giá trị gọi là bandgap. Bandgap là ngưỡng năng lượng chi phối tính chất của chất bán dẫn. Mô hình chỉ chọn và vẽ những tinh thể phù hợp với bandgap đó. Khi đặt điều kiện như vậy, chỉ các ứng viên hợp mục tiêu được gom lại. Nhờ đó, lãng phí do tạo ra ứng viên vô ích giảm đi đáng kể.

Một bản thảo trước bình duyệt năm 2026 đã đẩy sinh có điều kiện đi xa hơn. Nó sinh đồng thời cấu trúc tinh thể và tính chất điện tử. Vì học cả tính chất điện tử, mô hình khớp tính chất mục tiêu tốt hơn. Trong thử nghiệm dùng bandgap và năng lượng hình thành làm điều kiện, tỷ lệ thành công tăng lên. Tỷ lệ ứng viên ổn định và mới cũng cao hơn. Dòng nghiên cứu xử lý đồng thời đối xứng và điều kiện đang tiếp tục.

Cần nhìn xem vì sao công nghệ này cần cho vật liệu mới. Đối xứng gần rất sâu với tính chất vật chất. Cấu trúc giả bị mất đối xứng cũng làm dự đoán tính chất lệch đi. Hợp kim dùng trong hàng không vũ trụ hoặc vật liệu pin chỉ có ý nghĩa khi có thể đi tiếp đến tổng hợp thực tế. Mô hình sinh giữ đối xứng làm giảm ứng viên vô ích. Nhờ đó, chất lượng ứng viên chuyển sang thí nghiệm được nâng lên.

### 3. Khớp dòng chảy để tạo tinh thể nhanh hơn

Mô hình khuếch tán chính xác nhưng chậm vì cần nhiều bước. Khớp dòng chảy giảm số bước để tăng tốc. Ta cũng sẽ xem cách xử lý không gian cong nơi tinh thể tồn tại.

Trước hết, cần hiểu vì sao mô hình khuếch tán chậm. Mô hình khuếch tán tạo cấu trúc bằng cách xóa nhiều từng chút một. Mỗi lần chỉ xóa một ít. Vì vậy, để hoàn tất cần từ hàng trăm đến hàng nghìn bước. Nhiều bước thì tính toán lâu. Tạo một ứng viên cũng mất thời gian. Trong tìm kiếm vật liệu, nơi phải tạo nhiều ứng viên, thời gian này trở thành gánh nặng.

Khớp dòng chảy (Flow Matching) giải vấn đề này từ một góc khác. Thay vì xóa nhiều từng chút một, nó học đường tắt. Nó học con đường đi thẳng từ trạng thái đầu đến trạng thái mục tiêu. Có thể hiểu dễ hơn nếu nghĩ đến dòng sông. Khi dòng chảy đã được định hình, chiếc lá trôi theo dòng đó. Khớp dòng chảy học chính đường nước mà nước chảy qua.

Khi biểu diễn đường nước bằng công thức, ta có phương trình vi phân thường. Phương trình vi phân thường (Ordinary Differential Equation) là công thức ghi lại sự thay đổi của giá trị. Nó cho thấy giá trị thay đổi như thế nào theo thời gian. Khớp dòng chảy đi từ điểm đầu đến điểm cuối theo công thức này. Khuếch tán dao động ngẫu nhiên ở mỗi bước, nhưng đường đi này thì không. Khi điểm đầu đã được xác định, nó đi theo một đường đã định đến cuối. Vì vậy, đường đi thẳng hơn và cần ít bước hơn. Tuy nhiên, bản thân điểm đầu vẫn được lấy ngẫu nhiên. Việc đường đi đã được xác định không có nghĩa là méo lệch góc và thể tích của mạng tinh thể tự động giảm xuống.

Có thể so sánh khác biệt giữa khuếch tán và khớp dòng chảy với việc đi xuống núi. Khuếch tán dò dẫm từng bước trong sương mù. Cần thận nhưng chậm. Khớp dòng chảy đi xuống theo lối mòn đã học trước. Vì biết đường, nó đi nhanh bằng những bước dài. Cả hai cách đều đến được điểm đích dưới chân núi, nhưng thời gian cần dùng khác nhau.

Khi xử lý tinh thể, khó khăn nằm ở chỗ không gian không phẳng. Mạng tinh thể có chiều dài và góc. Tọa độ nguyên tử là vị trí trong chiếc hộp lập lại. Không gian như vậy không phẳng như tờ giấy, mà cong như quả cầu. Không gian cong này được gọi là đa tạp Riemann (Riemannian Manifold).

Nếu xử lý không gian cong như tờ giấy phẳng, sai số sẽ xuất hiện. Nó giống như trái quả địa cầu thành bản đồ phẳng thì diện tích bị méo. Vùng cực được vẽ lớn hơn thực tế. Tinh thể cũng vậy. Góc và thể tích của mạng tinh thể bị vặn lệch. Vì vậy, khớp dòng chảy dùng thước đo của không gian cong ngay từ đầu. Thước đo này gọi là metric Riemann (Riemannian Metric). Vì xử lý không gian cong đúng như nó vốn cong, méo lệch giảm xuống.

#### 3.1 Tìm đường tắt trên không gian cong

Ta sẽ xem khớp dòng chảy xử lý không gian cong như thế nào qua các mô hình tiêu biểu. Phần này tập trung vào các bài báo đã được kiểm chứng.

Tìm đường tắt trong không gian cong giống với chuyến bay trên Trái Đất. Đường tắt từ Seoul đến New York không phải là đường thẳng trên bản đồ. Vì Trái Đất có dạng tròn. Đường tắt thực sự là đường cong men theo bề mặt Trái Đất. Đường cong như vậy gọi là đường trắc địa (Geodesic). Sinh tinh thể cũng di chuyển theo đường trắc địa trên không gian cong.

Mô hình đầu tiên áp dụng tốt ý tưởng này vào tinh thể là FlowMM. FlowMM là mô hình mở rộng khớp dòng chảy Riemann cho phù hợp với tinh thể. Nó cũng xét đến đối xứng riêng của tinh thể. Nó xử lý đồng thời tịnh tiến ngang, quay, hoán đổi thứ tự nguyên tử và điều kiện biên tuần hoàn. Mô hình này đã giảm mạnh số bước cần để tìm vật liệu ổn định. Bài báo công khai báo cáo rằng nó tìm được ứng viên ổn định với số bước ít hơn khoảng ba lần so với phương pháp cũ.

FlowMM làm hai việc. Một là dự đoán cấu trúc ổn định khi thành phần đã được xác định. Hai là đề xuất đồng thời thành phần mới và cấu trúc của nó. Việc trước gọi là dự đoán cấu trúc tinh thể (Crystal Structure Prediction). Việc sau mới là phát hiện vật liệu mới đúng nghĩa. Khớp dòng chảy thực hiện nhanh cả hai việc. Vì vậy, nó được dùng ở nhiều vị trí trong tìm kiếm vật liệu.

Một trường hợp đã được kiểm chứng khác là CrystalFlow. Nghiên cứu này được đăng trên Nature Communications vào năm 2025. Đây là bài báo chính thức đã qua bình duyệt. CrystalFlow tạo vật liệu tinh thể bằng mô hình sinh dựa trên dòng chảy. Nó sinh đồng thời mạng tinh thể và bố trí nguyên tử. Nghiên cứu này cho thấy khớp dòng chảy làm việc huấn luyện dễ hơn mô hình khuếch tán.

Khớp dòng chảy cũng được dùng cho tinh thể hữu cơ. Một bản thảo trước bình duyệt năm 2026 dùng khớp dòng chảy ô đơn vị (Unit Cell Flow Matching). Bằng phương pháp này, nghiên cứu dự đoán nhanh cấu trúc tinh thể hữu cơ. Đây không phải là mô hình dành riêng cho tinh thể vô cơ. Vì vậy, cần thận trọng để không mở rộng kết quả sang vật liệu vô cơ. Dù vậy, nghiên cứu cho thấy khớp dòng chảy có thể được dùng cho nhiều loại tinh thể.

Ta hãy nối vấn đề tốc độ với tên lửa. Thiết kế ngược vật liệu phải tạo nhiều ứng viên rồi sàng lọc. Dù tạo ra mười nghìn ứng viên, có thể chỉ vài ứng viên dùng được. Nếu thời gian tạo một ứng viên quá dài, toàn bộ quá trình tìm kiếm bị nghẽn. Khớp dòng chảy giải nút thắt này. Trong cùng một thời gian, nó tạo ra nhiều ứng viên hơn. Ở các lĩnh vực có điều kiện khắt khe như vật liệu hàng không vũ trụ, tốc độ này có giá trị lớn.

Khớp dòng chảy giảm số bước được bao nhiêu. Bài báo FlowMM báo cáo rằng số bước để tìm ứng viên ổn định giảm khoảng ba lần. Nó đi con đường mà mô hình khuếch tán từng đi bằng ít bước hơn. Thời gian tạo một ứng viên vì thế ngắn lại. Những con số lớn kiểu nhanh hơn một trăm lần hoặc từ một ngày xuống còn vài phút không có căn cứ. Chỉ nhìn các giá trị đã được báo cáo cũng thấy thời gian tiết kiệm sẽ tích lũy khi số ứng viên cần tạo tăng lên.

Khớp dòng chảy vẫn còn bài toán chưa giải. Muốn học đường tắt, mô hình phải biết tốt phân bố mục tiêu. Nếu dữ liệu huấn luyện thiếu, nó học nhầm đường. Tính toán trên không gian cong vẫn còn khó. Vì vậy, khớp dòng chảy không thay thế mọi mô hình khuếch tán. Tùy bài toán mà chọn một trong hai phương pháp. Khi cần độ chính xác, dùng khuếch tán. Khi cần tốc độ, dùng khớp dòng chảy.

## 4. Chỉ vẽ lại một phần và để mô hình tự quyết định số nguyên tử

Không phải lúc nào cũng cần vẽ tinh thể từ trang trắng. Cũng có con đường chỉ thay đổi một phần từ cấu trúc sẵn có. Inpainting, tức lấp vào chỗ trống và khuyết tật, là một phương pháp như vậy. Cũng có ý tưởng gọi là nguyên tử ảo để xử lý bài toán trong đó số nguyên tử chưa được định sẵn.

Trước hết, cần nắm khái niệm inpainting (Inpainting). Inpainting là kỹ thuật lấp phần bị trống sao cho tự nhiên. Ta thường thấy trong chỉnh sửa ảnh. Khi xóa cột điện trong bức ảnh, chỗ đó bị trống. Trí tuệ nhân tạo nhìn bầu trời xung quanh rồi lấp phần trống. Trong cấu trúc tinh thể, nó cũng làm việc tương tự. Trong tinh thể, nó lấp tự nhiên các vị trí nguyên tử trống và các vị trí khuyết tật.

Vẽ tinh thể từ trang trắng khác với inpainting. Sinh từ trang trắng bắt đầu khi chưa đặt nguyên tử nào. Inpainting bắt đầu từ một cấu trúc đã có. Nó chỉ để trống một phần rồi vẽ lại riêng phần đó. Phần còn lại

được giữ nguyên. Hai cách này có cách dùng khác nhau. Khi tìm vật liệu hoàn toàn mới, ta dùng sinh từ trang trắng. Khi cải tiến một vật liệu tốt, ta dùng inpainting.

Có lý do khiến công nghệ này cần thiết trong nghiên cứu vật liệu. Trong thực tế, hiếm khi người ta phải vẽ lại toàn bộ tinh thể từ đầu. Thường đã có một tinh thể nền tốt (Host Matrix). Người ta thêm vào hoặc lấy bớt một lượng nhỏ nguyên tố khác để thay đổi tính chất. Việc thay một số nguyên tử bằng nguyên tử khác như vậy được gọi là pha tạp (Doping). Trạng thái nguyên tử bị thiếu hoặc chen vào vị trí khác được gọi là khuyết tật (Defect).

Pha tạp là phương pháp mà ngành bán dẫn đã dùng từ lâu. Silicon tinh khiết dẫn điện kém. Người ta cho vào đó một lượng rất nhỏ phosphor hoặc boron. Khi đó tính chất điện thay đổi lớn. Chip trong điện thoại thông minh vận hành theo nguyên lý này. Inpainting tinh thể giao việc thiết kế pha tạp này cho trí tuệ nhân tạo. Mô hình đề xuất nên đưa nguyên tố nào vào vị trí nào.

Thay đổi nhỏ này làm tính chất thay đổi lớn. Nó giống như chỉ trộn một ít chất khác vào muối mà vị đã khác đi. Nếu thay một nguyên tử trong tinh thể, độ dẫn điện sẽ thay đổi. Hiệu suất xúc tác cũng thay đổi. Vì vậy phải thiết kế thật chính xác sẽ đưa gì vào đâu. Inpainting giúp việc thiết kế chính xác này. Nó giữ nguyên nền và chỉ vẽ lại vị trí cần thay đổi.

Có lý do khiến inpainting thực dụng hơn tạo mới từ trang trắng. Bản thân việc tìm một tinh thể nền tốt đã mất nhiều thời gian. Nếu đã có nền được kiểm chứng, cải tiến trên đó sẽ nhanh hơn. Điều này giống như sửa một căn phòng thay vì xây mới một ngôi nhà vững chắc. Nền móng không lung lay nên rủi ro thấp. Vì vậy các nhà nghiên cứu thực tế thường dùng phương pháp inpainting.

Muốn lấp chỗ trống thì phải nhìn kỹ các nguyên tử xung quanh. Nếu đưa nguyên tố bất kỳ vào, cấu trúc sẽ sụp đổ. Nó phải khớp kích thước với nguyên tử lân cận. Cân bằng điện cũng phải phù hợp. Trí tuệ nhân tạo học quy luật của nền để đáp ứng các điều kiện này. Nó che riêng vị trí cần thay bằng một phương pháp gọi là che mặt nạ (Masking). Nó chỉ vẽ lại phần bị che và cố định phần còn lại.

## 4.1 Khuếch tán nguyên tử ảo MiAD tự quyết định số nguyên tử

Một mô hình tiền ấn phẩm có tên MiAD là phương pháp xử lý linh hoạt số nguyên tử. Ta sẽ xem vì sao ý tưởng nguyên tử ảo lại hữu ích.

MiAD là viết tắt của Mirage Atom Diffusion, tức khuếch tán nguyên tử ảo. Đây là một tiền ấn phẩm xuất hiện vào cuối năm 2025. Đây là nghiên cứu giai đoạn đầu, trước khi qua bình duyệt. Tên chính thức hàm ý một mô hình khuếch tán tạo tinh thể mới từ đầu. Mã định danh là arXiv:2511.14426.

Ý tưởng cốt lõi của mô hình này là nguyên tử ảo. Mirage có nghĩa là ảo ảnh. Nguyên tử ảo là một vị trí giả định có thể trở thành nguyên tử thật hoặc biến mất. Trước hết, mô hình lấp đầy trong hộp nhiều vị trí giả định. Sau đó, trong quá trình khuếch tán, nó phán đoán từng vị trí là nguyên tử thật hay chỗ trống. Vị trí thật sự còn lại sẽ được đặt nguyên tố vào, còn vị trí không phải thì biến mất.

Ưu điểm của cách này nằm ở chỗ nó không ấn định trước số nguyên tử. Các mô hình cũ tạo cấu trúc sau khi quyết định trước số lượng nguyên tử. MiAD tự quyết định số lượng bằng cách bật và tắt các vị trí giả định. Điều bài báo này thật sự cho thấy là khả năng tạo sinh tự quyết định số nguyên tử. Đây chưa phải là nghiên cứu đã kiểm chứng nồng độ khuyết tật, vị trí pha tạp hay tính phi hợp thức hóa học. Tuy vậy, vẫn có khoảng trống để thử gán phương pháp này vào các vật liệu có số nguyên tử linh động.

Vật liệu có tỷ lệ nguyên tố không khớp trọn vẹn theo số nguyên được gọi là hợp chất phi hợp thức hóa học. Những vật liệu như vậy thường gặp trong vật liệu thực tế. Một chỗ trống do thiếu một nguyên tử có thể thay đổi tính chất. Nhiều khuyết tật cũng có thể tụ lại thành cụm khuyết tật. Cũng có kiểu xen kẽ,

trong đó nguyên tử chen vào vị trí còn dư. Sự xáo trộn rất nhỏ này quyết định độ dẫn và độ bền. Phương pháp tạo sinh xử lý linh hoạt số nguyên tử trở thành manh mối để biểu diễn những xáo trộn như vậy.

MiAD đã báo cáo kết quả trong thử nghiệm benchmark. Nó dùng bộ thử nghiệm phổ biến MP-20. Chỉ số biểu thị tỷ lệ tinh thể ổn định, độc nhất và mới được gọi là tỷ lệ SUN. Bài báo báo cáo một mức thành quả nhất định theo chỉ số này. Tuy nhiên, giá trị này là số liệu ở giai đoạn tiền ấn phẩm. Vì vậy cần tiếp nhận nó như kết quả ban đầu, không phải hiệu năng đã được xác nhận.

Ta sẽ chỉ ra ý nghĩa của công nghệ này đối với tên lửa và vật liệu mới. Vật liệu dùng thật trong hàng không vũ trụ không hoàn toàn có trật tự. Luôn có khuyết tật, như một ít nguyên tử bị thiếu hoặc đổi chỗ. Những khuyết tật này quyết định độ bền và độ dẫn. Mô hình xử lý linh hoạt số nguyên tử biểu diễn tốt hơn các cấu trúc thực tế như vậy. Nhờ đó nó đưa ra ứng viên gần hơn với kết quả phòng thí nghiệm. Tuy nhiên, phương pháp này vẫn ở giai đoạn tiền ấn phẩm nên hiệu năng thiết kế khuyết tật cần được kiểm chứng thêm.

## 5. Dẫn hướng về phía ứng viên tốt bằng học tăng cường

Mô hình tạo sinh đưa ra nhiều ứng viên. Nhưng không phải tất cả đều có ích. Học tăng cường giúp định hướng về phía ứng viên tốt. Ta sẽ xem nguyên lý và cơ chế an toàn trong cách học tăng cường dẫn hướng.

Học tăng cường (Reinforcement Learning) là cách học sửa hành động dựa trên phần thưởng. Hãy nghĩ đến việc huấn luyện chó thì dễ hiểu. Làm tốt thì được cho đồ ăn. Làm không tốt thì không được cho. Con chó tăng dần những hành động giúp nó nhận đồ ăn. Trí tuệ nhân tạo cũng tăng những hành động nhận được phần thưởng cao.

Trong tạo sinh vật liệu, cấu trúc tốt được nhận phần thưởng cao. Tinh thể ổn định và có tính chất tốt được cho điểm lớn. Cấu trúc có nguyên tử va chạm nhau hoặc bất ổn được cho điểm thấp. Mô hình thay đổi thói quen tạo sinh theo hướng điểm cao. Vì vậy, càng thử nhiều thì ứng viên dùng được càng tăng.

Mô hình tạo sinh và học tăng cường có vai trò khác nhau. Mô hình tạo sinh rải rộng các ứng viên có vẻ hợp lý. Học tăng cường đẩy các ứng viên đó về phía mục tiêu. Nó giống như bàn tay gieo hạt và bàn tay dẫn dòng nước là hai bàn tay khác nhau. Hai bàn tay phải cùng chuyển động thì mới có quả tốt. Chỉ tạo sinh thôi cũng thiếu, chỉ học tăng cường thôi cũng thiếu. Khi nối hai thứ với nhau, ta tiến gần hơn đến tính chất mong muốn.

Có lý do khiến phương pháp này cần thiết. Mô hình tạo sinh tạo tốt các ứng viên có vẻ hợp lý. Nhưng về hợp lý và sự ổn định thật sự là hai chuyện khác nhau. Nhiều cấu trúc trông đẹp bên ngoài nhưng sụp đổ về mặt nhiệt động lực học. Vì vậy chỉ tạo sinh thôi là chưa đủ. Phải xem đồng thời độ ổn định, tính mới và khả năng tổng hợp. Học tăng cường đưa ba yếu tố này vào phần thưởng để dẫn mô hình.

Khả năng tổng hợp là một mục tiêu khó xử lý. Một vật liệu ổn định trên giấy vẫn có thể không tạo ra được trong phòng thí nghiệm. Con đường chế tạo có thể quá gian nan hoặc không có nguyên liệu. Vì vậy chỉ nhìn độ ổn định là chưa đủ. Cũng phải xem nó có thể thật sự tạo ra hay không. Phần thưởng của học tăng cường cố gắng chứa cả các điều kiện thực tế này. Có như vậy mới có ứng viên chuyển sang phòng thí nghiệm.

Cũng có nhiều cách gắn học tăng cường vào tạo sinh vật liệu. Một kỹ thuật tiêu biểu là tối ưu hóa chính sách gần (Proximal Policy Optimization). Kỹ thuật này chỉ thay đổi chính sách từng chút một mỗi lần. Lý do là nếu thay đổi quá lớn đột ngột, quá trình học sẽ dao động. Nó tiến lên thận trọng từng chút và tăng phần thưởng. Nhờ vậy hiệu năng tăng ổn định. Kỹ thuật này cũng được dùng rộng rãi để tinh chỉnh mô hình ngôn ngữ.

Một tiền ấn phẩm công bố năm 2025 cho thấy rõ hướng đi này. Nghiên cứu đã tinh chỉnh mô hình khuếch tán bằng học tăng cường và nâng tỷ lệ thành công của thiết kế ngược. Nó tạo ra nhiều cấu trúc phù hợp hơn với tính chất mục tiêu. Tỷ lệ đáp ứng các điều kiện ổn định, độc nhất và mới cũng tăng. Nghiên cứu này ở giai đoạn đầu. Vì vậy để ứng dụng công nghiệp, vẫn cần thêm tính toán DFT và tổng hợp thực nghiệm.

## 5.1 Cơ chế an toàn và chỉ số hiệu năng của học tăng cường

Học tăng cường có hai cơ chế an toàn. Một cơ chế giữ cho mô hình không đi lệch quá xa. Cơ chế còn lại lưu các ứng viên tốt để dùng lại. Ta cũng sẽ xem các chỉ số đo hiệu năng.

Học tăng cường có rủi ro ẩn. Nếu chỉ đuổi theo phần thưởng, nó sẽ nghiêng về cấu trúc kỳ lạ. Nó bị kẹt trong một vùng hẹp cho điểm cao. Nó đánh mất sự đa dạng của ứng viên và chỉ lặp lại một đáp án. Hiện tượng này được gọi là sụp đổ mode (Mode Collapse). Khi sụp đổ xảy ra, việc tìm kiếm dừng lại.

Cơ chế ngăn việc này là chính quy hóa KL (KL Regularization). KL là thước đo hai phân phối khác nhau bao nhiêu. Cơ chế này giữ cho mô hình đã tinh chỉnh không đi quá xa phân phối mà nó đã học ban đầu. Nó giống như bị buộc vào một sợi dây cao su. Mô hình tiến theo hướng mới, nhưng nếu rời quá xa gốc thì bị kéo lại. Nhờ vậy nó vừa tìm ứng viên tốt vừa giữ được sự đa dạng.

Một cơ chế nữa là phát lại kinh nghiệm (Experience Replay). Nó không bỏ đi các ứng viên từng tốt mà lưu lại. Các ứng viên đã lưu được lấy ra lần nữa để dùng cho học. Điều này giống như gom các bài đã làm tốt vào sổ ôn tập rồi giải lại. Làm như vậy giúp mô hình không quên những cấu trúc tốt đã khó khăn tìm được. Quá trình học tiếp tục ổn định hơn.

Chỉ số thường dùng để đo hiệu năng là tỷ lệ SUN. SUN chỉ các tính thể ổn định (Stable), độc nhất (Unique) và mới (Novel). Đây là tỷ lệ ứng viên thỏa mãn cả ba điều kiện. Ổn định có nghĩa là không dễ tách ra thành vật liệu khác. Độc nhất có nghĩa là các sản phẩm tạo sinh không trùng nhau. Mới có nghĩa là vật liệu chưa có trong cơ sở dữ liệu.

Khi đo độ ổn định, người ta xem một giá trị gọi là năng lượng trên bao lồi (Energy above Hull). Giá trị này được tính bằng DFT ở độ không tuyệt đối. Giá trị càng thấp thì càng ổn định theo tính toán. Nếu giá trị gần 0, nó được duy trì tốt trong phạm vi tính toán. Nhưng nhiệt độ và áp suất thực tế, nước và oxy, tốc độ phản ứng không nằm trong đó. Cũng có sai số tính toán. Vì vậy phải đọc giá trị này như một chỉ số chỉ ra ứng viên ổn định theo tính toán. Phần thưởng của học tăng cường dùng giá trị này như một thành phần quan trọng.

Phần thưởng đặt nhiều mục tiêu cùng lên bàn cân. Nếu chỉ đuổi theo độ ổn định, vật liệu tạo ra sẽ có tính chất nhạt. Nếu chỉ đuổi theo tính chất, vật liệu tạo ra sẽ khó tổng hợp. Vì vậy người ta gán trọng số riêng cho độ ổn định, tính chất và tính mới. Mục tiêu nào cần coi trọng hơn sẽ được điều chỉnh bằng trọng số. Việc điều chỉnh trọng số này là phần việc của người thiết kế. Khi chỉnh bàn cân phù hợp với mục tiêu, ứng viên sẽ tụ lại theo hướng mong muốn.

Ta sẽ chỉ ra ý nghĩa của các cơ chế an toàn này đối với thiết kế vật liệu tên lửa. Vật liệu hàng không vũ trụ chỉ có ý nghĩa khi thật sự tổng hợp được. Cấu trúc chỉ tốt trên giấy thì vô dụng. Học tăng cường đưa khả năng tổng hợp vào phần thưởng. Chính quy hóa KL và phát lại kinh nghiệm giữ cho quá trình tìm kiếm ổn định. Kết quả là tỷ lệ ứng viên đủ khả năng chuyển sang phòng thí nghiệm tăng lên.

## 6. Các trường hợp áp dụng thực tế của pipeline thiết kế ngược

Các công nghệ trên đã được dùng ra sao trong nghiên cứu thực tế. Ta sẽ xem chất điện ly rắn cho pin, xúc tác và hợp kim hàng không vũ trụ. Phần này dựa trên các nghiên cứu mới đã được kiểm chứng. Ta sẽ xem tạo sinh, tính toán và thực nghiệm ăn khớp với nhau thế nào.

Muốn dùng các công nghệ đã bàn ở trên trong thực tế, phải nối nhiều công cụ với nhau. Mô hình khuếch tán và flow matching tạo ra ứng viên. Nhận biết đối xứng và inpainting nâng chất lượng ứng viên. Học tăng cường dẫn ứng viên về phía mục tiêu. Khi các công cụ này nối thành một luồng, chúng mới phát huy sức mạnh. Luồng này được gọi là pipeline thiết kế ngược (Pipeline). Pipeline có nghĩa là nối nhiều công đoạn bằng đường ống.

Thiết kế ngược trong thực tế thường đi qua ba giai đoạn. Trước hết, ở giai đoạn tạo sinh, trí tuệ nhân tạo vẽ cấu trúc ứng viên phù hợp với tính chất mục tiêu. Tiếp theo, ở giai đoạn kiểm chứng bằng tính toán, DFT và thế năng liên nguyên tử học máy (MLIP) tính độ ổn định và tính chất. Cuối cùng, ở giai đoạn tổng hợp thực nghiệm, phòng thí nghiệm tự hành hoặc nhà nghiên cứu thật sự tạo ra và kiểm tra.

Ba giai đoạn này không chỉ chảy theo một hướng. Kết quả thu được từ thực nghiệm được đưa ngược lại cho mô hình tạo sinh. Mô hình dùng phản hồi này để vẽ ứng viên tiếp theo tốt hơn. Cách vận hành có vòng phản hồi như vậy được gọi là học chủ động (Active Learning). Một nghiên cứu công bố năm 2025 đã nâng đáng kể hiệu quả thiết kế ngược bằng học chủ động. Nghiên cứu lặp lại việc tạo và kiểm chứng ứng viên để tiến gần tính chất mục tiêu.

Ta xem kỹ hơn giai đoạn thứ hai, tức kiểm chứng bằng tính toán. Ứng viên do mô hình tạo sinh tạo ra vẫn chỉ là hình vẽ trên giấy. Phải dùng tính toán để xem hình vẽ này có thật sự ổn định hay không. DFT chính xác nhưng chậm. Vì vậy người ta dùng cùng thế năng liên nguyên tử học máy đã bàn ở chương trước. Phương pháp này cho độ chính xác gần DFT với tốc độ nhanh hơn nhiều. Nó nhanh chóng lọc rất nhiều ứng viên. Sau đó chỉ một số ít có triển vọng được kiểm tra lại bằng DFT.

Sự phối hợp này quyết định tốc độ tìm kiếm. Dù tạo sinh nhanh đến đâu, nếu kiểm chứng chậm thì cũng không có ích. Nếu kiểm chứng trở thành nút thắt, ứng viên chỉ chất đống. Tạo sinh nhanh và kiểm chứng nhanh phải ăn khớp với nhau. Vì vậy nghiên cứu mới nối ba giai đoạn thành một luồng tự động. Phòng thí nghiệm tự hành sẽ bàn ở chương sau đảm nhiệm phần cuối của luồng này.

## 6.1 Thiết kế chất điện ly rắn halide cho pin thể rắn

Trường hợp đầu tiên là thiết kế chất điện ly rắn cho pin. Ta sẽ xem vì sao vật liệu này quan trọng và AI giúp ra sao. Phần này tập trung vào các con số đã được kiểm chứng.

Pin thể rắn toàn phần (All-Solid-State Battery) là loại pin dùng chất điện phân rắn. Pin thông thường có chất điện phân ở dạng lỏng. Chất điện phân lỏng có nguy cơ bắt lửa. Chất điện phân rắn làm giảm nguy cơ này. Tuy nhiên, không phải cứ là chất rắn thì luôn an toàn. Nó cũng có thể phản ứng với kim loại lithium, hơi ẩm hoặc điện cực. Hệ sulfide và một số hệ halide có thể tạo ra sản phẩm phân hủy có hại, hoặc gây điện trở và vết nứt tại bề mặt tiếp xúc. Dù vậy, lợi ích giảm rò rỉ chất lỏng và nguy cơ cháy vẫn lớn, nên lĩnh vực này nhận được nhiều quan tâm.

Các tính chất đòi hỏi ở chất điện phân rắn rất khắt khe. Ion lithium phải đi qua bên trong chất rắn thật nhanh. Mức độ ion đi qua tốt đến đâu được gọi là độ dẫn ion (Ionic Conductivity). Giá trị càng cao càng tốt. Đồng thời, vật liệu phải chịu được điện áp cao. Nó không được quá mềm và phải bám ép tốt. Rất khó đáp ứng tất cả các điều kiện này.

Chất điện phân rắn hệ halide (Halide Solid Electrolyte) được xem là một ứng viên đầy triển vọng. Halide chỉ các nguyên tố như chlorine, bromine và iodine. Nhóm này tương đối ổn định ở điện áp cao. Sự cân bằng giữa độ mềm và độ dẫn cũng khá tốt. Vì vậy, nghiên cứu đang diễn ra sôi động. Một vật liệu tiêu biểu là hợp chất gồm lithium, yttrium và chlorine.

Trí tuệ nhân tạo hỗ trợ thiết kế vật liệu này theo nhiều cách. Một nghiên cứu công bố năm 2025 đã kết hợp học máy với tính toán DFT. Nghiên cứu đó dự đoán nhiều tính chất cùng lúc để lọc ra các ứng viên

halide tốt. Họ dự đoán đồng thời độ dẫn ion, mô đun đàn hồi và cửa sổ ổn định điện hóa. Cách tiếp cận xem nhiều tính chất cùng lúc như vậy đã làm tăng tính thực dụng.

Một nghiên cứu khác công bố năm 2026 đã dùng phương pháp rút ra quy luật từ dữ liệu. Nghiên cứu đó tìm ra những đặc trưng hóa học then chốt chi phối độ dẫn ion. Dựa trên các đặc trưng này, họ thiết kế thành phần mới. Bài báo cho biết một vật liệu được thiết kế đã cho thấy độ dẫn ion khoảng 12 millisiemens trên mỗi centimeter ở nhiệt độ phòng. Giá trị này thuộc nhóm cao trong hệ vật liệu này. Nó cho thấy dòng chảy từ việc rút ra quy luật trong dữ liệu đến thiết kế vật liệu.

Rất khó tính toán ion đi qua bên trong chất rắn như thế nào. Ion lithium nhảy từ vị trí này sang vị trí khác để di chuyển. Rào cản của cú nhảy này phải thấp thì độ dẫn mới cao. Muốn tính rào cản này, phải theo dõi chuyển động của nguyên tử trong thời gian dài. Ở đây cũng dùng thể năng liên nguyên tử học máy (MLIP). Một nghiên cứu năm 2025 báo cáo rằng phương pháp này có thể dùng được trong tính toán chất dẫn ion rắn. Ta tạo ra ứng viên bằng mô hình sinh, rồi dùng phương pháp này để kiểm tra sự di chuyển của ion.

Hãy xem ý nghĩa của vật liệu này trong hàng không vũ trụ và tên lửa. Vệ tinh và tàu thăm dò phụ thuộc rất nhiều vào pin. Khi không có ánh sáng Mặt Trời, chúng sống bằng điện đã lưu trữ. Nếu pin không ổn định, toàn bộ nhiệm vụ sẽ bị lung lay. Pin thể rắn toàn phần an toàn và nhẹ giúp giảm rủi ro này. Thiết kế ngược bằng AI giúp tìm ra vật liệu pin như vậy nhanh hơn.

## 6.2 Thiết kế tìm chất xúc tác tách nước

Ví dụ tiếp theo là thiết kế chất xúc tác. Ta xem chất xúc tác là gì và AI tìm chất xúc tác tốt như thế nào. Ta cũng xem một ví dụ mới có sự phối hợp với phòng thí nghiệm robot.

Chất xúc tác (Catalyst) là vật liệu giúp phản ứng hóa học diễn ra nhanh hơn. Bản thân nó hầu như không thay đổi. Nhưng nó làm tăng tốc độ phản ứng. Thiết bị tạo hydrogen và oxygen từ nước nhất thiết cần chất xúc tác. Thiết bị này gọi là thiết bị điện phân nước. Có chất xúc tác tốt thì với cùng lượng điện, ta thu được nhiều hydrogen hơn.

Hiệu năng của chất xúc tác phụ thuộc vào cách sắp xếp nguyên tử trên bề mặt. Phản ứng diễn ra trên bề mặt chất xúc tác. Chất phản ứng bám lên bề mặt, rồi rời ra trong khi biến đổi. Việc một chất bám lên bề mặt gọi là hấp phụ (Adsorption). Bám quá mạnh cũng là vấn đề. Chất đó khó rời ra, nên phản ứng bị cản lại. Bám quá yếu cũng là vấn đề. Nó rời khỏi bề mặt trước khi phản ứng xảy ra. Chất xúc tác tốt giữ chất phản ứng với độ mạnh vừa đủ.

Độ mạnh hấp phụ do cách sắp xếp các nguyên tử bề mặt quyết định. Chỉ cần khoảng cách giữa các nguyên tử thay đổi một chút, độ mạnh cũng thay đổi. Vì vậy, cách thiết kế bề mặt rất quan trọng. Trí tuệ nhân tạo thay đổi vị trí nguyên tử trên bề mặt để điều chỉnh độ mạnh hấp phụ. Nó tìm cách sắp xếp giữ chất trung gian phản ứng ở mức phù hợp. Cách sắp xếp này là bí quyết của chất xúc tác tốt.

Phản ứng tạo oxygen rất khó. Phản ứng này gọi là phản ứng tiến hóa oxygen (Oxygen Evolution Reaction). Để gây ra phản ứng này, phải đặt điện áp cao hơn điện áp cần thiết theo lý thuyết. Phần điện áp dư đó gọi là quá thế (Overpotential). Quá thế là khái niệm khác với năng lượng hoạt hóa, tức rào cản của tốc độ phản ứng. Quá thế càng thấp thì chất xúc tác càng tốt. Quá thế thấp có nghĩa là dùng ít điện hơn. Vì vậy, giảm quá thế là một mục tiêu lớn trong thiết kế chất xúc tác.

Một nghiên cứu công bố năm 2026 đã thiết kế ngược chất xúc tác bằng AI tạo sinh. Nghiên cứu này xử lý chất xúc tác entropy cao (High-Entropy Catalyst), trong đó nhiều nguyên tố được trộn với nhau. Họ kết hợp thông tin quang phổ, mô hình sinh và thiết bị thí nghiệm robot thành một hệ thống. Robot tạo ứng viên, đo hiệu năng, rồi trả kết quả về cho mô hình. Vòng phản hồi này được chạy nhanh.

Các con số nghiên cứu này báo cáo cho thấy tính thực dụng. Thời gian để tạo và đo một mẫu đã giảm mạnh. Công việc từng mất khoảng 20 giờ được rút xuống còn 78 phút. Và ứng viên tốt nhất tìm được có quá thế giảm thêm 32 millivolt. Ví dụ này vừa là nghiên cứu học thuật, vừa cho thấy tự động hóa bằng robot. Ta thấy tốc độ tăng lên ra sao khi mô hình sinh và thí nghiệm ăn khớp với nhau.

Dòng chảy này làm thay đổi công việc của con người. Trước đây, nhà nghiên cứu mất cả ngày để tạo một mẫu. Giờ robot làm việc đó cả ngày lẫn đêm. Con người suy nghĩ nên đặt mục tiêu nào. Con người phán đoán kết quả nào có ý nghĩa. Công việc chuyển từ dùng tay sang dùng suy nghĩ. Vị trí của con người không biến mất mà chuyển lên cao hơn. Chương sau sẽ bàn kỹ hơn về phòng thí nghiệm tự hành này.

Có lý do khiến chất xúc tác entropy cao trộn nhiều nguyên tố được chú ý. Khi trộn nhiều nguyên tố, bề mặt sinh ra nhiều loại vị trí khác nhau. Trong số đó, xác suất xuất hiện vị trí phù hợp với phản ứng tăng lên. Nhưng số tổ hợp quá lớn, con người khó thử hết. Mô hình sinh thay con người rà soát không gian rộng này. Nó thu hẹp thành phần tốt và giảm số lần thí nghiệm. Robot nhanh chóng tạo các ứng viên đó để kiểm tra.

Hãy xem ý nghĩa của chất xúc tác trong tên lửa và thám hiểm vũ trụ. Trong các nhiệm vụ vũ trụ dài ngày, con người phải tự tạo nước và không khí. Chất xúc tác được dùng để tách nước và lấy oxygen. Chất xúc tác cũng được dùng để tạo nhiên liệu. Chất xúc tác tốt tạo ra nhiều oxygen và nhiên liệu hơn với cùng lượng điện. Nhờ vậy, khối lượng hàng tàu vũ trụ phải chở sẽ nhẹ hơn. Thiết kế ngược bằng AI giúp đưa những chất xúc tác này đến sớm hơn.

Hydrogen cũng được dùng làm nhiên liệu tên lửa. Hydrogen lỏng tạo lực đẩy mạnh. Nếu công nghệ tách nước để lấy hydrogen phát triển, việc sản xuất nhiên liệu sẽ dễ hơn. Khi đó, chất xúc tác vẫn giữ chìa khóa của phản ứng. Nếu tạo được chất xúc tác tốt từ nguyên tố rẻ tiền, chi phí sẽ giảm. Mục tiêu là dùng ít platinum quý hiếm hơn. AI tạo sinh phát huy sức mạnh trong việc tìm các chất xúc tác không dùng kim loại quý như vậy.

### 6.3 Thiết kế ngược hợp kim và siêu hợp kim dùng trong hàng không vũ trụ

Ví dụ cuối cùng là thiết kế hợp kim cho động cơ tên lửa. Ta xem vì sao cần hợp kim mới. Ta tìm hiểu mô hình sinh tìm thành phần hợp kim như thế nào. Nội dung dựa trên các nghiên cứu mới đã được kiểm chứng.

Động cơ tên lửa phải chịu nhiệt cực hạn. Bên trong buồng đốt nóng tới hàng nghìn độ. Vòi phun bị dòng khí tốc độ cao bào mòn. Trong thời gian dài, siêu hợp kim nền nickel (Superalloy) đã giữ vị trí này. Siêu hợp kim là hợp kim không mất độ bền ngay cả ở nhiệt độ cao. Nhưng siêu hợp kim nickel cũng có giới hạn nhiệt độ. Trong môi trường nóng hơn, nó nóng chảy hoặc mềm đi.

Một lựa chọn thay thế là hợp kim entropy cao (High-Entropy Alloy). Hợp kim thông thường trộn một lượng nhỏ nguyên tố khác vào một nguyên tố chính. Hợp kim entropy cao trộn nhiều nguyên tố với lượng tương tự nhau. Khi trộn như vậy, cấu trúc ổn định theo cách khó ngờ. Loại được tạo từ các nguyên tố có điểm nóng chảy cao gọi là hợp kim entropy cao chịu lửa (Refractory HEA). Hợp kim này được xem là ứng viên thay thế siêu hợp kim nickel trong môi trường trên 1100 độ.

Lý do hợp kim entropy cao ổn định nằm ở sự vô trật tự. Khi nhiều nguyên tố trộn đều, số cách sắp xếp tăng lên. Rất nhiều khả năng đó giữ cấu trúc lại. Tính chất này được gọi là mức vô trật tự sắp xếp cao. Nói cách khác, sự vô trật tự lại bảo vệ trật tự. Nhờ tính chất này, vật liệu chịu được tốt trong môi trường cực hạn. Vòi phun và buồng đốt tên lửa là những vị trí tiêu biểu đòi hỏi khả năng chịu đựng như vậy.

Vấn đề là số tổ hợp lại nhiều như sao trên trời. Chỉ trộn năm nguyên tố thôi, số trường hợp về tỷ lệ đã nhiều vô kể. Rất khó tạo từng mẫu để thử nghiệm. Ở đây mô hình sinh bước vào. Khi đặt tính chất mục

tiêu, nó tìm ngược thành phần phù hợp. Ta đưa vào điều kiện về độ bền và độ dai. Mô hình đề xuất tỷ lệ nguyên tố có khả năng tạo ra các giá trị đó.

Nhiều nghiên cứu công bố năm 2026 cho thấy hướng đi này. Một nghiên cứu đã hệ thống hóa nhiều phương pháp sinh thành phần hợp kim. Họ xem xét cùng lúc bộ tự mã hóa biến phân, mạng sinh đối nghịch và mô hình khuếch tán. Một nghiên cứu khác đã thu thập tài liệu và dữ liệu để thu hẹp các ứng viên siêu hợp kim entropy cao nhiệt độ cao. Họ chọn thành phần phù hợp với tính chất mục tiêu và giảm số ứng viên cần thử nghiệm. Cả hai nghiên cứu đều chưa ở giai đoạn tạo linh kiện thật, mà ở giai đoạn thu hẹp ứng viên.

Tính chất của hợp kim không chỉ do thành phần quyết định. Cùng một thành phần, nhưng phương pháp chế tạo khác nhau sẽ cho kết quả khác nhau. Cách cấp nhiệt và làm nguội làm thay đổi vi cấu trúc. Vi cấu trúc là kích thước và cách sắp xếp của các hạt tinh thể. Vì vậy, thiết kế hợp kim phải xem cả thành phần lẫn quy trình. Mô hình sinh cũng đang tiến tới hướng xử lý đồng thời hai yếu tố này. Các chương sau sẽ bàn thêm về câu chuyện quy trình này.

Tuy nhiên, có điểm cần thận trọng. Hợp kim entropy cao chịu lửa vẫn còn những bức tường phải vượt qua. Nó vẫn có tính dễ gãy ở nhiệt độ phòng. Ở nhiệt độ cao, nó cũng có thể phản ứng với oxygen khiến bề mặt bị hư hại. Vì vậy, vẫn có khoảng cách giữa hứa hẹn trong phòng thí nghiệm và linh kiện thật. Dù mô hình sinh giúp thu hẹp ứng viên, hai bức tường này vẫn phải vượt qua bằng thí nghiệm. Kiểm soát độ giòn ở nhiệt độ phòng và oxy hóa ở nhiệt độ cao là nhiệm vụ tiếp theo.

## 6.4 So sánh mô hình sinh và tầm quan trọng của kiểm chứng

Ta so sánh nhiều mô hình sinh như thế nào. Ta xem chúng được đo bằng chỉ số nào. Ta tìm hiểu vì sao kiểm chứng quan trọng hơn kết quả. Ta cũng nêu lý do phải thận trọng với những con số chưa được xác nhận.

Khi so sánh mô hình sinh, ta xem nhiều thước đo cùng lúc. Một thước đo là độ ổn định, xem tinh thể được tạo ra có tự duy trì được hay không. Ta cũng xem tính mới. Ta kiểm tra xem vật liệu đó có phải là vật liệu chưa có trong cơ sở dữ liệu hay không. Ta cũng đo tính hợp lý của khoảng cách nguyên tử. Ta kiểm tra xem các nguyên tử có ở quá gần nhau hay không. Ta cũng xem sự bảo toàn đối xứng. Ta kiểm tra xem đối xứng tinh thể có được giữ hay không. Cuối cùng là tốc độ, đo thời gian cần để tạo một ứng viên.

Không phải chỉ cần một chỉ số tốt là thành mô hình tốt. Dù tốc độ nhanh, nếu độ ổn định thấp thì giá trị sử dụng ít. Dù giữ đối xứng tốt, nếu không có tính mới thì ý nghĩa mờ nhạt. Vì vậy, ta xem sự cân bằng giữa nhiều chỉ số. Tùy vấn đề mà chỉ số quan trọng sẽ khác nhau. Với vật liệu pin, độ ổn định và độ dẫn đi trước. Nếu cần tìm kiếm nhanh, tốc độ đi trước.

Muốn so sánh công bằng thì phải dùng cùng một đề thi. Trong lĩnh vực sinh tinh thể, có các bộ kiểm thử được dùng rộng rãi. MP-20 là một ví dụ tiêu biểu. Đây là tập hợp khoảng hai mươi nghìn tinh thể lấy từ Materials Project. Cũng có bộ chỉ gồm tinh thể perovskite. Tất cả mô hình được thử trên cùng một bộ. Như vậy mới có thể so sánh mô hình nào tốt hơn. Nếu đề thi khác nhau, không thể đặt điểm số cạnh nhau.

Ở đây có một điểm cần thận trọng. Các con số ghi song song hiệu năng của nhiều mô hình thường chưa được xác nhận trong văn bản gốc. Vì vậy, cuốn sách này không dùng các con số trong những bảng chưa được kiểm chứng. Những con số chưa được xác nhận dễ dẫn con người đi sai hướng.

Thay vào đó, ta nắm cảm giác bằng một kết quả đã được kiểm chứng. Đó là nghiên cứu về mô hình sinh mà Microsoft công bố trên Nature năm 2025. Tên mô hình này là MatterGen. Mô hình này tạo vật liệu vô cơ trên toàn bộ bảng tuần hoàn nguyên tố. Nó cũng có thể được tinh chỉnh theo tính chất mục tiêu. Bài

báo cho biết sản phẩm do mô hình này tạo ra ổn định và mới thường xuyên gấp đôi so với mô hình cũ. Bài báo cũng cho biết chúng gần điểm cực tiểu năng lượng hơn gấp mười lần.

Mô hình này có thể học lại theo nhiều điều kiện. Nếu học theo giá trị đàn hồi, nó tạo ra vật liệu cứng. Nếu học theo giá trị từ tính, nó tạo ra vật liệu nam châm. Nó cũng có thể học để chỉ dùng các nguyên tố mong muốn. Đây là cách chỉnh sửa một mô hình nền theo nhiều mục tiêu. Cách này gọi là tinh chỉnh (Fine-tuning). Ta không xây lại nền móng, mà chỉ thay mục tiêu vào đó.

Nghiên cứu này đã tiến tới bước kiểm chứng bằng thực nghiệm. Nhóm nghiên cứu đặt mô đun đàn hồi khối mục tiêu là 200 gigapascal rồi tạo ra cấu trúc vật chất mới. Họ đã thực sự tổng hợp vật chất này. Tuy nhiên, thứ được tổng hợp là dung dịch rắn trong đó các nguyên tử trộn lẫn một cách mất trật tự. Nhóm nghiên cứu đo mô đun Young bằng phép ấn nano, rồi dùng giá trị thu được từ tính toán để ước tính mô đun đàn hồi khối. Giá trị ước tính thu được theo cách này đạt tối đa 169 gigapascal trong bốn lần, còn trung bình là  $158 \pm 11$  gigapascal. Cần đọc tách bạch giữa giá trị đo và giá trị quy đổi, giữa cấu trúc tính toán và cấu trúc đã tổng hợp. Dù vậy, đây vẫn là một minh chứng hiếm có nối từ khâu sinh tạo đến khâu tổng hợp.

Khi xử lý bài báo tiền ấn phẩm, cần thận trọng hơn. Nhiều mô hình được giới thiệu trong chương này đang ở giai đoạn tiền ấn phẩm. Tiền ấn phẩm là trước khi qua bình duyệt đồng cấp. Kết quả có thể thay đổi về sau. Vì vậy, cuốn sách này giới thiệu tiền ấn phẩm như nghiên cứu giai đoạn đầu. Dù con số có xuất sắc đến đâu, trước khi kiểm chứng thì không viết như điều đã xác định. Thái độ này bảo vệ người đọc khỏi niềm tin sai.

Hãy xem vì sao kiểm chứng quan trọng hơn kết quả. Mô hình sinh tạo có thể tạo ra vô hạn ứng viên. Nhưng không phải cứ tạo ra là tất cả đều có thật. Phải sàng lọc bằng tính toán. Phải xác nhận bằng thực nghiệm. Thiết kế ngược vật liệu chỉ có ý nghĩa khi sinh tạo, tính toán và thực nghiệm cùng vận hành. Trong những lĩnh vực liên quan đến an toàn như vật liệu dùng cho hàng không vũ trụ, nguyên tắc này càng phải nghiêm ngặt hơn. Chỉ sau khi qua kiểm chứng tính toán và tổng hợp thực nghiệm, ta mới có thể tin vào ứng viên.

## Tổng kết

Chương này là câu chuyện đảo ngược hướng tìm kiếm vật liệu. Cách cũ chọn vật liệu trước rồi đo tính chất sau. Cách mới xác định tính chất trước rồi vẽ vật liệu sau. Cách mới này được gọi là thiết kế ngược vật liệu. Chỉ đổi một hướng mà sức mạnh tìm kiếm đã thay đổi lớn. Con đường để vẽ ra những vật chất chưa từng có đã mở ra.

Nhân vật chính của thiết kế ngược là trí tuệ nhân tạo sinh tạo. Mô hình khuếch tán tạo nên tinh thể bằng cách xóa nhiễu khỏi một đám nguyên tử mờ. Nguyên lý của nó giống AI tạo ảnh. Sinh tạo tinh thể khó hơn. Vì phải khớp đồng thời mạng tinh thể, vị trí nguyên tử và loại nguyên tử.

Tinh thể có một quy tắc nghiêm ngặt gọi là đối xứng. Tinh thể 3 chiều thuộc một trong 230 nhóm không gian. CDVAE, một mô hình ban đầu, không thể cố định trực tiếp đối xứng nên cấu trúc thường sụp đổ. Mô hình khuếch tán có điều kiện biết đối xứng đã giảm vấn đề này bằng cách tạo một hạt giống nhỏ rồi trải nó ra theo đối xứng. Điều quan trọng là phải tách bạch các nghiên cứu có tên giống nhau, không trộn lẫn chúng.

Nỗ lực tăng tốc cũng tiếp diễn. Flow matching học đường tắt thay vì xóa nhiễu từng chút một. FlowMM và CrystalFlow là các ví dụ tiêu biểu. Chúng tạo tinh thể nhanh bằng cách đi theo đường trắc địa trên không gian cong. Tốc độ này rất hữu ích trong tìm kiếm vật liệu, nơi cần rút ra nhiều ứng viên.

Công nghệ chỉ thay đổi một phần cũng đã phát triển. Inpainting lấp các chỗ trống và khuyết tật. Mirage atom diffusion xử lý số lượng nguyên tử một cách tự do. Nó đặt các vị trí ảo có thể tồn tại hoặc không tồn tại, rồi tự quyết định số lượng. Nhờ đó, nó biểu diễn tốt hơn cấu trúc khuyết tật của vật liệu thực.

Học tăng cường dẫn mô hình sinh tạo về phía các ứng viên tốt. Tính ổn định, tính mới và khả năng tổng hợp được đưa vào phần thưởng. Chuẩn hóa KL và phát lại kinh nghiệm giữ cho quá trình tìm kiếm ổn định. Hiệu quả được đo bằng tỷ lệ SUN và năng lượng trên hull.

Các trường hợp thực tế cho thấy công nghệ này không chỉ là chuyện trên giấy. Trong thiết kế chất điện phân rắn halide, phân tích dữ liệu nối tiếp với thiết kế. Trong thiết kế chất xúc tác entropy cao, thí nghiệm robot và mô hình sinh tạo phối hợp với nhau, giúp giảm mạnh thời gian. Mô hình sinh tạo của Microsoft đã tổng hợp thực tế vật chất do nó thiết kế và xác nhận tính chất.

Ý nghĩ xuyên suốt chương này chỉ có một. Trí tuệ nhân tạo là người dẫn đường thay ta đi qua không gian vật chất rộng lớn. Nó rà soát trong vài ngày những con đường mà con người cả đời không thể đi hết. Nó chọn ra và đưa cho ta những ứng viên có vẻ tốt. Nhưng tính toán và thực nghiệm mới quyết định có tin ứng viên đó hay không. Người dẫn đường thu hẹp con đường, còn con người mở cánh cửa cuối cùng.

Cuối cùng, cần nhắc lại nguyên tắc kiểm chứng. Mô hình sinh tạo đưa ra vô hạn ứng viên. Nhưng không phải cứ tạo ra là tất cả đều có thật. Không dùng những con số chưa được xác nhận. Phải sàng lọc bằng tính toán và xác nhận bằng thực nghiệm. Thiết kế ngược vật liệu chỉ có ý nghĩa khi sinh tạo, tính toán và thực nghiệm cùng vận hành. Trong những lĩnh vực liên quan đến an toàn như tên lửa và thám hiểm không gian, nguyên tắc này càng phải nghiêm ngặt hơn. Cấu trúc do AI vẽ ra không phải là đáp án cuối cùng, mà là ứng viên thực nghiệm đã được chọn lọc kỹ.

## Tài liệu căn cứ liên quan

- 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>

# Chương 3. Phòng thí nghiệm tự hành và hệ thống đa tác nhân khoa học

## Mở đầu

Hãy hình dung cảnh nướng bánh mì ở nhà. Ta quyết định lượng bột mì và nước. Ta trộn bột và chỉnh nhiệt độ lò. Khi bánh chín, ta cắt ra và nếm thử. Nếu bánh quá cứng, ta tăng nước và hạ nhiệt độ. Lần sau, chiếc bánh sẽ khá hơn một chút. Cứ như vậy, ta chọn nguyên liệu, làm, nếm, rồi sửa, lặp đi lặp lại nhiều lần.

Nghiên cứu tìm vật liệu mới cũng giống như vậy. Nhà nghiên cứu quyết định tỷ lệ thành phần. Họ trộn bột và nung ở nhiệt độ cao. Họ đo cấu trúc và tính chất của chất đã tạo ra. Nếu kết quả không tốt, họ thay đổi điều kiện. Phải lặp lại quá trình này hàng trăm lần mới có được một vật liệu tốt.

Vấn đề là thời gian. Nếu con người làm việc này bằng tay, mỗi ngày chỉ làm được vài lần. Ban đêm, thí nghiệm dừng lại. Nếu muốn thay đổi thành phần từng chút một và tìm trong một phạm vi rộng, sẽ mất nhiều năm. Vật liệu dùng cho tên lửa phải chịu các điều kiện khắc khe. Vì vậy số vật liệu ứng viên rất lớn, và khó có thể thử hết chỉ bằng sức người.

Phòng thí nghiệm tự hành (Self-Driving Laboratory) là phòng thí nghiệm giao việc lặp lại này cho máy móc. Trí tuệ nhân tạo (AI) chọn điều kiện cần thử tiếp theo. Robot cân, trộn và nung bột. Thiết bị phân tích đo kết quả. Kết quả đó lại quay về với trí tuệ nhân tạo. Thí nghiệm tự tiếp tục mà không cần con người ra lệnh từng lần. Giống như robot hút bụi tự lau dọn một mình, phòng thí nghiệm tự vận hành suốt đêm.

Cấu trúc trong đó kết quả quay trở lại lựa chọn tiếp theo được gọi là vòng kín (Closed-loop). Vòng kín nghĩa là một mạch khép lại như chiếc vòng. Thí nghiệm, đo lường, học tập và thí nghiệm tiếp theo quay liên tục không đứt đoạn. Chương này bàn về cách phòng thí nghiệm tự hành vận hành. Ta xem trí tuệ nhân tạo chọn thí nghiệm tiếp theo như thế nào. Ta xem máy đọc câu chữ trong bài báo khoa học ra sao. Ta tìm hiểu robot và trí tuệ nhân tạo trao đổi với nhau an toàn như thế nào. Cuối cùng, ta xem các trường hợp phòng thí nghiệm tự hành có thật ở nhiều nơi trên thế giới.

Cũng cần xem vì sao phòng thí nghiệm tự hành đang nổi lên lúc này. Gần đây, trí tuệ nhân tạo dự đoán nhanh tính chất của vật liệu. Tính toán có thể đưa ra hàng nghìn ứng viên tốt mỗi ngày. Nhưng tốc độ chế tạo và kiểm chứng ngoài thực tế vẫn chậm. Dự đoán đã nhanh hơn, nhưng kiểm chứng không theo kịp. Phòng thí nghiệm tự hành lấp khoảng cách này. Robot nhanh chóng chế tạo và kiểm tra các ứng viên do tính toán đưa ra. Có thể xem đó là cây cầu làm khớp tốc độ giữa dự đoán và kiểm chứng.

Nhiều hệ thống xuất hiện trong chương này vẫn còn ở giai đoạn nghiên cứu ban đầu. Có cả các bản tiền ấn phẩm vừa được công bố tại hội nghị. Tiền ấn phẩm là nghiên cứu sơ bộ chưa qua bình duyệt đồng cấp. Vì vậy chương này không khoe các con số hiệu năng, mà xem cả nguyên lý lẫn giới hạn. Mục tiêu của phòng thí nghiệm tự hành không phải là thí nghiệm nhanh, mà là thí nghiệm đáng tin cậy.

## 3.1 Phòng thí nghiệm tự hành vận hành bằng những bộ phận nào

### Nội dung của mục này

Mục này xem toàn bộ phòng thí nghiệm tự hành gồm những bộ phận nào. Có bốn phần lớn. Đó là cái đầu quyết định thí nghiệm tiếp theo, bàn tay thực sự chế tạo, con mắt đo kết quả, và trí nhớ cập nhật điều đã học. Bốn phần này nối với nhau như một chiếc vòng, chuyển kết quả thí nghiệm sang thí nghiệm tiếp theo và tiếp tục vận hành.

Trước hết, ta nhắc lại ý nghĩa của từ vòng kín. Sau đó, ta xem từng phần làm việc gì. Cuối cùng, ta xem vì sao cấu trúc này cần cho nghiên cứu vật liệu tên lửa.

Phòng thí nghiệm tự hành vận hành nhờ bốn phần ăn khớp với nhau. Ở tầng lập kế hoạch, trí tuệ nhân tạo chọn thành phần và điều kiện quy trình cần thử tiếp theo. Ở đây, điều kiện quy trình là nhiệt độ nung, thời gian giữ nhiệt, tốc độ gia nhiệt và khí môi trường. Ở tầng thực thi, robot cân bột, trộn bột, ép thành hạt. Rồi nó đưa vào lò điện để nung. Ở tầng đo lường, các công cụ như thiết bị nhiễu xạ tia X đo cấu trúc của vật chất đã tạo ra. Ở tầng học tập, kết quả đo đi vào mô hình trí tuệ nhân tạo và cập nhật tri thức.

Điều nối bốn phần này là dòng chảy thông tin. Tầng lập kế hoạch gửi bản chỉ định công việc cho tầng thực thi. Bản chỉ định công việc là tài liệu chỉ dẫn phải làm gì, làm bao nhiêu và làm như thế nào. Tầng thực thi gửi mẫu đã chế tạo sang tầng đo lường. Tầng đo lường gửi các giá trị như độ tinh khiết pha và hằng số mạng sang tầng học tập. Độ tinh khiết pha nghĩa là tinh thể mong muốn đã xuất hiện tinh khiết đến mức nào. Tầng học tập dùng các giá trị đó để chỉnh kế hoạch tiếp theo. Khi thông tin đi hết một vòng như vậy, một thí nghiệm kết thúc.

Sức mạnh của vòng kín nằm ở sự lặp lại nhanh. Dù một thí nghiệm thất bại, dữ liệu vẫn còn. Điều kiện thất bại cũng cho biết điều gì không hiệu quả. Trí tuệ nhân tạo học cả thành công lẫn thất bại. Vì vậy điều kiện tiếp theo dần tốt hơn. Con người chỉ cần xem hướng lớn vài ngày một lần, còn những lặp lại chi tiết do máy đảm nhiệm.

Ta xem vì sao vật liệu tên lửa cần cấu trúc này. Vòi phun và buồng đốt tên lửa chịu điều kiện cực hạn. Nhiệt độ tăng đến hàng nghìn độ. Áp suất và dao động cũng lớn. Vật liệu dùng ở những nơi như vậy phải đáp ứng điều kiện rất khắt khe. Chỉ cần thay đổi thành phần một chút, tính chất có thể đổi rất nhiều. Vì vậy có hàng chục nghìn ứng viên cần thử. Con người không thể tìm hết không gian rộng này bằng tay. Phòng thí nghiệm tự hành có thể nhanh chóng thu hẹp không gian rộng đó.

Năm 2026, các nghiên cứu nhằm đưa phòng thí nghiệm tự hành lên một bậc cao hơn tiếp tục xuất hiện. Hiệp hội Hóa học Hoàng gia Anh đã hệ thống hóa khái niệm phòng thí nghiệm tự hành 2.0. Thế hệ phòng thí nghiệm tự hành đầu tiên chỉ giải một vấn đề đã định sẵn. Thế hệ tiếp theo muốn xử lý linh hoạt nhiều loại thiết bị và vấn đề. Một bài tổng quan năm 2025 khảo sát rộng về phòng thí nghiệm tự hành cũng đã xuất hiện. Bài báo này xem cả mức độ trưởng thành của công nghệ lẫn những rào cản còn lại. Truyền thông của Hiệp hội Hóa học Hoa Kỳ cho biết phòng thí nghiệm tự hành đang thay đổi cách các nhà hóa học làm việc.

Cấu trúc này cũng có giới hạn rõ ràng. Trang bị thiết bị tốn rất nhiều tiền. Việc nối robot với thiết bị phân tích cũng khó. Ngay cả cân bột thật chính xác cũng không dễ. Tự động nhận biết thí nghiệm thất bại cũng khó. Vì vậy phòng thí nghiệm tự hành hiện nay chỉ phù hợp tốt với một số nhóm vật liệu nhất định. Chưa có phòng thí nghiệm vận năng xử lý được mọi vật liệu. Các mục còn lại của chương này xem cách vượt qua từng rào cản như vậy.

### 3.1.1 Học chủ động Bayes chọn thí nghiệm tiếp theo

Tiểu mục này bàn về phần tương ứng với cái đầu của phòng thí nghiệm tự hành. Trọng tâm là trí tuệ nhân tạo chọn thí nghiệm tiếp theo như thế nào. Phương pháp được dùng ở đây là học chủ động Bayes. Ta sẽ giải thích bằng ngôn ngữ dễ hiểu phương pháp này là gì và vì sao nó làm giảm số thí nghiệm.

Giả thuyết là suy nghĩ về việc vì sao thí nghiệm tiếp theo sẽ tốt. Một phòng thí nghiệm tự hành tốt đặt ra giả thuyết tốt. Nếu thử bừa mọi điều kiện, thời gian sẽ bị lãng phí. Phải chọn nơi có khả năng thành công để thử. Học chủ động Bayes (Bayesian Active Learning) là phương pháp giúp việc lựa chọn này. Học chủ động nghĩa là cách máy tự chọn dữ liệu để học.

Phương pháp này trước hết tạo một mô hình thay thế (Surrogate Model). Mô hình thay thế là mô hình tính toán nhỏ bất chước thí nghiệm đắt tiền. Mô hình này được tạo từ các kết quả thí nghiệm đã có đến nay. Mô hình thay thế dự đoán tính chất của các điều kiện chưa thử. Dự đoán đưa ra hai giá trị cùng lúc. Một là hiệu năng dự kiến. Giá trị còn lại là mức độ bất định của dự đoán đó.

Điểm cốt lõi là xem hai giá trị này cùng nhau. Nơi có hiệu năng dự kiến cao đáng để thử. Nơi có độ bất định lớn cũng đáng để thử. Vì vật liệu tốt ngoài dự đoán có thể đang ẩn ở nơi ta chưa biết rõ. Nếu chỉ đuổi theo hiệu năng, ta sẽ quanh quẩn ở chỗ đã biết. Nếu chỉ đuổi theo độ bất định, ta sẽ lang thang ở nơi vô ích. Một phương pháp tốt giữ cân bằng giữa hai khuynh hướng này. Ta gọi đó là cân bằng giữa khám phá và khai thác. Ở đây, khám phá là xem nơi chưa biết, còn khai thác là đào sâu nơi đã biết.

Hàm thu nhận (Acquisition Function) là công cụ đo sự cân bằng này bằng con số. Hàm thu nhận chấm điểm từng điều kiện ứng viên. Nếu xếp theo điểm từ cao xuống thấp, vị trí đứng đầu sẽ trở thành thí nghiệm tiếp theo. Một hàm thu nhận tiêu biểu là mức cải thiện kỳ vọng (Expected Improvement). Mức cải thiện kỳ vọng đo xem thí nghiệm tiếp theo có thể tốt hơn kết quả tốt nhất hiện tại bao nhiêu. Một phương pháp khác là cận tin cậy trên (Upper Confidence Bound). Phương pháp này cho phép xem cùng lúc hiệu năng cao và vùng chưa biết. Phòng thí nghiệm tự hành dùng các điểm số như vậy để quyết định thành phần, nhiệt độ và thời gian tiếp theo.

Ta xem vì sao phương pháp này quan trọng trong vật liệu tên lửa. Gốm siêu cao nhiệt hoặc hợp kim đặc biệt tốn chi phí lớn cho mỗi thí nghiệm. Chúng dùng nguyên liệu đắt tiền và phải nung trong thời gian dài. Vì vậy giảm số thí nghiệm chính là tiết kiệm tiền và thời gian. Học chủ động Bayes chỉ chọn những điều kiện có khả năng thành công để thử. Số thí nghiệm giảm mạnh so với việc thử tất cả một cách mù quáng. Một số phòng thí nghiệm tự hành báo cáo rằng số thí nghiệm cần thiết đã giảm hàng chục lần.

Hiệu quả đó hiện rõ trong các trường hợp thực tế. Năm 2026, một nhóm nghiên cứu đã tạo ra phòng thí nghiệm tự hành để nuôi màng mỏng. Màng mỏng được tạo bằng cách xếp nguyên tử từng lớp một. Phòng thí nghiệm này đọc hoa văn nhiễu xạ điện tử bằng camera thời gian thực. Hoa văn nhiễu xạ là tín hiệu cho biết nguyên tử có xếp tốt hay không. Trí tuệ nhân tạo nhìn hoa văn này và chọn điều kiện tiếp theo. Kết quả là số thí nghiệm cần thiết giảm mạnh. Nhóm nghiên cứu cho biết số thí nghiệm giảm hơn ba mươi lần so với khi quét từng điều kiện một. Họ đã tìm được điều kiện tốt nhanh hơn rất nhiều. Nghiên cứu này cũng còn ở giai đoạn đầu, nhưng cho thấy rõ hướng đi.

Năm 2026, các nghiên cứu điều chỉnh phương pháp này cho phù hợp với phòng thí nghiệm tự hành diễn ra sôi động. Một tiền án phẩm đã hệ thống hóa tối ưu hóa Bayes cho phòng thí nghiệm vật liệu tự hành. Bài viết giải thích từ thuật toán đến dòng công việc thí nghiệm thực tế. Một tiền án phẩm khác đưa ra phương pháp tự thu hẹp không gian tìm kiếm. Phương pháp đó dùng học chủ động để giảm khoảng một nửa không gian ứng viên mà vẫn giữ phần lớn các nghiệm tốt. Làm như vậy giúp gặp vật liệu tốt nhanh hơn. Một hướng dẫn thực hành về tối ưu hóa Bayes cho khoa học thực nghiệm cũng xuất hiện trong năm 2026.

Phương pháp này cũng có giới hạn. Lúc đầu, mô hình thay thế có ít dữ liệu nên dự đoán còn thô. Vài thí nghiệm đầu có thể đi sai hướng. Nếu không gian thành phần quá rộng, khối lượng tính toán tăng lên. Nếu phép đo có nhiễu nhiều, dự đoán sẽ dao động. Vì vậy con người không tin nguyên xi dự đoán của mô hình thay thế. Họ tham khảo dự đoán, nhưng luôn xác nhận bằng đo lường thực tế. Ngay cả trong phòng thí nghiệm tự hành, sự giám sát của con người vẫn còn.

## 3.2 Trích xuất phương pháp tổng hợp trong bài báo và chuyển thành kế hoạch an toàn

### Nội dung của mục này

Mục này bàn về hai việc. Một là làm cho máy đọc được phương pháp tổng hợp trong bài báo khoa học. Hai là chuyển phương pháp đó thành kế hoạch để robot thực hiện an toàn.

Con người đọc bài báo và hình dung trình tự thí nghiệm trong đầu. Máy không làm được như vậy. Vì thế phải chuyển câu chữ thành dạng máy có thể xử lý. Mục này xem quá trình chuyển đổi đó và kế hoạch an toàn.

Trong bài báo khoa học, phương pháp tổng hợp được viết bằng câu văn. Bài báo ghi trộn loại bột nào với lượng bao nhiêu. Bài báo ghi nung ở bao nhiêu độ trong bao nhiêu giờ. Bài báo ghi xử lý trong loại khí nào. Trong các bài báo tích lũy suốt hàng chục năm có hàng triệu phương pháp như vậy. Tri thức này là vốn lớn của nghiên cứu vật liệu mới. Nhưng phần lớn nằm rải rác trong cách nói của con người.

Vấn đề là mỗi người viết những câu này theo một cách khác nhau. Có bài báo viết rằng đã nung qua đêm. Có bài báo chỉ viết rằng đã xử lý ở nhiệt độ cao. Họ không ghi chính xác qua đêm là mấy giờ, nhiệt độ cao là bao nhiêu độ. Con người lấp chỗ trống này bằng kinh nghiệm. Máy không làm được như vậy. Vì thế cần chuyển các câu rải rác thành những con số chính xác.

Trong nghiên cứu vật liệu tên lửa, việc này có giá trị lớn. Phương pháp tổng hợp gồm siêu cao nhiệt và hợp kim đặc biệt nằm rải rác trong các bài báo cũ. Nếu con người đọc và sắp xếp tất cả, sẽ mất rất nhiều thời gian. Nếu máy tự động trích xuất và sắp xếp câu văn, thời gian giảm đáng kể. Phòng thí nghiệm tự hành chuyển ngay các phương pháp thu được như vậy thành chỉ dẫn cho robot, qua đó rút ngắn khoảng cách từ bài báo đến thí nghiệm.

Khi tri thức này nằm rải rác, nghiên cứu lặp lại sẽ xuất hiện. Một nhà nghiên cứu khác lại thử điều kiện mà người đi trước đã thất bại. Họ không biết phương pháp đã thành công và phải tìm lại từ đầu. Nếu gom các phương pháp trong bài báo vào một nơi, sự lãng phí này giảm đi. Con người khó đọc hết hàng triệu bài. Vì vậy việc máy trích xuất và sắp xếp câu văn có giá trị lớn. Phương pháp được sắp xếp tốt trở thành điểm xuất phát tốt cho phòng thí nghiệm tự hành.

Ở đây có hai bước. Một là trích xuất phương pháp từ câu văn. Mô hình ngôn ngữ lớn (Large Language Model) đảm nhiệm việc này. Mô hình ngôn ngữ lớn là trí tuệ nhân tạo học và xử lý văn bản của con người. Bước còn lại là chuyển phương pháp đã trích xuất thành kế hoạch an toàn. Ngôn ngữ lập kế hoạch đảm nhiệm việc này. Hai tiểu mục tiếp theo sẽ xem kỹ từng bước.

### 3.2.1 Mô hình ngôn ngữ lập kế hoạch tổng hợp từ bài báo (MSP-LLM)

Tiểu mục này bàn về cách rút ra phương pháp tổng hợp từ các bài báo. Trường hợp tiêu biểu là nghiên cứu MSP-LLM. Ta sẽ xem nghiên cứu này muốn giải quyết vấn đề gì và vận hành theo những bước nào.

MSP-LLM là một nghiên cứu về mô hình ngôn ngữ lớn dùng để lập kế hoạch tổng hợp vật liệu. Tên này xuất phát từ tiêu đề: một khung mô hình ngôn ngữ lớn tích hợp cho kế hoạch tổng hợp vật liệu hoàn chỉnh. Đây là nghiên cứu ban đầu được công bố vào năm 2026. Nhóm nghiên cứu gồm Roh Hee-woong, Na Kyung-soo, Lee Nam-kyung và Park Chan-young. Nghiên cứu này là bản tiền ấn phẩm đã được trình bày tại hội nghị. Vì vậy, nên xem đây là một nỗ lực đi trước, chứ không phải tiêu chuẩn trong thực tiễn.

Vấn đề mà nghiên cứu này giải quyết là kế hoạch tổng hợp. Kế hoạch tổng hợp là việc lập ra toàn bộ trình tự để tạo ra vật liệu mục tiêu. Việc này được chia thành hai vấn đề nhỏ. Một là dự đoán tiền chất. Tiền chất là nguyên liệu được đưa vào lúc đầu để tạo ra vật liệu mục tiêu. Hai là dự đoán thao tác tổng

hợp. Thao tác tổng hợp là trình tự các hành động như trộn, nung và làm nguội. Phải giải quyết cả hai vấn đề cùng lúc thì mới có được một kế hoạch hoàn chỉnh.

MSP-LLM nối hai vấn đề này thành một luồng thống nhất. Trước hết, nó xác định vật liệu mục tiêu thuộc loại nào. Loại này trở thành tiêu chí phán đoán trung gian. Sau đó, nó chọn tiền chất phù hợp với loại đó. Cuối cùng, nó lập trình tự thao tác phù hợp với các tiền chất ấy. Như vậy, nó tạo ra một chuỗi nối từ loại vật liệu đến nguyên liệu, rồi từ nguyên liệu đến trình tự. Khi nối thành chuỗi, ta có được một kế hoạch hợp lý về mặt hóa học từ đầu đến cuối.

Có thể hiểu luồng này qua một ví dụ dễ hơn. Việc quyết định sẽ nấu món gì tương ứng với loại vật liệu mục tiêu. Chẳng hạn như quyết định nấu canh kimchi. Tiếp theo, việc quyết định mua những gì tương ứng với dự đoán tiền chất. Đó là chọn kimchi, thịt heo và đậu phụ. Cuối cùng, việc quyết định nấu theo thứ tự nào tương ứng với dự đoán thao tác. Đó là lập trình tự xào, đổ nước và đun sôi. MSP-LLM nối ba bước này để hoàn tất kế hoạch.

Ta xét vì sao kiểu lập kế hoạch tự động này có giá trị trong vật liệu tên lửa. Muốn tạo hợp kim mới hoặc gốm mới thì ngay từ khâu chọn nguyên liệu đã khó. Cùng một mục tiêu, nhưng nếu đổi nguyên liệu thì kết quả sẽ khác. Nếu đổi trình tự nung thì kết quả cũng khác. Con người cần rất nhiều thời gian để cân nhắc hết các tổ hợp này. Bộ lập kế hoạch tự động nhanh chóng đưa ra điểm xuất phát tốt dựa trên tri thức trong các bài báo. Phòng thí nghiệm tự hành có thể bắt đầu thí nghiệm từ điểm xuất phát đó.

Phương pháp này cũng có giới hạn rõ ràng. Mô hình ngôn ngữ lớn đôi khi bịa ra sự thật không tồn tại. Điều này được gọi là ảo giác. Nếu đưa nguyên liệu hoặc điều kiện bịa ra vào thí nghiệm mà không kiểm tra thì sẽ lãng phí tài nguyên. Nó cũng có thể đưa ra tổ hợp nguy hiểm. Vì vậy, kế hoạch được rút ra nhất thiết phải được con người hoặc công cụ kiểm chứng khác xác nhận. Trước hết phải xét xem nó có phù hợp với quy tắc hóa học hay không, có an toàn hay không. Lập kế hoạch tự động là công cụ hỗ trợ con người, không phải công cụ đẩy con người ra ngoài.

### 3.2.2 Ngăn tai nạn robot bằng ngôn ngữ lập kế hoạch (PDDL)

Tiểu mục này bàn về việc chuyển phương pháp tổng hợp đã rút ra thành kế hoạch robot an toàn. Công cụ được dùng ở đây là ngôn ngữ lập kế hoạch PDDL. Ta sẽ giải thích bằng lời dễ hiểu ngôn ngữ này ngăn tai nạn robot như thế nào.

Kế hoạch thí nghiệm không chỉ là một bảng thứ tự. Mỗi hành động đều có điều kiện phải được đáp ứng trước. Lò điện phải trống thì mới đưa mẫu vào. Lò điện phải nguội thì mới lấy mẫu ra. Việc cân bột phải xong thì mới trộn. Nếu vi phạm các điều kiện trước sau như vậy thì tai nạn sẽ xảy ra. Nếu robot tùy tiện mở cửa lò điện đang nóng thì rất nguy hiểm.

PDDL(Planning Domain Definition Language) là ngôn ngữ dùng để ghi các kế hoạch như vậy. Dịch sang tiếng Việt là ngôn ngữ định nghĩa miền lập kế hoạch. Ngôn ngữ này ghi hai điều cho mỗi hành động. Một là điều kiện tiên quyết. Nó ghi những gì phải đúng trước khi thực hiện hành động này. Hai là hiệu quả thực thi. Nó ghi điều gì thay đổi khi thực hiện hành động này. Ví dụ, điều kiện tiên quyết của hành động nung là lò điện đang ở trạng thái chờ và nhiệt độ thấp. Và hiệu quả thực thi sau khi hoàn tất hành động này là mẫu đã được xử lý nhiệt.

Khi ghi các điều kiện như vậy, máy sẽ tự tìm trình tự an toàn. Hành động không đáp ứng điều kiện sẽ không được thực hiện. Lệnh tiếp cận lò điện đang nóng sẽ bị chặn. Mẫu chưa cân xong sẽ không chuyển sang bước trộn. Quy tắc này trở thành hàng rào ngăn tai nạn robot từ trước. Có thể xem đây là việc chuyển các quy tắc an toàn phòng thí nghiệm mà con người tuân thủ sang ngôn ngữ máy.

Tuy nhiên, hàng rào này không tự động bảo đảm an toàn. Điều kiện tiên quyết chỉ hoạt động khi con người nhập chính xác từng điều kiện một. Nếu bỏ sót hoặc ghi sai điều kiện, hàng rào sẽ có lỗ hổng. Vì

vậy, PDDL không phải là thiết bị bảo đảm thực thi an toàn. Nó là công cụ để con người lường trước rủi ro và chuyển rủi ro ấy thành quy tắc. Độ dày của các quy tắc chỉ dày đến mức phán đoán của con người.

Kế hoạch này có thể biểu diễn bằng hình vẽ. Đặt mỗi hành động thành một điểm và nối quan hệ trước sau bằng mũi tên. Hình vẽ như vậy được gọi là đồ thị có hướng không chu trình. Nghĩa là mũi tên chỉ chảy theo một hướng và không quay lại. Nhìn vào hình này, ta biết việc nào có thể làm đồng thời. Ví dụ, trong lúc nung một mẫu thì cân bột cho mẫu khác. Khi đặt song song những việc không cản trở nhau, toàn bộ phòng thí nghiệm sẽ vận hành nhanh hơn chừng ấy.

Trong thí nghiệm vật liệu tên lửa, kế hoạch an toàn này có giá trị lớn. Lò điện siêu cao nhiệt nếu xử lý sai sẽ rất nguy hiểm. Nguyên liệu có tính phản ứng mạnh không được tiếp xúc với không khí. Nhiều thiết bị cùng chuyển động trong một phòng thí nghiệm. Trong một phòng thí nghiệm vận hành suốt đêm không có người, cần phải có quy tắc ngăn tai nạn. Ngôn ngữ lập kế hoạch như PDDL giúp dệt các quy tắc này thật chặt. Tốc độ của phòng thí nghiệm tự hành chỉ có ý nghĩa khi đặt trên thiết bị an toàn ấy.

Cách làm này cũng có giới hạn. Rất khó ghi toàn bộ điều kiện của thế giới thành quy tắc. Một tình huống chưa kịp ghi có thể thật sự xảy ra. Nếu có quá nhiều quy tắc, lượng tính toán để tìm kế hoạch sẽ tăng lên. Hồng học không dự đoán trước không thể bị chặn chỉ bằng quy tắc. Vì vậy, chỉ ngôn ngữ lập kế hoạch là chưa đủ. Cần có thiết bị khác cùng theo dõi trạng thái thiết bị theo thời gian thực, và thiết bị đó sẽ được bàn ở mục sau.

### 3.3 Truyền thông tiêu chuẩn trí tuệ nhân tạo với thiết bị thí nghiệm (MCP)

Mục này bàn về cách trí tuệ nhân tạo và thiết bị thí nghiệm trò chuyện với nhau. Mô hình ngôn ngữ xử lý câu văn. Cân, lò điện và cánh tay robot đòi hỏi lệnh chính xác. Cần có một tầng trung gian nối hai bên. Mục này xem xét tầng trung gian đó. Ta sẽ cùng xem quy tắc truyền thông tiêu chuẩn và thiết bị an toàn đặt trên nó.

Mô hình ngôn ngữ xử lý tốt chữ viết của con người. Nó cũng hiểu câu yêu cầu tăng nhiệt độ từ từ. Nhưng lò điện không hiểu câu văn. Lò điện phải nhận tín hiệu chính xác rằng cần tăng bao nhiêu độ mỗi phút. Cân và cánh tay robot cũng như vậy. Vì thế cần có một bộ phiên dịch để chuyển câu văn sang ngôn ngữ của thiết bị.

Trước đây, người ta tạo riêng bộ phiên dịch cho từng thiết bị. Bộ phiên dịch cho cân, bộ phiên dịch cho lò điện, bộ phiên dịch cho robot được viết riêng từng cái. Nếu thay một thiết bị, bộ phiên dịch cũng phải viết lại. Nhiều phòng thí nghiệm lặp lại cùng một việc. Đó là một sự lãng phí lớn. Vì không có tiêu chuẩn nên cũng khó chia sẻ mã của nhau.

Quy tắc truyền thông tiêu chuẩn nhằm giải quyết vấn đề này. Khi có quy tắc tiêu chuẩn, mỗi thiết bị được kết nối theo cùng một cách. Thiết bị mới chỉ cần phù hợp với quy tắc là có thể gắn vào ngay. Các phòng thí nghiệm cũng dễ chia sẻ mã với nhau hơn. Nghiên cứu gần đây về phòng thí nghiệm tự hành tìm tiêu chuẩn này ở Model Context Protocol.

Có một quy tắc mở gọi là Model Context Protocol. Viết tắt là MCP. Protocol là cách truyền thông mà các bên đã thỏa thuận với nhau. MCP là quy tắc trí tuệ nhân tạo với các công cụ phần mềm. Nó cho trí tuệ nhân tạo biết theo cách tiêu chuẩn rằng có những công cụ nào. Trí tuệ nhân tạo tự nhận ra và gọi các công cụ đó.

Ở đây có một điểm dễ hiểu nhầm cần nói rõ. MCP là quy tắc cấp cao nối các phần mềm với nhau. Nó không trực tiếp vận hành thiết bị vật lý như cân và lò điện. Việc quay motor và van do driver riêng của từng thiết bị đảm nhiệm. Driver là chương trình cấp thấp trực tiếp xử lý tín hiệu thiết bị. MCP nằm trên driver này và nối nó với trí tuệ nhân tạo. Vì vậy, chỉ có MCP thì thiết bị không thể gắn vào ngay. MCP

cũng không phải tiêu chuẩn điều khiển thời gian thực công nghiệp, và cũng không phải tiêu chuẩn an toàn. Muốn gắn thiết bị, bên dưới vẫn phải đặt riêng tầng driver.

Một trường hợp đưa quy tắc này vào phòng thí nghiệm tự hành là NIMO controller. Đây là nghiên cứu ban đầu được công bố vào năm 2026. Nhóm nghiên cứu gồm Yoshikawa Naruki và Tamura Ryo. Nghiên cứu này bộc lộ nhiều chức năng của phòng thí nghiệm tự hành dưới dạng MCP server. Ở đây, NIMO là engine quyết định thí nghiệm tiếp theo. Controller là tầng điều phối nối engine đó với nhiều công cụ bằng MCP. Trong trường hợp này, trí tuệ nhân tạo điều khiển một phòng thí nghiệm tự hành để khớp màu. Nó tự thay đổi thành phần và thí nghiệm cho đến khi đạt màu mục tiêu. Điều mà nghiên cứu này thật sự chứng minh là việc điều phối này và thí nghiệm khớp màu.

Mục đích của NIMO mà nghiên cứu này nêu ra rất rõ. Trong phòng thí nghiệm tự hành có nhiều phương pháp tối ưu hóa. Cũng có nhiều loại công cụ. Việc nối từng phương pháp với từng công cụ là một nút nghẽn lớn. NIMO controller nối hai bên theo một cách chung. Con người có thể lập luồng thí nghiệm mà không cần viết mã. Đó là cách chọn công cụ trên màn hình rồi nối chúng với nhau. Tuy nhiên, nghiên cứu này vẫn ở giai đoạn đầu nên còn sớm để xem nó là tiêu chuẩn được dùng rộng rãi.

Ta xét vì sao tầng truyền thông này quan trọng trong phòng thí nghiệm vật liệu tên lửa. Trong phòng thí nghiệm vật liệu có nhiều thiết bị của các công ty khác nhau trộn lẫn với nhau. Nhà sản xuất cân, lò điện và thiết bị nhiễu xạ đều khác nhau. Muốn một trí tuệ nhân tạo điều khiển tất cả, phải có ngôn ngữ chung. Quy tắc truyền thông tiêu chuẩn buộc nhiều thiết bị này vào một luồng. Khi đưa thiết bị mới vào, cũng có thể gắn vào mà không cần sửa đổi lớn. Việc mở rộng hoặc thay đổi phòng thí nghiệm trở nên dễ hơn.

Truyền thông tiêu chuẩn còn có một lợi ích khác. Đó là con người có thể lập thí nghiệm dù không biết viết mã. Trí tuệ nhân tạo tự nhận ra có những công cụ nào. Nó có thể bày các công cụ đó trên màn hình. Con người chỉ cần chọn công cụ rồi nối chúng theo thứ tự. Họ lập luồng như một bức tranh: cân bằng cân, trộn, nung, rồi đo bằng nhiễu xạ. Khi ngưỡng lập thí nghiệm thấp xuống, nhiều nhà nghiên cứu hơn sẽ dùng phòng thí nghiệm tự hành.

Tầng này cũng có giới hạn. Quy tắc truyền thông vẫn còn mới và đang định hình. Mức độ tuân thủ quy tắc khác nhau tùy từng công ty thiết bị. Thiết bị cũ khó chỉnh cho phù hợp với quy tắc. Nếu truyền thông bị ngắt hoặc tín hiệu đến chậm, thí nghiệm sẽ lệch. Vì vậy, bản thân tầng truyền thông cũng cần thiết bị an toàn. Thiết bị này không thực hiện lệnh ngay, mà kiểm tra trước. Đó chính là máy trạng thái.

### 3.3.1 Tuyến an toàn cuối cùng do máy trạng thái bảo vệ

Tiểu mục này bàn về máy trạng thái bảo vệ an toàn thiết bị. Ta sẽ xem máy trạng thái là gì và nó chặn lệnh nguy hiểm như thế nào. Ta cũng xem vì sao thiết bị này là tuyến an toàn cuối cùng trong phòng thí nghiệm tự hành. Máy trạng thái được bàn ở đây không phải là chức năng mà một sản phẩm cụ thể đã có đầy đủ. NIMO controller đã nói ở trên là nghiên cứu tập trung vào việc điều phối để nối các công cụ. Máy trạng thái và interlock là thiết kế an toàn chung phải đặt riêng trên tầng điều phối như vậy. Cuốn sách này xem thiết kế an toàn đó là một tầng mà phòng thí nghiệm tự hành phải có.

Máy trạng thái là phương pháp chia trạng thái hiện tại của thiết bị thành các ô đã định trước. Máy trạng thái hữu hạn (Finite State Machine) định trước các ô có thể có. Ví dụ có các ô như chờ, chuẩn bị, cân đo, chuyển mẫu, phản ứng, phân tích, lỗi, dừng khẩn cấp. Thiết bị luôn nằm trong một trong các ô này. Đường đi từ một ô sang ô tiếp theo cũng được định trước.

Điểm cốt lõi là chỉ thực hiện hành động được phép trong một trạng thái. Ở trạng thái phản ứng, lò điện đang nóng. Lúc này, lệnh để cánh tay robot mở cửa lò điện sẽ bị chặn. Chỉ khi nhận được tín hiệu rằng nhiệt độ đã nguội xuống dưới giá trị an toàn thì mới mở cửa. Cảm biến xác nhận tín hiệu này. Trước khi

có tín hiệu xác nhận, lệnh bị chặn ở bước máy trạng thái. Sự chặn này được gọi là interlock. Interlock là khóa an toàn sẽ khóa lại khi điều kiện không phù hợp.

Có thể hiểu vì sao cách này an toàn qua một ví dụ dễ hiểu. Lò vi sóng dừng hoạt động khi cửa mở. Nó chỉ chạy khi cửa đóng. Khi cửa đang mở, dù bấm nút bao nhiêu cũng không bật lên. Đó là hạn chế hành động theo trạng thái. Máy trạng thái của phòng thí nghiệm tự hành cũng vận hành theo cùng nguyên lý. Ở trạng thái nguy hiểm, lệnh nguy hiểm hoàn toàn không được thực hiện.

Lệnh bằng câu văn của trí tuệ nhân tạo nhất thiết phải đi qua quy tắc này. Dù trí tuệ nhân tạo ra lệnh tự tin đến đâu, quy tắc vẫn đứng trước. Chỉ lệnh vượt qua quy tắc mới đi đến driver thiết bị. Driver là chương trình vận hành motor và van thật. Lệnh bị quy tắc giữ lại sẽ không được thực hiện, chỉ để lại bản ghi. Bản ghi này về sau trở thành tư liệu để tìm nguyên nhân.

Trong thí nghiệm vật liệu tên lửa, đường an toàn này rất có giá trị. Vì người ta phải xử lý lò điện siêu cao nhiệt và nguyên liệu có tính phản ứng. Một sai sót nhỏ cũng có thể dẫn đến tai nạn lớn. Ban đêm, khi không có người, rủi ro còn cao hơn. Máy trạng thái là cửa an toàn cuối cùng giữa trí tuệ nhân tạo và thiết bị. Ngay cả khi trí tuệ nhân tạo mắc lỗi, thiết bị cũng không thực hiện thao tác nguy hiểm. Sự tự do của phòng thí nghiệm tự hành chỉ được cho phép bên trong cửa này.

Thiết bị này cũng có giới hạn. Những trạng thái không thể định trước thì khó xử lý. Sự cố ngoài dự kiến không vừa với những ô đã định sẵn. Nếu cảm biến gửi tín hiệu sai, phán đoán cũng sai. Vì vậy, chỉ có máy trạng thái là chưa đủ. Cần có con mắt quan sát thiết bị theo thời gian thực và nhận ra bất thường. Con mắt đó là hợp nhất cảm biến, nội dung sẽ được xem ở mục tiếp theo.

### 3.4 Khung thực thi robot chịu được môi trường thí nghiệm xáo trộn (PUDA)

Mục này bàn về những khó khăn robot gặp trong môi trường thí nghiệm thực tế. Phòng thí nghiệm khác với nhà máy. Bột chảy, vón cục và mang tĩnh điện. Chén nung lệch đi từng chút một. Cảm biến lẫn nhiễu. Mục này xem xét khung thực thi robot có thể chịu được môi trường xáo trộn như vậy. Trường hợp tiêu biểu là nghiên cứu PUDA.

Robot phòng thí nghiệm gặp công việc khó hơn robot nhà máy. Trong nhà máy, cùng một linh kiện đi đến vị trí đã định. Trong phòng thí nghiệm, tính chất của bột mỗi lần lại khác. Có loại bột chảy tốt. Có loại bột vón lại và làm tắc vòi phun. Có loại bột mang tĩnh điện nên dính khắp nơi. Chén nung mỗi lần đặt xuống cũng lệch vị trí từng chút một.

Sự xáo trộn như vậy được gọi là nhiễu loạn. Nhiễu loạn là trở ngại không có trong kế hoạch. Khi có nhiễu loạn, robot không thể làm việc chỉ bằng động tác đã định. Robot phải có mắt và cảm giác. Mắt là camera, còn cảm giác là cảm biến lực. Robot đọc thông tin này theo thời gian thực và sửa động tác.

PUDA là một trường hợp xử lý môi trường thực thi robot như vậy. Tên đầy đủ là Physical Unified Device Architecture, tức cấu trúc thiết bị thống nhất tích hợp vật lý. Nó cũng được gọi là harness phần cứng thân thiện với trí tuệ nhân tạo cho phòng thí nghiệm tự hành. Harness là khung gộp nhiều thiết bị lại để xử lý như một hệ thống. Đây là nghiên cứu giai đoạn đầu được công bố năm 2026. Nhóm nghiên cứu gồm Zenkun Ren, Hongzhao Tan, Jiaen Yi và Kedar Hippalgaonkar. Điều nghiên cứu này thật sự cho thấy là môi trường thực thi dòng lệnh dễ cho trí tuệ nhân tạo xử lý và hệ thống ghi chép bên trong nó. Hợp nhất cảm biến và phục hồi lỗi được trình bày bên dưới không phải là thành quả hiệu năng mà bài báo này đã chứng minh. Đó là hướng mà loại khung thực thi này muốn tiến tới.

Cách làm của PUDA hơi khác các hệ thống trước đó. Nó không đặt màn hình dành cho con người lên trước. Thay vào đó, nó tạo môi trường dòng lệnh dễ cho tác tử trí tuệ nhân tạo xử lý. Mỗi thiết bị hiện ra như một công cụ có thể tìm thấy trong dòng lệnh. Lệnh và phản hồi được lưu thành bản ghi chuẩn. Mẫu

nào nhận lệnh nào và cho ra kết quả nào được nối lại như một chuỗi. Lịch sử này được gọi là provenance. Provenance là lai lịch cho biết dữ liệu đến từ đâu và bằng cách nào.

Hãy xem vì sao bản ghi lai lịch này có giá trị trong nghiên cứu vật liệu tên lửa. Linh kiện hàng không vũ trụ đòi hỏi truy xuất chất lượng nghiêm ngặt. Phải để lại toàn bộ thông tin về nguyên liệu nào đã được dùng và được chế tạo trong điều kiện nào. Sau này nếu có vấn đề, người ta tìm nguyên nhân bằng những bản ghi này. Hệ thống như PUDA tự động để lại lai lịch đó. Nó ít bỏ sót hơn khi con người ghi bằng tay. Dữ liệu do phòng thí nghiệm tự hành tạo ra cũng có thể được tin cậy và dùng được.

PUDA chọn cách dòng lệnh là có lý do. Màn hình dành cho con người nhìn thì tốt. Nhưng tác tử trí tuệ nhân tạo xử lý thì bất tiện. Nút trên màn hình khó cho máy bấm. Công cụ dòng lệnh dễ để trí tuệ nhân tạo gọi trực tiếp. Trí tuệ nhân tạo quan sát tình huống, phán đoán và ra lệnh. Thiết bị thực thi lệnh đó một cách chính xác và có thể truy hồi. Có thể xem như đã tách bộ não lập kế hoạch thí nghiệm của con người khỏi bàn tay vận hành thiết bị. Khi tách như vậy, dù thay bộ não thì bàn tay vẫn dùng như cũ.

Cách làm này cũng có giới hạn. Môi trường dòng lệnh bất tiện để con người nhìn ngay. Mỗi thiết bị cần được tạo công cụ dòng lệnh riêng. Thiết bị cũ khó khớp với khung như vậy. Muốn để lại bản ghi chuẩn thì phải tốn sức cho lưu trữ và quản lý. Nghiên cứu này cũng vẫn ở giai đoạn đầu. Muốn được dùng rộng rãi trong nhiều phòng thí nghiệm thì cần thêm thời gian.

### 3.4.1 Gộp nhiều cảm biến để phán đoán và đảo ngược thất bại

Tiểu mục này bàn về hợp nhất cảm biến, tức việc gộp thông tin từ nhiều cảm biến để phán đoán. Sau đó xem quá trình phục hồi tự đảo ngược thất bại. Ta sẽ xem vì sao hai chức năng này giữ được chất lượng thí nghiệm. Hai chức năng này là những phương pháp đã được nghiên cứu rộng rãi trong robot học. Khung thực thi như PUDA ở mục trước là cái khuôn muốn chứa các phương pháp đó. Ở đây, ta xem nguyên lý của phương pháp, chứ không xem thành quả của một hệ thống riêng lẻ.

Hợp nhất cảm biến là việc gộp thông tin từ nhiều cảm biến để phán đoán. Camera nhìn vị trí và hình dạng. Cảm biến lực nhìn lực tác động lên cổ tay robot. Cảm biến nhiệt độ và cảm biến áp suất xem thiết bị có an toàn hay không. Nếu chỉ nhìn một loại cảm biến, có nhiều thứ bị bỏ lỡ. Khi xem nhiều cảm biến cùng nhau, phán đoán chính xác hơn.

Hãy xem cảnh robot gấp chén nung làm ví dụ. Nếu camera bị che, robot không thấy được vị trí. Lúc này cảm biến lực phát hiện va chạm. Nếu lực nhỏ đến mức khó cảm nhận, camera nhìn thấy bột bị rơi vãi. Các cảm biến khác nhau bù vào chỗ trống của nhau. Điều này giống việc con người dùng cả tay và mắt. Trong phòng tối thì sờ bằng tay, ở nơi sáng thì nhìn bằng mắt.

Khi robot va vào vật gì đó, giá trị cảm biến lực tăng vọt. Nếu vượt giá trị giới hạn đã định, robot lập tức dừng động tác. Sau khi dừng, nó xem lại tình huống. Nó dùng camera kiểm tra bột có đổ ra không, chén nung có trượt không. Quá trình kiểm tra lại này là khởi đầu của phục hồi lỗi.

Phục hồi lỗi thường diễn ra theo ba bước. Trước hết, hệ thống tìm bất thường để biết điều gì đã sai. Tiếp theo, robot quay về trạng thái an toàn ngay trước khi thất bại. Cuối cùng, nó thay đổi vị trí gấp hoặc lực một chút rồi thử thêm một lần. Qua ba bước này, sai sót nhỏ có thể được vượt qua mà không cần con người.

Phục hồi lỗi cũng quan trọng đối với chất lượng thí nghiệm. Nếu tiếp tục xử lý mẫu đã cân sai, dữ liệu xấu sẽ tích tụ. Dữ liệu xấu dẫn mô hình trí tuệ nhân tạo đi sai hướng. Vì vậy, phòng thí nghiệm tự hành không che giấu thất bại mà ghi lại. Thất bại được ghi lại trở thành dữ liệu cho lần học tiếp theo. Nếu biết thất bại xảy ra ở đâu và vì sao, lần sau hệ thống tránh được sai lầm đó.

Chức năng này cũng có giới hạn. Giá trị cảm biến cũng lẫn nhiễu. Nếu đọc nhầm nhiễu là bất thường, hệ thống sẽ dừng một thí nghiệm đang bình thường. Nếu dùng quá thường xuyên, thí nghiệm chậm lại.

Ngược lại, nếu đặt giá trị giới hạn quá lỏng, hệ thống bỏ lỡ tai nạn thật. Vì vậy, phải chỉnh độ nhạy của thiết bị này cho phù hợp. Việc điều chỉnh này hiện vẫn dựa nhiều vào kinh nghiệm của con người. Phục hồi tự hành không phải là thiết bị để loại con người ra. Nó là thiết bị để giảm những lần lặp nguy hiểm và gọi con người khi cần.

### 3.5 Phòng thí nghiệm tự hành đã vận hành thật sự, Berkeley A-Lab

Mục này xem một trường hợp tiêu biểu về phòng thí nghiệm tự hành đã vận hành thật sự. Đó là A-Lab của Phòng thí nghiệm Quốc gia Lawrence Berkeley ở Hoa Kỳ. Ta sẽ xem trường hợp này đã đạt được gì và vấp phải bức tường nào. Ta cũng sẽ cùng xem cách đọc các con số thành quả.

A-Lab là nền tảng tổng hợp vật liệu vô cơ tự hành do Phòng thí nghiệm Quốc gia Lawrence Berkeley xây dựng. Vật liệu vô cơ là vật chất tạo bởi các nguyên tố như kim loại và oxy. Oxide hoặc phosphate thuộc nhóm này. A-Lab chọn các ứng viên có thể chế tạo từ dữ liệu tính toán. Sau đó nó dùng robot xử lý bột và nhiệt luyện. Nó xác nhận kết quả bằng thiết bị nhiễu xạ tia X. Quá trình này được nối tiếp với rất ít bàn tay con người.

Việc đọc chính xác các con số thành quả rất quan trọng. A-Lab được công bố trong một bài báo Nature năm 2023. Con số thành công ban đầu sau đó được xem xét lại và hạ xuống. Các nhà hóa học vật liệu khác đã phân tích lại vân nhiễu xạ. Năm 2026, Nature đăng bản đính chính. Sau đính chính, con số là 36 trong 57 mục tiêu. Sách này lấy con số đã đính chính làm chuẩn. Không dùng con số cao hơn ban đầu.

Chính quá trình đính chính này là một bài học có giá trị. Trọng tâm của đính chính có hai điểm. Một là phán đoán sai mà phân tích tự động đã gặp. Việc phán định thành công do phân tích nhiễu xạ tia X tự động đảm nhiệm. Phân tích này đã đọc nhầm tín hiệu của nguyên liệu còn sót lại chưa phản ứng thành tín hiệu của vật liệu mới mục tiêu. Nhìn bên ngoài, có vẻ như tinh thể mới đã được tạo ra. Chỉ khi con người xem xét lại kỹ vân nhiễu xạ, phán đoán sai này mới lộ ra. Điểm còn lại là ý nghĩa của cụm từ vật liệu mới. Ban đầu, nó được đọc như thể đã tạo ra vật liệu chưa từng có trên đời. Khi xem lại, nhiều trường hợp chỉ có nghĩa là vật liệu chưa có trong dữ liệu của phòng thí nghiệm này. Điều đó khác với vật liệu lần đầu xuất hiện trong toàn bộ giới khoa học. Bản đính chính cho biết các thành công được báo cáo đã được con người xác nhận lại, và một phần vẫn để ở trạng thái chưa xác định. Vì vậy, phán định của phòng thí nghiệm tự hành không phải câu trả lời cuối cùng. Chỉ sau phân tích chính xác và thử nghiệm tái lập, nó mới trở thành thành công thật. Nếu thí nghiệm tự hành nhanh mà bỏ qua kiểm chứng, sai sót sẽ tích tụ nhiều hơn.

Hãy xem trường hợp này có ý nghĩa gì trong nghiên cứu vật liệu tên lửa. A-Lab cho thấy tổng hợp tự hành không phải là tưởng tượng bên ngoài phòng thí nghiệm. Nhóm oxide và ceramic liên quan đến vật liệu cách nhiệt của tên lửa. Một con đường đã mở ra để tìm kiếm những vật liệu như vậy nhanh hơn con người. Tuy nhiên, sự nghiêm ngặt trong phán định thành công phải đi kèm. Với linh kiện hàng không vũ trụ, một tạp chất nhỏ cũng có thể trở thành vấn đề lớn. Vì vậy, kiểm chứng quan trọng ngang với tốc độ.

A-Lab bộc lộ ba bức tường rõ rệt. Tiểu mục tiếp theo sẽ xem từng bức tường một. Những bức tường này cũng là vấn đề mà các phòng thí nghiệm tự hành khác cùng gặp. Biết chính xác các bức tường là nền tảng để xây hệ thống tiếp theo.

#### 3.5.1 Cấu trúc hệ thống và thành quả vận hành

Tiểu mục này xem A-Lab được tạo thành từ những bộ phận nào. Sau đó xem các bộ phận đó ăn khớp và vận hành ra sao. Ta sẽ hình dung cụ thể đáng về thật của phòng thí nghiệm tự hành.

Cấu trúc của A-Lab có thể chia thành ba phần. Tầng tính toán chọn vật liệu ứng viên và phương pháp tổng hợp. Ở tầng thực thi, robot và thiết bị thật sự tạo mẫu. Ở tầng đo lường, thiết bị nhiễu xạ tia X xác nhận kết quả. Ba tầng này nối từ tính toán sang thực thi, từ thực thi sang đo lường, tạo thành một vòng.

Tầng tính toán chọn ứng viên từ cơ sở dữ liệu lớn. Nó dùng dữ liệu công khai mang tên Materials Project. Nó cũng dùng dữ liệu tính toán của Google. Các dữ liệu này dự đoán những vật liệu có thể ổn định về mặt nhiệt động lực học. Ổn định có nghĩa là vật liệu đó không dễ phân rã và có thể duy trì. A-Lab lấy những vật liệu được dự đoán như vậy làm mục tiêu.

Tầng thực thi có nhiều cánh tay robot. Robot chuyển mẫu giữa nơi cân bột, nơi trộn, nơi ép và nơi nung. Nó cân bột chính xác và trộn bằng máy nghiền bi. Máy nghiền bi là thiết bị nghiền và trộn bột trong một bình có các viên bi nhỏ. Bột đã trộn được ép thành viên nhỏ. Những viên đó được nung trong lò điện ở nhiệt độ cao.

Thiết bị cốt lõi của tầng đo lường là máy nhiễu xạ tia X. Nhiễu xạ tia X (X-ray Diffraction) là phương pháp chiếu tia X để xem cấu trúc tinh thể. Mỗi tinh thể có vân tia X bật ra khác nhau. Vân này giống như dấu vân tay của vật liệu. Trí tuệ nhân tạo đọc vân này để phán đoán vật liệu mục tiêu đã được tạo ra hay chưa. Nếu vân khớp với mục tiêu, nó được tính là thành công.

A-Lab vận hành quá trình này liên tục trong nhiều ngày. Thí nghiệm vẫn tiếp tục khi con người ngủ vào ban đêm. Điều này cho thấy phòng thí nghiệm tự hành đã chuyển từ ý tưởng trong bài báo sang thiết bị thực tế. Nó đã tạo ra nhiều vật liệu mục tiêu mà không cần con người can thiệp. Thành quả này được ghi nhận như một trường hợp mở cửa cho tổng hợp vật liệu tự hành.

Tuy nhiên, đó không phải là vận hành hoàn toàn không người. Nhà nghiên cứu con người chuẩn bị và giám sát hệ thống. Các thí nghiệm thất bại được con người cùng xem xét. Phán định thành công sau đó cũng được con người xác nhận lại. Phòng thí nghiệm tự hành giảm việc của con người, nhưng không xóa con người đi. Con người phụ trách phán đoán cấp cao hơn và kiểm chứng. Máy móc phụ trách những thao tác lặp chi tiết. Sự phân công này là hình ảnh thật của phòng thí nghiệm tự hành hiện nay.

### 3.5.2 Ba bức tường mà A-Lab bộc lộ

Tiểu mục này xem từng bức tường trong ba bức tường mà A-Lab bộc lộ. Những bức tường này là bài toán chung mà phòng thí nghiệm tự hành phải vượt qua. Phải biết chính xác vấn đề thì mới xây tốt hệ thống tiếp theo.

Bức tường thứ nhất là khó khăn trong phán định kết quả tổng hợp. Một số vật liệu được tính là thành công không phải là đơn pha. Đơn pha là trạng thái chỉ có một tinh thể mong muốn một cách tinh khiết. Có mẫu còn sót nguyên liệu chưa phản ứng. Có mẫu chứa cả pha thứ cấp. Pha thứ cấp là tinh thể khác không phải mục tiêu. Dù vân nhiễu xạ giống mục tiêu, tạp chất vẫn có thể bị trộn vào. Vì vậy, phải xác nhận nghiêm ngặt xem có phải đơn pha hay không.

Bức tường thứ hai là xử lý bột. Mỗi loại bột có tính chất khác nhau. Có loại bột mang tính điện. Titan dioxide hoặc bột nano là như vậy. Những loại bột này bám vào bên trong vòi phun. Khi đó vòi định lượng bị tắc. Nguyên liệu dễ hút ẩm cũng khó xử lý. Máy định lượng tự động thường dừng lại trước nhiều loại bột như vậy. Việc con người chỉ cần gõ nhẹ bằng tay là qua, máy móc lại bị kẹt.

Bức tường thứ ba là kiểm soát môi trường. A-Lab vận hành trong môi trường mở ra khí quyển. Trong không khí có oxy và hơi nước. Vật liệu yếu trước oxy và hơi nước không thể được xử lý trong môi trường như vậy. Bởi vì tính chất của chúng thay đổi hoặc chúng bị phân hủy ngay sau khi được tạo ra. Vật liệu sulfide và halide dùng cho pin thể rắn gặp vấn đề này. Những vật liệu này không dùng được nếu tiếp xúc với không khí. Vì vậy A-Lab mở ra khí quyển không thể khám phá nhóm vật liệu này.

Những bức tường này cho thấy rõ vì sao tự hành hóa lại khó. Con người vượt qua các vấn đề này bằng kinh nghiệm của mắt và tay. Khi bột vón cục, họ gõ nhẹ vào bình. Khi mẫu có gì lạ, họ nhận ra qua mùi và màu. Với máy móc, từng kinh nghiệm như vậy phải được dạy riêng. Việc con người làm mà không

cần nghỉ lại là việc khó đối với máy. Sự phát triển của phòng thí nghiệm tự hành là quá trình chuyển những cảm giác nhỏ như vậy sang máy móc.

Ba bức tường này đan xen với nhau. Nếu xử lý bột kém, định lượng sẽ sai. Nếu định lượng sai, việc đánh giá kết quả cũng dao động. Nếu không kiểm soát môi trường, có những vật liệu không thể tạo ra ngay từ đầu. Muốn mở rộng phòng thí nghiệm tự hành, phải vượt qua các bức tường này cùng lúc. Chỉ sửa một vấn đề thì khó tạo ra bước tiến lớn.

Trong nghiên cứu vật liệu tên lửa, những bức tường này hiện ra lớn hơn. Vật liệu pin mới hoặc hợp chất đặc biệt dùng cho tên lửa thường có độ phản ứng mạnh. Cũng có nhiều vật liệu yếu trước không khí. Muốn xử lý các vật liệu này, kiểm soát môi trường phải đi trước. Phòng thí nghiệm mở ra khí quyển khó đựng đến nhóm vật liệu này. Vì vậy bức tường thứ ba, kiểm soát môi trường, trở thành trở ngại đặc biệt lớn trong nghiên cứu vật liệu tên lửa.

Nỗ lực vượt qua bức tường này là chủ đề của mục tiếp theo. Muốn xử lý vật liệu yếu trước không khí, cần một môi trường kín. Môi trường đó là glovebox. Mục tiếp theo xem xét phòng thí nghiệm tự hành vận hành trong môi trường kín. Đó là hệ thống nhắm thẳng vào bức tường thứ ba của A-Lab.

### 3.6 Tìm vật liệu yếu trước không khí trong môi trường kín(A-Lab GPSS)

Mục này bàn về phòng thí nghiệm tự hành vận hành trong môi trường kín. Đây là hệ thống cố gắng vượt qua bức tường thứ ba đã thấy ở mục trước. Ta xem cách xử lý vật liệu yếu trước không khí.

Glovebox là một hộp kín có gắn gang tay. Trên thành hộp có gắn gang tay cao su dày. Con người đưa tay vào gang từ bên ngoài để thao tác bên trong. Bên trong hộp thường được nạp khí trơ như argon. Khí trơ là khí khó phản ứng với chất khác. Vì vậy bên trong hộp hầu như không có oxy và hơi nước.

Tên của hệ thống này là A-Lab GPSS. GPSS có nghĩa là hệ thống tổng hợp được bảo vệ bằng glovebox. Đây là thế hệ tiếp theo được tạo ra để vượt qua giới hạn của A-Lab mở ra khí quyển. Thiết bị được đặt trong một glovebox lớn nạp argon siêu tinh khiết. Hệ thống giữ hơi nước và oxy ở mức thấp hơn một phần triệu. Trong đó, tất cả các bước lưu trữ nguyên liệu, định lượng, tạo hình, niêm kín và xử lý nhiệt đều được thực hiện. Điểm cốt lõi là loại bỏ mọi khoảnh khắc mẫu tiếp xúc với không khí.

Trường hợp này được xác nhận qua một bản thảo trước xuất bản năm 2026. Nhóm nghiên cứu gồm Feiyu và cộng sự. Đây là nghiên cứu ban đầu xử lý chất dẫn spinel lithium halide yếu trước không khí. Bài báo cho biết họ đã tạo ra ba trăm năm mươi hai mẫu. Số tổ hợp cặp kim loại khả dĩ là một trăm bảy mươi một. Trong số đó, bảy mươi hai phần trăm đã được xử lý bằng thí nghiệm. Ở đây, bảy mươi hai phần trăm không phải là tỷ lệ thành công. Đó là tỷ lệ các tổ hợp khả dĩ đã được bao phủ bằng thí nghiệm. Không được đọc nhầm con số này là tỷ lệ thành công.

Bài báo cho biết kết quả tốt lên khi thí nghiệm tiếp tục. Trong bảy mươi lăm mẫu đầu tiên, tỷ lệ kết quả tốt còn thấp. Trong bảy mươi lăm mẫu cuối cùng, tỷ lệ đó tăng lên. Trí tuệ nhân tạo đã học từ các thí nghiệm trước để cải thiện các thí nghiệm sau. Điều này cho thấy học vòng kín thật sự hoạt động. Tuy nhiên nghiên cứu này vẫn ở giai đoạn ban đầu, nên không nên mở rộng thành kết luận quá rộng.

Ta xem vì sao phòng thí nghiệm tự hành kín này có giá trị trong ngành tên lửa. Tàu vũ trụ cần pin có hiệu năng tốt. Pin thể rắn là ứng viên an toàn và có năng lượng cao. Cốt lõi của nó là chất điện phân rắn. Nhưng chất điện phân rắn tốt thường yếu trước không khí. Phòng thí nghiệm tự hành kín có thể khám phá an toàn những vật liệu như vậy. Môi trường tương tự cũng được dùng trong nghiên cứu vật liệu tên lửa có độ phản ứng mạnh.

### 3.6.1 Chất điện phân rắn lithium halide yếu trước không khí

Tiểu mục này xem kỹ vật liệu được bàn ở đây. Nó giải thích bằng lời dễ hiểu chất điện phân rắn lithium halide là gì. Nó xem vì sao vật liệu này yếu trước không khí và vì sao nó có giá trị.

Trước hết, hãy giải nghĩa các thuật ngữ trong pin. Ion lithium là hạt nhỏ mang điện tích bên trong pin. Khi pin được sạc và xả, các ion này di chuyển qua lại. Chất điện phân là con đường để các ion này đi qua. Pin thông thường dùng chất điện phân lỏng. Chất lỏng có thể rò rỉ hoặc bắt lửa.

Chất điện phân rắn là việc làm con đường này bằng chất rắn. Chất điện phân rắn phải cho ion lithium đi qua tốt. Nhưng nó phải chặn electron tốt. Chất rắn cho ion đi qua và chặn electron là chất điện phân tốt. Chất rắn không rò rỉ và chịu lửa tốt. Vì vậy pin thể rắn được xem là loại pin an toàn hơn.

Chất điện phân lithium halide là chất điện phân rắn chứa nguyên tố halogen. Halogen là các nguyên tố như chlorine, bromine và iodine. Về công thức hóa học, lithium, kim loại và halogen liên kết theo một tỷ lệ xác định. Vật liệu này có khả năng phù hợp với điện cực điện áp cao. Khi điện áp cao, cùng một kích thước có thể chứa nhiều năng lượng hơn. Vì vậy nó thu hút chú ý như một vật liệu pin thế hệ tiếp theo.

Vấn đề là vật liệu này rất yếu trước nước. Tính chất này được gọi là tính chảy rữa. Tính chảy rữa là tính chất hút hơi nước trong không khí rồi tự hòa tan. Chỉ cần tiếp xúc với một lượng nước rất nhỏ, pha của vật liệu cũng phân hủy. Khi phân hủy, nó có thể sinh ra khí có hại như hydrogen chloride. Vì vậy lưu trữ nguyên liệu, tổng hợp và đo lường đều phải được thực hiện trong môi trường kín.

Điều kiện này làm cho phòng thí nghiệm tự hành trở nên khó hơn. Bên trong hộp kín hẹp và khó thao tác. Robot phải di chuyển chính xác trong không gian hẹp. Không được cho phép phơi nhiễm không khí dù chỉ trong khoảnh khắc. Trí tuệ nhân tạo đề xuất thành phần ứng viên và điều kiện tổng hợp. Robot lặp lại cùng quy trình mà không để mẫu tiếp xúc với không khí. Thiết bị phân tích nhanh chóng kiểm tra độ dẫn ion và pha tinh thể. Tất cả những việc này ăn khớp với nhau bên trong môi trường kín.

Việc đo độ dẫn ion cũng quan trọng. Độ dẫn ion là giá trị cho biết ion lithium đi qua tốt đến mức nào. Giá trị này phải cao thì pin mới phát công suất tốt. Phòng thí nghiệm tự hành đo ngay độ dẫn của mẫu vừa tạo ra. Rồi dùng giá trị đó làm mục tiêu cho trí tuệ nhân tạo. Nó chọn thí nghiệm tiếp theo theo hướng các thành phần có độ dẫn cao. Nó vừa xem pha tinh thể vừa xem độ dẫn để thu hẹp dần về chất điện phân tốt.

Việc khám phá vật liệu này vẫn còn giới hạn. Môi trường kín khiến việc đưa thiết bị vào và lấy thiết bị ra trở nên khó. Không gian hẹp làm hạn chế thiết bị có thể xử lý. Vật liệu có tính chảy rữa chỉ cần sai sót nhỏ cũng không dùng được. Vì vậy từng thí nghiệm đều phải thận trọng. Dù tự hành hóa làm tăng tốc độ, sự thận trọng này không giảm đi. Phòng thí nghiệm tự hành kín hiện vẫn xử lý phạm vi vật liệu còn hẹp.

### 3.6.2 Vòng lặp qua lại giữa quy nạp và suy đoán nguyên nhân

Tiểu mục này bàn về hai cách suy nghĩ mà phòng thí nghiệm tự hành sử dụng. Đó là quy nạp và suy đoán nguyên nhân. Ta xem vì sao việc dùng luân phiên hai cách này làm tăng tốc khám phá. Đây là cấu trúc suy luận mà nghiên cứu A-Lab GPSS trong mục này đã nêu ra.

Quy nạp là cách suy nghĩ tìm quy luật chung từ nhiều kết quả thí nghiệm. Nó gom các vân nhiễu xạ và kết quả đo độ dẫn. Nó xem thành phần nào đã được tạo ra tốt. Nó đặt nhiều mẫu cạnh nhau và đọc xu hướng. Nó xem độ dẫn thay đổi thế nào khi mạng tinh thể mở rộng. Những quy luật tích lũy như vậy trở thành nền tảng cho thí nghiệm tiếp theo.

Suy đoán nguyên nhân là cách suy nghĩ tìm nguyên nhân khi gặp kết quả bất ngờ. Suy đoán nguyên nhân là suy luận chọn nguyên nhân giải thích tốt một kết quả lạ. Giả sử một mẫu xuất hiện pha tạp chất không dự đoán trước. Ta xem nguyên liệu phản ứng chưa đủ hay nhiệt độ quá thấp. Trong số đó, ta chọn

nguyên nhân hợp lý nhất rồi thay đổi thí nghiệm tiếp theo. Suy đoán nguyên nhân tìm ngay nước đi kế tiếp từ một kết quả gây bất ngờ. Nó có hướng khác với quy nạp, vốn rút ra quy luật từ nhiều kết quả.

Phòng thí nghiệm tự hành dùng luân phiên hai cách suy nghĩ này. Tính toán lý thuyết phiếm hàm mật độ (Density Functional Theory) thu hẹp trước các ứng viên có khả năng. Tính toán này tìm năng lượng của vật chất ở cấp độ nguyên tử. Thí nghiệm kiểm tra các ứng viên được chọn ở đây. Khi kết quả thí nghiệm tích lũy, quy nạp tìm quy luật. Khi có kết quả bất ngờ, suy đoán nguyên nhân chỉ ra nguyên nhân và đối hướng. Như vậy tính toán và thí nghiệm chỉnh sửa lẫn nhau. Suy nghĩ không chảy theo một hướng duy nhất, mà quay vòng như một vòng lặp.

Ta xem cách này bằng một ví dụ dễ hiểu. Hãy hình dung cảnh tìm đường trong một thành phố lạ. Việc đi nhiều lần rồi tìm ra quy luật của đường tắt là quy nạp. Khi đường đột nhiên bị chặn, việc đoán nguyên nhân rồi đi vòng là suy đoán nguyên nhân. Nếu chỉ tin vào quy luật, ta sẽ dừng lại trước tình huống bất ngờ. Nếu chỉ chạy theo phỏng đoán nguyên nhân, ta sẽ mất bức tranh lớn. Dùng cả hai cùng nhau thì sẽ đến đích nhanh hơn. Phòng thí nghiệm tự hành cũng dùng cả hai cách suy nghĩ để đến nhanh hơn với vật liệu tốt.

Ta xem vì sao vòng kép này có giá trị trong nghiên cứu vật liệu tên lửa. Vật liệu mới không thể được hiểu hết chỉ bằng tính toán. Tính toán giả định các điều kiện lý tưởng. Tổng hợp thật có tạp chất và nhiễu. Chỉ thí nghiệm cũng không thể biết hết. Thí nghiệm tốn nhiều thời gian và chi phí. Khi nối hai phương pháp với nhau, ta thu hẹp phạm vi bằng tính toán và xác nhận bằng thí nghiệm. Trong việc khám phá vật liệu tên lửa đắt đỏ, tổ hợp này tạo ra sức mạnh lớn.

Cách này cũng có giới hạn. Ngôn ngữ của tính toán và thí nghiệm khác nhau. Việc khớp kết quả tính toán với kết quả thí nghiệm khá khó. Nếu tính toán sai, thí nghiệm bị lãng phí cho ứng viên sai. Nếu nhiễu thí nghiệm lớn, việc sửa tính toán sẽ sai. Vì vậy phải đặt cây cầu thật tốt giữa hai vòng lặp. Việc đặt cây cầu này hiện vẫn phụ thuộc nhiều vào phán đoán của con người.

### 3.6.3 Hợp tác giữa nhiều trí tuệ nhân tạo chia vai trò

Tiểu mục này bàn về cách nhiều trí tuệ nhân tạo chia vai trò và hợp tác. Cách này được gọi là đa tác tử. Ta xem vì sao nhiều trí tuệ nhân tạo tốt hơn một trí tuệ nhân tạo và có những rủi ro nào.

Đa tác tử là cấu trúc trong đó nhiều trí tuệ nhân tạo chia vai trò. Tác tử là đơn vị trí tuệ nhân tạo tự phán đoán và hành động. Nó giống như một nhóm chia việc thay vì một người làm tất cả. Trong phòng thí nghiệm tự hành, sự phân vai như vậy là tự nhiên. Bởi vì thí nghiệm cần cả kiến thức hóa học, lập kế hoạch, an toàn và phân tích.

Khi chia vai trò, mỗi tác tử tập trung vào một việc. Tác tử hóa học xem phương trình phản ứng có đúng không. Nó kiểm tra khối lượng có được bảo toàn không, số oxy hóa có khớp không. Nó cũng xem nguyên liệu có phản ứng xấu với chén nung không. Tác tử kế hoạch sắp xếp thứ tự thiết bị. Nó chia trước các công đoạn gây nghẽn như xử lý nhiệt dài. Nó cũng quyết định những việc nào có thể làm song song.

Tác tử an toàn là con mắt bảo vệ phòng thí nghiệm. Nó theo dõi nồng độ oxy và hơi nước trong glovebox. Nó cũng xem biến đổi áp suất và dòng nước làm mát của lò điện. Nó lặp lại việc kiểm tra các giá trị này theo chu kỳ rất ngắn. Nếu thấy giá trị lạ, nó đưa ra cảnh báo và lưu ghi chép. Tuy nhiên tác tử này không trực tiếp kích hoạt dừng khẩn cấp thật sự. Bởi vì tác tử tạo sinh có thể phản hồi chậm hoặc phán đoán sai. Khi oxy hoặc áp suất đạt đến giá trị nguy hiểm, cảm biến độc lập và bộ điều khiển an toàn đã được kiểm chứng sẽ dừng thiết bị ngay lập tức. Khóa liên động phần cứng này không chờ phán đoán của mô hình ngôn ngữ. Tác tử đảm nhiệm giám sát và ghi chép, còn việc dừng lại do thiết bị an toàn đã được kiểm chứng đảm nhiệm. Tác tử phân tích diễn giải kết quả đo. Nó khớp các vân nhiễu xạ với cơ sở dữ liệu. Nó lượng hóa lượng tạp chất bằng con số. Rồi đưa con số đó vào giá trị mục tiêu của bộ học.

Năm 2026, nghiên cứu về phòng thí nghiệm đa tác tử như vậy tăng mạnh. Một nghiên cứu đã lắp ghép nhiều tác tử trí tuệ nhân tạo dưới một bộ điều phối. Mỗi tác tử chia nhau dùng các công cụ khác nhau. Chúng được chia thành nhóm tìm tri thức, nhóm chuyển thiết kế thành quy trình thực hiện, nhóm xử lý robot và nhóm diễn giải dữ liệu. Bằng cách chia nhóm như vậy, chúng tạo nên khám phá vật liệu vòng kín. Một nghiên cứu khác, ChatMat, đã tự động nối luồng dự đoán tính chất vật liệu bằng nhiều tác tử. Cũng có bài báo bàn về ý nghĩa của đa tác tử trong quản lý phòng thí nghiệm tự hành.

Khi chia vai trò, lợi ích hiện ra rõ. Có thể thấy rõ việc mỗi tác tử phụ trách. Khi phát sinh vấn đề, dễ tìm nó xảy ra ở bước nào. Dù sửa một tác tử, không cần thay đổi toàn bộ. Việc gắn thêm công cụ mới cũng dễ. Điều này giống với ưu điểm của làm việc nhóm giữa con người.

Hệ đa tác tử cũng làm thay đổi chính cách thức vận hành của khoa học. Trước đây, nhà nghiên cứu là con người ghi lại từng thí nghiệm một. Giờ đây, tác tử để lại quá trình phán đoán dưới dạng dữ liệu. Vì sao chọn điều kiện này, vì sao xem kết quả này là thành công, tất cả đều được lưu lại. Những ghi chép này về sau được dùng để lần lại nghiên cứu. Tuy nhiên, phán đoán của trí tuệ nhân tạo không phải lúc nào cũng minh bạch. Con người khó biết hết vì sao nó đưa ra quyết định như vậy. Sự thiếu minh bạch này vẫn là bài toán cần giải trong thời gian tới.

Nhưng rủi ro cũng lớn lên cùng lúc. Mô hình ngôn ngữ lớn có thể bịa ra những sự thật không tồn tại. Sai sót của một tác tử có thể lan sang tác tử khác. Nếu nhiều tác tử tin nhau rồi chuyển tiếp công việc, chúng có thể không phát hiện lỗi. Vì vậy, quyền quyết định cuối cùng phải thuộc về quy tắc an toàn và người giám sát là con người. Ngay cả khi trí tuệ nhân tạo nói một cách tự tin, thiết bị vẫn có thể nguy hiểm. Càng có nhiều tác tử, việc kiểm tra chéo của con người và thiết bị an toàn phần cứng càng quan trọng.

## Tổng kết

Chương này đã xem xét phòng thí nghiệm tự hành vận hành như thế nào. Cốt lõi là cấu trúc vòng khép kín. Trí tuệ nhân tạo chọn thí nghiệm tiếp theo. Robot chế tạo. Thiết bị đo. Kết quả quay trở lại với trí tuệ nhân tạo. Vòng này chạy cả ngày lẫn đêm, nhanh chóng thu hẹp không gian vật liệu rộng lớn.

Phương pháp để trí tuệ nhân tạo chọn thí nghiệm tiếp theo là học chủ động Bayesian. Phương pháp này xem xét đồng thời hiệu năng dự đoán và độ bất định. Nó chọn cân bằng giữa nơi có triển vọng và nơi chưa biết rõ. Vì vậy, nó có thể chạm tới vật liệu tốt chỉ với ít thí nghiệm. Điểm này trở thành sức mạnh lớn trong việc tìm kiếm vật liệu tên lửa đất đỏ.

Phương pháp tổng hợp trong các bài báo được mô hình ngôn ngữ lớn rút ra. Những nghiên cứu như MSP-LLM làm việc này. Phương pháp được rút ra sau đó được chuyển đổi an toàn sang ngôn ngữ lập kế hoạch như PDDL. Ngôn ngữ lập kế hoạch trở thành hàng rào ngăn trước tai nạn của robot. Trí tuệ nhân tạo và thiết bị được nối với nhau bằng các tiêu chuẩn như giao thức ngữ cảnh mô hình. Bộ điều khiển NIMO là một ví dụ. Máy trạng thái là cổng chặn cuối cùng đối với các lệnh nguy hiểm.

Robot phải chịu được môi trường thí nghiệm lộn xộn. Hợp nhất cảm biến, tức việc đọc nhiều cảm biến cùng lúc, giúp làm điều đó. Khả năng tự khôi phục lỗi cũng hoạt động cùng với nó. Các hệ thống như PUDA để lại lệnh và kết quả dưới dạng hồ sơ. Hồ sơ này có giá trị lớn trong việc truy vết chất lượng linh kiện hàng không vũ trụ.

Ta đã xem A-Lab và A-Lab GPSS như các trường hợp thực tế. A-Lab đã tạo ra 36 trong số 57 mục tiêu trong 17 ngày. Con số này là giá trị đã qua đỉnh chính năm 2026. Trường hợp này cho thấy phán đoán của phòng thí nghiệm tự hành cũng cần con người xem xét lại. A-Lab GPSS xử lý vật liệu nhạy với không khí trong môi trường kín. Với 352 mẫu, nó bao phủ 72 phần trăm các tổ hợp cặp kim loại có thể có. Con số này không phải tỷ lệ thành công, mà là độ bao phủ tổ hợp.

Bước sang năm 2026, lĩnh vực này đã mở rộng nhanh chóng. Phòng thí nghiệm đa tác tử, nơi nhiều trí tuệ nhân tạo chia nhau vai trò, đã tăng lên. Các nỗ lực kết nối thiết bị bằng truyền thông tiêu chuẩn cũng tiếp diễn. Nghiên cứu điều chỉnh tối ưu hóa Bayesian cho phù hợp với luồng thí nghiệm cũng diễn ra sôi động. Các bài tổng quan và hướng dẫn khảo sát rộng về phòng thí nghiệm tự hành cũng đã xuất hiện. Xu hướng lớn rất rõ. Đó là giảm bớt bàn tay con người khỏi những thao tác lặp lại tỉ mỉ. Tuy nhiên, vị trí của phán đoán và kiểm chứng vẫn thuộc về con người.

Các hệ thống trong chương này phần lớn vẫn là nghiên cứu giai đoạn đầu. Có nhiều bản thảo tiền công bố, còn tiêu chuẩn hiện trường vẫn ít. Vì vậy, thay vì đặt các con số hiệu năng lên trước, chương này đã xem cả nguyên lý lẫn giới hạn. Phòng thí nghiệm tự hành không phải là công cụ xóa bỏ con người. Nó giao những lặp lại nguy hiểm cho máy móc, còn con người đảm nhận phán đoán và kiểm chứng. Ý nghĩa thật sự của phòng thí nghiệm tự hành không phải là thí nghiệm nhanh, mà là thí nghiệm đáng tin. Chỉ trên nguyên tắc này, tự động hóa trong nghiên cứu vật liệu tên lửa mới có thể phát triển an toàn.

### Tài liệu căn cứ liên quan

- 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>

# Chương 4. Dự đoán nhất quán vật lý và tối ưu hóa thành phần của hợp kim entropy cao (HEA) và siêu hợp kim

## Mở đầu

Khi dùng một chiếc nồi trong bếp lâu ngày, đáy nồi sẽ cháy đen. Phần thường xuyên tiếp xúc với lửa sẽ hồng trước. Kim loại càng nóng thì càng mềm và yếu. Tính chất yếu đi vì nhiệt này có ở mọi kim loại như nhau.

Động cơ tên lửa gặp vấn đề này nghiêm trọng hơn bất kỳ cỗ máy nào khác. Trong buồng đốt, nhiên liệu cháy và tạo ra khí ở nhiệt độ hàng nghìn độ. Khí đó phun nhanh ra ngoài qua một đường hẹp gọi là vòi phun. Thành vòi phun phải chịu luồng khí nóng này ngay sát bên cạnh. Đó là vị trí khắc nghiệt khó có thể so sánh với đáy nồi.

Các phương tiện bay quay trở lại khí quyển cũng mang cùng một nỗi lo. Khi lao vào không khí với tốc độ cao, không khí bị nén và nóng lên. Mặt trước của phương tiện bay đỏ rực. Nếu không chịu được nhiệt này, kết cấu sẽ nóng chảy hoặc gãy vỡ.

Ta có thể cảm nhận sự nung nóng không khí này bằng lòng bàn tay. Khi đi xe đạp nhanh, gió đập mạnh vào mặt. Khi đưa tay ra ngoài cửa sổ, gió đẩy vào tay. Phương tiện bay hứng luồng gió đó nhanh hơn hàng nghìn lần. Không khí bị đẩy nén lại, nóng lên, rồi truyền nhiệt vào mặt trước. Nhiệt độ điểm dừng trên bề mặt tái nhập vượt quá 1.600 độ C.

Phương tiện bay siêu vượt âm cũng gặp cùng loại nhiệt này. Siêu vượt âm là tốc độ nhanh hơn âm thanh trên năm lần. Khi bay ở tốc độ này, mép trước của cánh nóng hơn các phần khác. Lý do là nhiệt tập trung vào những phần hẹp và nhọn. Vì vậy, ở phần này cần dùng vật liệu có khả năng chịu nhiệt độ cao rất tốt.

Vì vậy, con người tìm kiếm kim loại không mất độ bền ngay cả ở nơi rất nóng. Kim loại đã giữ vị trí đó trong thời gian dài là siêu hợp kim nền niken (Nickel-based superalloy). Siêu hợp kim là hợp kim được thiết kế để chịu tốt ngay cả ở nhiệt độ cao. Nhưng kim loại này cũng có nhiệt độ giới hạn.

Chương này bàn về một loại kim loại mới nhằm vượt qua giới hạn đó. Nhân vật chính là hợp kim entropy cao (High-Entropy Alloy, HEA). Đây là hợp kim trộn nhiều kim loại với lượng gần tương đương nhau. Trong số đó, nếu chỉ chọn và trộn những kim loại có điểm nóng chảy cao, ta có một ứng viên vật liệu mới cho tên lửa. Loại hợp kim này gọi là hợp kim entropy cao chịu lửa (Refractory High-Entropy Alloy, RHEA).

Chương này giải thích một phương pháp mới để tìm ra những kim loại như vậy. Thay vì con người thí nghiệm từng mẫu một, trí tuệ nhân tạo (AI) nhanh chóng chọn lọc các ứng viên. Ta sẽ lần lượt xem cách sắp xếp nguyên tử, các định luật nhiệt động lực học và những phép tính nguyên tử quy mô lớn gặp gỡ trí tuệ nhân tạo như thế nào. Đây là câu chuyện về kim loại, nhưng tôi sẽ giải thích bằng những ví dụ gần gũi như mùa trong năm và trường học.

Trước hết, hãy hình dung vì sao cần trí tuệ nhân tạo. Để tạo ra một hợp kim tốt, ta phải quyết định loại nguyên tố và tỷ lệ của chúng. Chỉ cần có năm nguyên tố, số tổ hợp tỷ lệ đã khó đếm hết. Nếu chế tạo và thử từng mẫu một, cả đời cũng không thể xem hết. Trí tuệ nhân tạo giúp thu hẹp trước không gian ứng viên rộng lớn này.

Vai trò của trí tuệ nhân tạo gần với một người dẫn đường. Trong núi, người dẫn đường biết lối sẽ chỉ cho ta đường tắt. Nhưng người leo lên đỉnh cuối cùng vẫn là con người. Khi trí tuệ nhân tạo thu hẹp ứng

viên, con người xác nhận bằng thí nghiệm. Chương này sẽ cho thấy sự hợp tác ấy diễn ra như thế nào qua các ví dụ.

## 4.1 Yêu cầu của môi trường hàng không vũ trụ cực hạn và giới hạn kết cấu của hợp kim nhiệt độ cao

### Nội dung của mục này

Mục này trước hết phác họa môi trường nhiệt mà tên lửa và phương tiện tái nhập phải đối mặt khắc nghiệt đến mức nào. Sau đó, mục này xem xét siêu hợp kim nền niken, loại vật liệu đã được dùng lâu năm, gặp giới hạn ở đâu. Cuối cùng, mục này giải thích vì sao hợp kim entropy cao chịu lửa nổi lên như một lựa chọn mới.

Nói dễ hiểu, mục này là nơi xác định vấn đề. Phải hiểu đúng vấn đề thì mới thấy được công dụng của các công cụ trí tuệ nhân tạo ở phần sau. Vì vậy, trọng tâm đặt vào việc hiểu tình huống trước khi đi vào con số.

Buồng đốt tên lửa là một bình chịu áp suất rất cao. Nhiên liệu và chất oxy hóa cháy cùng nhau trong đó, tạo ra áp suất lớn. Áp suất này lên tới hàng chục megapascal. Áp suất lớp ô tô khoảng 0,2 megapascal. Buồng đốt tên lửa chịu áp suất lớn gấp hàng trăm lần con số đó.

Khí tạo ra trong buồng đốt thoát ra qua vòi phun. Nhiệt độ của khí này, tùy vật liệu, có thể tăng đến hàng nghìn độ. Tuy nhiên, nhiệt độ của kim loại thành không tăng cao bằng nhiệt độ khí. Làm mát và lớp biên chặn nhiệt, giữ nhiệt độ kim loại thấp hơn. Dù vậy, lượng nhiệt mà thành nhận vẫn rất lớn. Nếu không chịu được nhiệt này, thành sẽ nóng chảy, xuất hiện lỗ thủng, và động cơ không thể sử dụng được.

Nhiệt không chỉ đến một lần. Động cơ lặp lại quá trình bật và tắt. Khi bật, nó nóng lên rất nhanh; khi tắt, nó nguội đi rất nhanh. Hiện tượng nhiệt độ tăng giảm như vậy gọi là sốc nhiệt (Thermal shock). Nguyên lý giống như khi đổ nước nóng đột ngột vào cốc thủy tinh thì cốc bị nứt.

Lực cũng cộng thêm vào đó. Khi kim loại chịu lực trong thời gian dài ở trạng thái nóng, nó từ từ giãn ra. Hiện tượng này gọi là rão (Creep). Nó giống như thanh kẹo kéo khi cầm lâu sẽ tự chảy xệ xuống. Ở nhiệt độ cao, rão và tải trọng lặp lại cùng tác động, bào mòn vật liệu.

Siêu hợp kim nền niken là kim loại đã chịu đựng lâu trong môi trường như vậy. Sức mạnh của kim loại này đến từ tổ chức vi mô. Trong pha nền gamma ( $\gamma$ ), những hạt nhỏ gọi là gamma prime ( $\gamma'$ ) cắm dày đặc. Những hạt này đóng vai trò vật cản, ngăn kim loại biến dạng. Nó giống như những hạt quả nhỏ trong khối bột làm cho bột khó bị kéo giãn.

Mỗi siêu hợp kim có loại vật cản hơi khác nhau. Inconel 718 chứa nhiều niobi (Nb) là một ví dụ. Hợp kim này tạo độ bền chủ yếu nhờ một pha gọi là gamma double prime ( $\gamma''$ ,  $\text{Ni}_3\text{Nb}$ ), hơn là gamma prime. Tên của vật cản có thể khác, nhưng nguyên lý thì giống nhau. Những hạt nhỏ ngăn biến dạng và làm kim loại cứng chắc.

Tên lửa dùng kim loại khác nhau tùy từng vị trí. Trong buồng đốt nơi thành được làm mát bằng nhiên liệu, hợp kim đồng thường được dùng. Lý do là chúng có khả năng đưa nhiệt ra ngoài rất nhanh. Siêu hợp kim nền niken được dùng cho các bộ phận chịu nhiệt độ cao trong thời gian dài, như turbine. Ở cổ vòi phun nóng hơn hoặc bề mặt tái nhập, kim loại chịu lửa hoặc gốm trở thành ứng viên. Chương này bàn về kim loại mới nhắm tới những vị trí có nhiệt độ cao nhất trong số đó.

Vấn đề xuất hiện khi nhiệt độ tăng cao hơn. Khi vượt khoảng 1.150 độ C, những hạt nhỏ này lớn lên hoặc bị phá vỡ hình dạng. Hiện tượng này gọi là thô hóa. Khi vật cản biến mất, kim loại dễ giãn ra và độ bền giảm nhanh. Điểm nóng chảy của siêu hợp kim nền niken khoảng 1.350 độ C, nên giới hạn vận hành hữu hiệu thấp hơn nhiều so với điểm nóng chảy.

Thêm vào đó, các pha có hại cũng có thể hình thành. Khi đặt lâu ở nhiệt độ cao, các nguyên tử tự lại theo một quy luật nhất định và tạo ra những khối cứng nhưng dễ vỡ. Những pha như vậy gọi là pha có hại. Khi pha có hại xuất hiện, kim loại dễ gãy. Siêu hợp kim càng tốt thì càng được thiết kế để kìm hãm những pha như vậy.

Vì những vấn đề này, siêu hợp kim nền niken cần được làm mát. Trong động cơ thực tế, người ta tạo các kênh nhỏ bên trong bộ phận. Nhiên liệu lạnh chảy qua các kênh đó để làm mát thành. Nhờ làm mát này, bộ phận được dùng ở nhiệt độ thấp hơn điểm nóng chảy. Nhưng khi tạo kênh làm mát, kết cấu bộ phận phức tạp hơn và khối lượng cũng tăng.

Mục tiêu của hợp kim entropy cao chịu lửa là giảm việc làm mát này. Nếu bản thân vật liệu chịu được nơi nóng hơn, nhu cầu làm mát sẽ giảm. Khi làm mát giảm, kết cấu nhẹ hơn và gọn hơn. Các nhà nghiên cứu đặt mục tiêu phát triển hợp kim mới là vật liệu chịu được nơi trên 1.400 độ C mà không cần làm mát.

Vì vậy, con người đã tìm một con đường khác hẳn. Ý tưởng là dựng bộ khung bằng những kim loại có điểm nóng chảy vốn đã cao. Các kim loại như niobi (Nb), molybden (Mo), tantali (Ta), vonfram (W) thuộc nhóm này. Điểm nóng chảy của chúng vượt xa 2.000 độ C. Những kim loại như vậy gọi là kim loại chịu lửa (Refractory metal).

Hợp kim entropy cao chịu lửa là hợp kim trộn nhiều kim loại chịu lửa này với lượng gần tương đương nhau. Các nghiên cứu gần đây báo cáo rằng hợp kim này cho độ bền cao hơn siêu hợp kim nền niken ngay cả trong thử nghiệm nén ở 1.600 độ C. Vì vậy, nó được xem là ứng viên mới cho hệ đẩy tên lửa, mép trước cánh của phương tiện bay siêu vượt âm và tấm chắn nhiệt tái nhập.

Nhưng không phải chỉ có ưu điểm. Hợp kim này có xu hướng dễ gãy ở nhiệt độ phòng. Tính chất này gọi là độ dẻo thấp. Và khi gặp oxy ở nhiệt độ cao, bề mặt xảy ra oxy hóa khiến nó vỡ vụn. Hai vấn đề này phải được giải quyết trước khi dùng làm bộ phận thực tế. Trí tuệ nhân tạo được dùng chính để giải nhanh bài toán đó.

Hãy xem kỹ hơn vấn đề oxy hóa. Hợp kim có nhiều vonfram hoặc molybden phản ứng mạnh với oxy ở nhiệt độ cao. Màng oxit bề mặt không bám chắc mà rơi ra như bột. Kiểu oxy hóa làm bề mặt vụn ra như vậy gọi là oxy hóa pest (Pest oxidation). Tên này xuất phát từ ý nghĩa bề mặt bị gặm mòn như cây lương thực bị sâu hại gặm ăn.

Cách ngăn vấn đề này là tạo màng oxit bảo vệ. Nếu thêm nhôm hoặc crom với lượng phù hợp, một màng oxit dày đặc sẽ hình thành trên bề mặt. Màng này ngăn oxy xâm nhập vào bên trong. Nó giống như lớp cháy dính ở đáy nồi đôi khi lại bảo vệ phần kim loại bên trong. Cho bao nhiêu và cho như thế nào để tạo màng tốt là trọng tâm của thiết kế.

Độ tương thích với lớp phủ cũng quan trọng. Trên hợp kim, người ta thường phủ thêm lớp phủ chống oxy hóa. Vấn đề nằm ở chỗ kim loại và lớp phủ giãn nở khác nhau theo nhiệt độ. Mức độ giãn nở này gọi là hệ số giãn nở nhiệt. Nếu hệ số giãn nở nhiệt của hai vật liệu chênh lệch lớn, ranh giới giữa chúng sẽ nứt khi nhiệt độ tăng. Vì vậy, hợp kim không chỉ cần độ bền mà còn phải khớp với lớp phủ.

Như vậy, hợp kim lót vòi phun phải thỏa mãn nhiều điều kiện cùng lúc. Điểm nóng chảy cao, mật độ thấp, độ bền nhiệt độ cao, khả năng chống oxy hóa và độ tương thích với lớp phủ đều cần thiết. Các điều kiện này cũng thường va chạm với nhau. Con người khó có thể giải bài toán chằng chịt như vậy chỉ bằng thí nghiệm. Trí tuệ nhân tạo bước vào như một công cụ sắp xếp nhanh bài toán chằng chịt này.

## 4.2 Mô hình hóa vật lý của bốn hiệu ứng lõi và nhiệt động lực học cân bằng pha

### Nội dung của mục này

Mục này bàn về khung giải thích ban đầu cho việc vì sao hợp kim entropy cao lại đặc biệt. Khi lĩnh vực này mới mở ra, các nhà nghiên cứu nêu ra bốn hiệu ứng. Bốn hiệu ứng này được gọi là bốn hiệu ứng lõi. Đó là entropy cấu hình, biến dạng mạng tinh thể, khuếch tán chậm và hiệu ứng cocktail.

Bốn hiệu ứng này không phải là các định luật vật lý đã được xác lập, mà gần với các giả thuyết ban đầu. Như sẽ thấy ở phần sau, khuếch tán chậm và sự ổn định hóa pha đơn do entropy cao chưa được chứng minh là quy luật phổ quát. Dù vậy, phải hiểu khung này thì mới hiểu trí tuệ nhân tạo đang cố tính toán điều gì. Ta sẽ chuyển từng hiệu ứng thành ví dụ trong đời sống và giải thích lần lượt.

Trước hết, cần chỉ rõ hợp kim entropy cao là gì. Hợp kim thông thường thêm một lượng nhỏ kim loại khác vào một kim loại chính. Thép có nhân vật chính là sắt và có thêm một ít carbon. Hợp kim entropy cao khác với cách này. Nó trộn từ năm kim loại trở lên với lượng gần tương đương nhau. Có thể xem là có nhiều nhân vật chính.

Khái niệm này được hai nhóm nghiên cứu công bố riêng rẽ vào năm 2004. Đó là nhóm Yeh ở Đài Loan và nhóm Cantor ở Anh. Họ cho thấy có thể tạo ra kim loại ổn định mà không cần dựa vào một nguyên tố chính duy nhất. Phát hiện này đã mở cánh cửa cho thiết kế hợp kim mới.

Nguyên lý thứ nhất là entropy cấu hình (Configurational entropy). Entropy là giá trị liên quan đến số cách trộn. Càng có nhiều cách đặt đồ vào tủ, mức độ hỗn độn càng lớn. Khi trộn nhiều kim loại với lượng gần tương đương nhau, số cách sắp xếp nguyên tử trở nên rất lớn. Mức hỗn độn lớn này giúp kim loại duy trì trạng thái trộn thành một khối.

Trạng thái trộn này gọi là dung dịch rắn (Solid solution). Dung dịch rắn là trạng thái nhiều nguyên tố hòa đều vào trong một mạng tinh thể. Nó giống như đường tan trong nước và trở thành một chất lỏng. Khi entropy lớn, nó ngăn các nguyên tố tụ riêng lại và trở thành hợp chất cứng. Vì vậy, hợp kim entropy cao dễ tạo thành dung dịch rắn thay vì hợp chất.

Nguyên lý thứ hai là biến dạng mạng tinh thể (Lattice distortion). Nguyên tử kim loại có kích thước khác nhau. Khi nguyên tử lớn và nguyên tử nhỏ trộn trong cùng một mạng, mạng tinh thể bị méo mó lồi lõm. Nó giống như một bức tường xếp bằng đá có kích thước khác nhau thay vì gạch phẳng đều. Mạng tinh thể méo mó này trở thành bức tường ngăn biến dạng.

Biến dạng xảy ra khi một khuyết tật nhỏ gọi là lệch mạng (Dislocation) trượt trong mạng tinh thể. Nếu mạng tinh thể lồi lõm, lệch mạng khó trượt. Nó giống như kéo xe trên bãi sỏi thì khó. Vì vậy, hợp kim có biến dạng mạng tinh thể lớn khó giãn ra và cứng chắc.

Nguyên lý thứ ba là khuếch tán chậm (Sluggish diffusion). Khuếch tán là chuyển động trong đó nguyên tử đổi chỗ. Trong hợp kim entropy cao, môi trường láng giềng của mỗi nguyên tử khác nhau. Có vị trí dễ dịch chuyển, có vị trí khó dịch chuyển. Người ta cho rằng sự không đều này, xét trung bình, làm chậm chuyển động của nguyên tử.

Nếu nguyên tử chuyển động chậm, cấu trúc vi mô không dễ thay đổi ngay cả ở nhiệt độ cao. Vì vậy, hạt tinh thể có thể lớn lên chậm và có thể chống creep tốt. Tuy nhiên, hiệu ứng này không phải lúc nào cũng xuất hiện mạnh trong mọi hợp kim entropy cao. Tốc độ khuếch tán thực tế thay đổi theo tổ hợp nguyên tố, nhiệt độ và cấu trúc tinh thể. Vì vậy, các nghiên cứu gần đây xem hiệu ứng này không phải là quy luật chung, mà là một tính chất cần kiểm chứng tùy từng trường hợp.

Nguyên lý thứ tư là hiệu ứng cocktail (Cocktail effect). Cocktail là đồ uống tạo ra hương vị mới bằng cách trộn nhiều nguyên liệu. Vị của từng nguyên liệu cộng lại không giống vị sau khi trộn. Kim loại cũng vậy. Khi nhiều nguyên tố cùng tồn tại, các tính chất khó giải thích bằng cách cộng riêng tính chất của từng nguyên tố sẽ xuất hiện.

Tính chất phi tuyến này xuất hiện ở độ bền, khả năng chống ăn mòn và độ dai. Vấn đề nằm ở chỗ con người khó biết trước hiệu ứng tổ hợp như vậy. Vì số lượng tổ hợp nguyên tố quá lớn. Trí tuệ nhân tạo được dùng như công cụ tìm ra các hiệu ứng tổ hợp ẩn như vậy từ dữ liệu.

Bốn nguyên lý này nối trực tiếp với vật liệu tên lửa. Entropy cấu hình giúp cấu trúc vi mô không bị tách ra ở nhiệt độ cao. Méo mạng tinh thể làm tăng độ bền. Khuếch tán chậm có thể có lợi cho khả năng chống creep. Hiệu ứng cocktail đem lại lợi ích bất ngờ như khả năng chống oxy hóa. Phải có cả bốn nguyên lý cùng hoạt động thì vật liệu mới chịu được môi trường cực hạn.

Tuy nhiên, các nguyên lý này không phải chìa khóa vạn năng. Trong một số tổ hợp, chúng xuất hiện mạnh; trong tổ hợp khác, chúng xuất hiện yếu. Vì vậy, dù biết nguyên lý, vẫn phải kiểm tra từng thành phần cụ thể. Nguyên lý chỉ ra hướng đi, còn tính toán và thí nghiệm đảm nhận việc xác nhận. Trí tuệ nhân tạo làm cho bước xác nhận này diễn ra nhanh hơn nhiều.

Cuối cùng, bốn nguyên lý này quy về một câu hỏi. Thành phần nào sẽ ổn định và còn lại dưới dạng một pha? Ngành học trả lời câu hỏi này là nhiệt động lực học cân bằng pha. Trong hai tiểu mục tiếp theo, chúng ta xem hai chỉ số giúp đưa ra phán đoán này.

#### 4.2.1 Tham số omega đo khả năng còn lại dưới dạng một pha

Nhiệt động lực học là ngành nghiên cứu năng lượng và độ ổn định. Một trạng thái có ổn định hay không được phán đoán bằng năng lượng tự do (Free energy). Trạng thái có năng lượng tự do thấp ổn định hơn. Giống như nước chảy từ nơi cao xuống nơi thấp, vật chất có xu hướng đi về trạng thái có năng lượng tự do thấp.

Năng lượng tự do do hai giá trị cùng quyết định. Một là enthalpy (Enthalpy). Enthalpy là giá trị liên quan đến việc các nguyên tử thích hay không thích nhau. Những nguyên tử hợp nhau sẽ muốn trộn lẫn, còn những nguyên tử không hợp nhau sẽ muốn tách ra. Giá trị còn lại là entropy đã xem ở phần trước. Entropy biểu thị mức độ hỗn loạn của sự trộn lẫn.

Cuộc giằng co giữa hai lực này thay đổi theo nhiệt độ. Nhiệt độ càng cao, lực của entropy càng mạnh. Điều này giống như người trong một căn phòng nóng sẽ tản ra tự do hơn. Vì vậy, ở nhiệt độ cao, dung dịch rắn hỗn loạn dễ trở nên ổn định. Đây là lý do hợp kim entropy cao thường còn lại tốt dưới dạng một pha ở nhiệt độ cao.

Entropy cấu hình tăng lên khi trộn các nguyên tố theo tỷ lệ đẳng mol. Tỷ lệ đẳng mol nghĩa là cho mỗi nguyên tố vào với lượng như nhau. Khi trộn năm nguyên tố với lượng như nhau, entropy cấu hình đạt giá trị cực đại. Số nguyên tố càng tăng, giá trị này càng lớn. Tên gọi entropy cao bắt nguồn từ tính chất có entropy cấu hình lớn này.

Tham số omega là giá trị dùng để xem nhanh khả năng hình thành dung dịch rắn một pha. Giá trị này xét cùng lúc ba yếu tố. Đó là nhiệt độ nóng chảy trung bình, entropy cấu hình và độ lớn của enthalpy trộn. Nếu ảnh hưởng có lợi của entropy lớn hơn ảnh hưởng bất lợi của enthalpy, giá trị omega tăng. Khi omega lớn, người ta xem hợp kim dễ còn lại dưới dạng một pha.

Ở đây cũng cần kiểm tra thêm chênh lệch kích thước nguyên tử. Nếu kích thước nguyên tử khác nhau quá nhiều, mạng tinh thể không chịu nổi và bị tách ra. Vì vậy, chênh lệch kích thước phải nhỏ hơn một giá trị nhất định thì dung dịch rắn một pha mới dễ hình thành. Các nhà nghiên cứu dùng quy tắc kinh nghiệm rằng dung dịch rắn một pha hình thành khi omega lớn và chênh lệch kích thước nhỏ.

Quy tắc này tiện lợi, nhưng không phải công cụ kết luận chắc chắn. Hợp kim thực tế chịu ảnh hưởng của trật tự cục bộ, tốc độ làm nguội, tạp chất và xử lý nhiệt. Cùng một thành phần, nhưng cấu trúc vi mô có thể khác nhau tùy cách chế tạo. Vì vậy, quy tắc omega chỉ được dùng cho bước lọc đầu tiên.

Việc phán đoán chính xác do các phương pháp cần nhiều tính toán hơn đảm nhận. Phương pháp tiêu biểu là CALPHAD. CALPHAD là phương pháp tích lũy dữ liệu năng lượng tự do của nhiều pha rồi tính cân bằng pha. Vào đó còn cộng thêm tính toán bằng lý thuyết phiếm hàm mật độ (Density Functional Theory, DFT) và thí nghiệm. Lý thuyết phiếm hàm mật độ là phương pháp tính toán lượng tử tính năng lượng từ hành vi của electron.

Hãy chuyển cuộc giảng co giữa enthalpy và entropy thành một ví dụ đời thường. Giả sử học sinh nam và học sinh nữ ngồi lẫn nhau trong lớp học. Lực khiến các nhóm thân nhau tụ lại giống với enthalpy. Lực khiến chỗ ngồi được trộn tự do giống với entropy. Vào giờ yên tĩnh, các nhóm tụ lại. Vào giờ sôi nổi, mọi người trộn lẫn tự do. Điều này giống như nhiệt độ càng cao thì lực muốn trộn lẫn càng mạnh.

Trí tuệ nhân tạo làm quá trình này nhanh hơn. Các chỉ số gọn như omega được dùng làm đặc trưng đầu vào của trí tuệ nhân tạo khi dữ liệu rất lớn. Trí tuệ nhân tạo nhìn các đặc trưng này và nhanh chóng sàng lọc thành phần nào sẽ còn lại dưới dạng một pha. Sau đó, chỉ những ứng viên đã được thu hẹp mới được kiểm tra bằng CALPHAD và thí nghiệm. Sự phối hợp này là chìa khóa để xử lý không gian thành phần rộng.

Các nghiên cứu gần đây đã nâng mạnh độ chính xác dự đoán này. Khi đưa nhiều đặc trưng vật lý vào cùng nhau, độ chính xác dự đoán pha tăng từ khoảng 70 phần trăm lên hơn 90 phần trăm. Ở mức này, có thể nhanh chóng lọc hơn 1 triệu thành phần. Như vậy, trí tuệ nhân tạo giảm rất nhiều công việc thủ công của con người. Thời gian và nguồn lực được dồn vào số ít ứng viên đã lọc ra như vậy.

#### 4.2.2 Ước lượng cấu trúc tinh thể bằng nồng độ electron hóa trị

Tiểu mục này bàn về chỉ số dùng để ước lượng hợp kim sẽ có cấu trúc tinh thể nào. Chỉ số đó là nồng độ electron hóa trị (Valence Electron Concentration, VEC). Nghĩa là số electron ngoài cùng trung bình. Trước hết, cần nói rõ cấu trúc tinh thể là gì. Nguyên tử kim loại không nằm tùy tiện. Chúng xếp thành mạng tinh thể có quy luật. Giống như xếp gạch theo quy tắc. Tính chất thay đổi theo cách xếp.

Trong vật liệu tên lửa, có hai cấu trúc quan trọng. Một là lập phương tâm khối (Body-Centered Cubic, BCC). Đây là cách sắp xếp có nguyên tử ở tám đỉnh của khối lập phương và ở chính giữa khối. Cấu trúc còn lại là lập phương tâm mặt (Face-Centered Cubic, FCC). Đây là cách sắp xếp có nguyên tử ở tám đỉnh và ở trung tâm của sáu mặt.

Hai cấu trúc này có tính chất khác nhau. Lập phương tâm khối thường có lợi cho độ bền ở nhiệt độ cao. Nhưng ở nhiệt độ phòng, nó dễ thể hiện tính giòn và dễ gãy. Lập phương tâm mặt thường có độ dẻo tốt, dễ kéo giãn. Nếu biết trước cấu trúc nào sẽ hình thành, việc thiết kế sẽ dễ hơn.

Nồng độ electron hóa trị là chỉ số ước lượng cấu trúc này. Electron hóa trị là electron ở phía ngoài của nguyên tử. Electron này tham gia vào liên kết. Giá trị trung bình của số electron ngoài cùng mà mỗi nguyên tử cung cấp, tính theo tỷ lệ thành phần, là nồng độ electron hóa trị. Giá trị này cho biết hướng của cấu trúc tinh thể.

Quy tắc kinh nghiệm như sau. Nếu giá trị này nhỏ hơn khoảng 6,87, lập phương tâm khối ổn định. Nếu lớn hơn 8, lập phương tâm mặt ổn định. Ở khoảng giữa, hai cấu trúc trộn lẫn. Hợp kim entropy cao chịu lửa thường có giá trị này nhỏ nên có nhiều lập phương tâm khối. Hợp kim Cantor nổi tiếng có giá trị này lớn nên là lập phương tâm mặt.

Có thể thấy ích lợi của quy tắc này cả trong nghiên cứu trí tuệ nhân tạo gần đây. Có một nghiên cứu tiền ấn phẩm được công bố vào đầu năm 2026. Nghiên cứu tiền ấn phẩm là nghiên cứu giai đoạn đầu chưa

qua bình duyệt. Nghiên cứu này dự đoán nhiệt độ nóng chảy của hợp kim entropy cao chịu lửa. Nó không đưa trực tiếp cấu trúc tinh thể vào, mà chỉ dùng các đặc trưng thu được từ thành phần nguyên tố. Những đặc trưng có ý nghĩa vật lý như nồng độ electron hóa trị đã đóng góp lớn vào dự đoán. Có thể xem tri thức vật lý do con người tích lũy và đặc trưng do trí tuệ nhân tạo chọn đã gặp nhau.

Trường hợp này cho thấy hai điều. Một là các chỉ số gọn vẫn còn hữu ích. Hai là trí tuệ nhân tạo có thể học mà không đi ngược vật lý. Hai điều này là ý tưởng xuyên suốt toàn bộ chương này.

Ở đây cần nói thêm một chút vì sao cấu trúc chi phối tính chất. Biến dạng xảy ra khi lệch mạng trượt bên trong mạng tinh thể. Lập phương tâm mặt có các mặt nơi nguyên tử xếp dày đặc, nên lệch mạng đi qua trơn tru. Vì vậy, nó dễ kéo giãn và không dễ gãy. Lập phương tâm khối không có những mặt dày đặc như vậy. Không phải vì số đường trượt ít, mà vì lệch mạng phải vượt qua rào cản năng lượng lớn mới đi qua được. Rào cản này nhạy với nhiệt độ. Vì vậy, lập phương tâm khối dễ gãy ở nhiệt độ phòng, còn khi nhiệt độ tăng, lệch mạng chuyển động dễ hơn và vật liệu duy trì được độ bền.

Khác biệt này nổi thẳng với lựa chọn vật liệu tên lửa. Với lớp lót vòi phun, độ bền ở nhiệt độ cao quan trọng hơn hết. Vì vậy, hợp kim entropy cao chịu lửa thuộc hệ lập phương tâm khối trở thành ứng viên. Nhưng nó cũng mang theo điểm yếu là tính giòn ở nhiệt độ phòng. Muốn giảm điểm yếu này, phải điều chỉnh thành phần rất tinh tế. Nồng độ electron hóa trị trở thành chiếc la bàn đầu tiên cho việc điều chỉnh đó.

Tuy nhiên, quy tắc nồng độ electron hóa trị cũng có giới hạn. Ranh giới 6,87 và 8 là giá trị kinh nghiệm rút ra từ một nhóm hợp kim kim loại chuyển tiếp nhất định. Nó không khớp nguyên xi với các nhóm hợp kim khác. Vì giá trị này là trung bình, nó cũng không chứa được môi trường cục bộ. Dù có cùng giá trị trung bình, cấu trúc thực tế vẫn có thể khác nhau tùy tổ hợp nguyên tố. Hơn nữa, quy tắc này không cho biết khi nào các hợp chất liên kim loại cứng và giòn sẽ hình thành. Pha Laves hoặc pha sigma là các ví dụ như vậy. Nếu các pha này trộn vào, hợp kim yếu đi. Vì vậy, chỉ số này cũng được dùng cho bước lọc đầu tiên rồi được kiểm tra bằng tính toán chính xác. Trí tuệ nhân tạo làm bước lọc này nhanh hơn và rộng hơn rất nhiều.

### 4.3 Trích xuất đặc trưng vật liệu từ bài báo khoa học

Trí tuệ nhân tạo không thể nhìn kim loại bằng mắt. Nó chỉ biết vật liệu thông qua các đặc trưng đã đổi thành con số. Đặc trưng này gọi là descriptor (Descriptor). Mục này xem cách trích xuất các đặc trưng này từ văn bản bài báo. Trước hết, cần nhìn xem không gian thành phần rộng đến mức nào. Khi trộn từ năm nguyên tố trở lên và thay đổi từng tỷ lệ, số trường hợp tăng bùng nổ. Nếu nguyên liệu cho vào cà phê chỉ có hai loại, số tổ hợp chỉ có vài trường hợp. Nếu nguyên liệu có mười loại, số tổ hợp khó đếm hết. Hợp kim đa thành phần đúng là ở trạng thái như vậy.

Không thể quét hết không gian rộng này bằng thí nghiệm. Làm ra một hợp kim và thử nghiệm nó tốn thời gian và tiền bạc. Tính toán CALPHAD cũng mất thời gian cho từng thành phần. Vì vậy, cần một người dẫn đường để nhanh chóng thu hẹp không gian rộng đó. Trí tuệ nhân tạo là người dẫn đường ấy.

Muốn dùng trí tuệ nhân tạo, phải đổi vật liệu thành con số. Descriptor chính là những con số đó. Bán kính nguyên tử, độ âm điện, nhiệt độ nóng chảy và số electron ngoài cùng là các descriptor tiêu biểu. Độ âm điện là lực mà nguyên tử hút electron. Các giá trị như vậy được tính theo thành phần rồi đưa vào trí tuệ nhân tạo.

Vấn đề là chỉ các con số như vậy thì chưa đủ. Dù cùng một thành phần, tính chất có thể khác nhau rất nhiều tùy cách chế tạo. Nhiệt độ xử lý nhiệt, thời gian giữ nhiệt và phương pháp làm nguội làm thay đổi cấu trúc vi mô. Thông tin quy trình như vậy khó chứa trong một con số. Nhưng trong bài báo, thông tin này được viết bằng văn bản.

Descriptor dẫn xuất từ ngôn ngữ tự nhiên (Natural language-derived descriptor) lấp vào khoảng trống này. Đây là giá trị đặc trưng được tạo ra khi trí tuệ nhân tạo đọc câu trong bài báo và đổi chúng thành đặc trưng. Có thể xem là chuyển văn bản con người đọc thành con số máy đọc. Các nghiên cứu gần đây đã dùng phương pháp này để đưa cả mô tả quy trình vào đặc trưng.

Một nghiên cứu tiền ấn phẩm được công bố năm 2026 là ví dụ tốt. Đây là kết quả của nhóm nghiên cứu tại Pennsylvania State University ở Mỹ. Họ đưa các câu mô tả quy trình xử lý nhiệt vào một mô hình ngôn ngữ lớn. Từ biểu diễn số của các câu đó, họ có thể khôi phục gần như nguyên vẹn các giá trị quy trình ban đầu. Và khi dùng biểu diễn này làm đặc trưng, dự đoán độ cứng của hợp kim được cải thiện khoảng 20 phần trăm. Nghiên cứu này vẫn là nghiên cứu giai đoạn đầu chưa qua bình duyệt.

Khai phá tài liệu là công việc thu thập các đặc trưng như vậy ở quy mô lớn. Nó tự động rút ra nhiệt độ chuyển pha, giới hạn độ hòa tan, độ bền và độ dẻo từ vô số bài báo kim loại. Trước đây, con người lập bảng bằng tay. Giờ đây, mô hình ngôn ngữ lớn đọc câu và điền bảng. Dữ liệu được gom như vậy trở thành nền tảng huấn luyện trí tuệ nhân tạo.

Cũng có nghiên cứu dùng trực tiếp mô hình ngôn ngữ để dự đoán tính chất. Đó là kết quả đăng trên tạp chí khoa học năm 2025. Nghiên cứu này dùng cấu trúc mô hình ngôn ngữ gọi là transformer.

Transformer là trí tuệ nhân tạo đọc quan hệ giữa các từ trong câu. Nó xử lý thành phần và điều kiện của hợp kim như văn bản rồi dự đoán tính chất. Đây là trường hợp công nghệ xử lý ngôn ngữ chuyển sang nghiên cứu vật liệu.

Gần đây cũng xuất hiện nghiên cứu tạo cơ sở dữ liệu hợp kim entropy cao bằng mô hình ngôn ngữ lớn. Đó là kết quả đăng trên tạp chí khoa học năm 2026. Cơ sở dữ liệu này tự động sắp xếp thành phần và tính chất từ các bài báo. Nó rút ngắn rất nhiều công việc vốn mất nhiều thời gian nếu làm bằng tay. Dữ liệu tích lũy như vậy trở thành nguyên liệu nền cho các mô hình dự đoán sẽ xem ở mục tiếp theo.

Phương pháp này cũng có giới hạn. Mỗi bài báo có điều kiện đo và cách ký hiệu khác nhau. Nếu máy đọc sai câu, giá trị sai sẽ lẫn vào dữ liệu. Vì vậy con người phải kiểm tra lại dữ liệu đã thu thập. Hiện nay, chuẩn làm việc là trí tuệ nhân tạo thu thập và con người xác nhận.

Chất lượng dữ liệu quyết định chất lượng dự đoán. Nếu học từ dữ liệu sai, câu trả lời sai sẽ xuất hiện. Nếu bản thiết kế sai, dù thợ mộc giỏi cũng không thể xây một ngôi nhà vững chắc. Vì vậy, bước thu thập dữ liệu quan trọng không kém bước dự đoán. Dữ liệu tốt là nền tảng của trí tuệ nhân tạo tốt.

Gần đây, trí tuệ nhân tạo còn tự mở rộng quy tắc. Một nghiên cứu đăng trên tạp chí học thuật năm 2026 đã giới thiệu một khung trí tuệ nhân tạo được bổ sung tri thức. Nghiên cứu này thoát khỏi ràng buộc phải trộn các nguyên tố với lượng bằng nhau. Nó tự do thay đổi tỷ lệ nguyên tố theo mục tiêu và thiết kế hợp kim entropy cao một pha. Đây là kết quả có được khi đưa tri thức luyện kim mà con người đã tích lũy vào trí tuệ nhân tạo.

Xu hướng này cho thấy sự hợp tác giữa trí tuệ nhân tạo và tri thức của con người. Nếu chỉ dùng trí tuệ nhân tạo, dự đoán sẽ dao động ngoài phạm vi dữ liệu. Nếu chỉ dùng quy tắc của con người, rất khó quét một không gian rộng. Khi kết hợp hai thứ, chúng bù đắp điểm yếu cho nhau. Khi khai phá văn và bổ sung tri thức đi cùng nhau, thiết kế trở nên vững hơn.

Sự kết hợp với CALPHAD cũng quan trọng. Trí tuệ nhân tạo có thể mô phỏng nhanh phép tính CALPHAD hoặc chọn ứng viên cho phép tính tiếp theo. Nếu nổi dữ liệu thu thập từ văn, phép tính CALPHAD, lý thuyết phiếm hàm mật độ (DFT) và thí nghiệm thành một mạch, thiết kế ứng viên sẽ nhanh hơn. Pipeline tích hợp này là phương pháp thực tế để xử lý không gian thành phần rộng.

## 4.4 Mạng nơ-ron đồ thị biểu diễn hợp kim mất trật tự

Ta sẽ xem trí tuệ nhân tạo biểu diễn sự mất trật tự của hợp kim entropy cao như thế nào. Công cụ cốt lõi là mạng nơ-ron đồ thị (Graph Neural Network, GNN). Đây là phương pháp biến vật liệu thành hình gồm điểm và đường rồi học từ đó. Trước hết cần hiểu đồ thị là gì. Đồ thị là hình biểu diễn quan hệ bằng điểm và đường. Sơ đồ tuyến tàu điện ngầm là một đồ thị. Nhà ga là điểm, đường ray là đường. Tinh thể cũng có thể được vẽ thành đồ thị. Nguyên tử là điểm, liên kết giữa nguyên tử là đường.

Mạng nơ-ron đồ thị đọc hình này để dự đoán tính chất của vật liệu. Mỗi điểm chứa thông tin của nguyên tử. Mỗi đường chứa khoảng cách và hướng giữa các nguyên tử. Mạng nơ-ron hiểu vật liệu bằng cách để các điểm lân cận trao đổi thông tin với nhau. Việc trao đổi thông tin này gọi là truyền thông điệp (Message passing). Nó giống như trao đổi giấy nhắn với bạn ngồi cạnh để nắm bầu không khí của cả lớp.

Phương pháp này phù hợp với đơn tinh thể lý tưởng. Đơn tinh thể là tinh thể mà mỗi vị trí được lấp đầy đúng bằng một nguyên tố. Nó có quy tắc rõ ràng như tinh thể muối. Các mạng nơ-ron đồ thị tinh thể trước đây được xây dựng trên giả định này. Khi đặt một nguyên tố vào mỗi điểm, phép tính trở nên gọn gàng.

Vấn đề nằm ở chỗ hợp kim entropy cao không như vậy. Trong hợp kim này, nhiều nguyên tố trộn theo xác suất tại một vị trí. Một vị trí có thể là niobi, cũng có thể là molypden, cũng có thể là tantali. Điều này gọi là chiếm chỗ một phần (Fractional occupancy). Đó là trạng thái một vị trí đồng thời có nhiều khả năng.

Ở đây còn có trật tự tầm ngắn cục bộ (Short-Range Order, SRO). Các nguyên tử không trộn hoàn toàn tùy tiện. Có những nguyên tố thích ở cạnh nhau, và có những nguyên tố tránh nhau. Sự ưa thích ở khoảng cách gần này ảnh hưởng lớn đến tính chất. Mô hình trước đây không chứa được trật tự tinh vi này.

Mạng nơ-ron đồ thị phân số tinh thể là một nỗ lực nhằm giải quyết vấn đề này. Ý tưởng theo hướng này như sau. Thay vì đặt một nguyên tố vào mỗi điểm, ta đặt một vector thành phần. Vector thành phần là danh sách số ghi xác suất mỗi nguyên tố có mặt ở vị trí đó. Ví dụ, một vị trí được biểu diễn là niobi 25 phần trăm, molypden 25 phần trăm, tantali 50 phần trăm. Làm như vậy sẽ chứa được chiếm chỗ một phần một cách tự nhiên.

Điều mà nghiên cứu tiền đề mang tên này thật sự cho thấy là hẹp và rõ. Nghiên cứu này đưa cùng lúc môi trường cục bộ khoảng 16 nguyên tử và phân suất nguyên tố của toàn hợp kim vào mô hình để dự đoán năng lượng của tinh thể. Đó là kết quả học từ khoảng một nghìn cấu trúc. Đây không phải là nghiên cứu khớp cùng lúc trật tự tầm ngắn, điểm nóng chảy, enthalpy trộn và độ bền nhiệt độ cao. Dự đoán các tính chất này cùng nhau là hướng cần giải trong tương lai, chứ chưa phải thành tựu đã đạt được. Cách thêm thông tin cấp nguyên tố vào đường để chứa trật tự tầm ngắn cũng đang ở cùng giai đoạn đề xuất.

Ta cần hiểu truyền thông điệp dẫn đến học như thế nào. Ban đầu, mỗi điểm chỉ biết thông tin của chính nó. Sau một lần trao đổi thông tin, nó biết đến các điểm lân cận. Nếu lặp lại nhiều lần, nó biết đến nơi xa hơn. Mô hình dùng thông tin đã mở rộng như vậy để dự đoán tính chất của toàn vật liệu. Khi dự đoán lệch khỏi giá trị thực, các giá trị trong mạng nơ-ron được sửa. Mô hình trở nên thông minh hơn bằng cách lặp lại việc sửa này vô số lần.

Một ưu điểm khác là một mô hình có thể dự đoán nhiều tính chất cùng nhau. Điều này gọi là dự đoán đa mục tiêu. Điểm nóng chảy và độ bền có liên quan với nhau. Thông tin học được khi học một tính chất giúp dự đoán tính chất khác. Nó giống như khi học nhiều môn cùng nhau thì hiểu biết giữa các môn sâu hơn.

Tuy nhiên, hướng này vẫn ở giai đoạn đầu. Mạng nơ-ron đồ thị phân số tinh thể chỉ mới được đề xuất gần đây như một nghiên cứu tiền đề. Ô lớn, trật tự tầm ngắn phức tạp và lịch sử xử lý nhiệt thực tế vẫn

là các vấn đề khó. Vì vậy, dự đoán của mô hình này phải được xác nhận bằng thí nghiệm và mô phỏng nguyên tử.

Trong khi đó, các nghiên cứu trí tuệ nhân tạo khác xử lý trật tự tầm ngắn đang tăng nhanh. Một nghiên cứu đăng trên tạp chí học thuật năm 2025 đã so sánh các phương pháp tái hiện trật tự tầm ngắn bằng thể năng học máy. Thể năng học máy là công cụ tính lực giữa các nguyên tử bằng trí tuệ nhân tạo. Nghiên cứu này cho thấy ngay cả khi độ chính xác năng lượng cao, mô hình vẫn có thể không khớp tốt tính chất. Điều đó có nghĩa là cách chọn dữ liệu học quyết định dự đoán tính chất.

Cùng năm đó, một nghiên cứu khác đã tăng tốc tính toán Monte Carlo bằng trí tuệ nhân tạo. Monte Carlo là phương pháp tính toán tìm trạng thái cân bằng bằng các thử nghiệm ngẫu nhiên. Nghiên cứu này dùng mô hình đại diện trí tuệ nhân tạo để làm lộ ra cấu trúc nano ở quy mô hàng chục tỷ nguyên tử. Có thể xem đây là việc mở ra con đường quan sát trật tự tinh vi trong hợp kim entropy cao ở quy mô lớn. Mạng nơ-ron đồ thị phân số tinh thể cũng đang phát triển trên dòng chảy này.

Ta cần hiểu vì sao biểu diễn mất trật tự như vậy quan trọng đối với vật liệu tên lửa. Độ bền của lớp lót vòi phun phụ thuộc vào việc lệch mạng bị chặn tốt đến đâu. Trật tự tầm ngắn thay đổi độ khó của con đường mà lệch mạng đi qua. Khi một số nguyên tử tụ lại với nhau, chúng chặn lệch mạng mạnh hơn. Vì vậy, dù cùng thành phần, độ bền vẫn thay đổi theo trật tự tầm ngắn.

Trật tự tinh vi này thay đổi theo cách chế tạo. Nếu làm nguội nhanh, sự mất trật tự bị đóng băng. Nếu làm nguội chậm, các nguyên tử ổn định vị trí và trật tự hình thành. Nhiệt độ và thời gian xử lý nhiệt cũng thay đổi trật tự. Vì vậy, để khớp tính chất, phải xem xét không chỉ thành phần mà cả quy trình.

Trí tuệ nhân tạo được dùng để học các quan hệ đan xen này. Nó nối thành phần, quy trình, trật tự và tính chất cùng lúc. Con người khó xử lý đồng thời số biến lớn như vậy. Nếu có đủ dữ liệu, trí tuệ nhân tạo tìm ra các liên kết đó. Mạng nơ-ron đồ thị phân số tinh thể là nỗ lực chứa các liên kết này ở cấp độ sắp xếp nguyên tử.

Hướng này vẫn ở giai đoạn đầu. Hiện nay cả độ chính xác dự đoán lẫn lượng dữ liệu đều còn thiếu. Khi dữ liệu thí nghiệm tích lũy nhiều hơn và tính toán nhanh hơn, dự đoán sẽ tốt hơn. Cho đến lúc đó, quy trình xác nhận dự đoán của trí tuệ nhân tạo bằng thí nghiệm bắt buộc phải đi kèm.

## 4.5 Động lực học phân tử quy mô lớn nhìn đến từng nguyên tử

Phương pháp mô phỏng chuyển động của từng nguyên tử bằng máy tính gọi là động lực học phân tử (Molecular Dynamics, MD). Đây là công cụ quan sát quá trình kim loại bị phá hủy ở cấp độ nguyên tử. Chìa khóa để mở rộng phép tính này lên quy mô rất lớn là một dạng trí tuệ nhân tạo gọi là thể năng tiến hóa thần kinh (Neuroevolution Potential, NEP). Trước hết cần hiểu vì sao cần quy mô lớn. Vết nứt của kim loại bắt đầu từ vài nguyên tử. Khe nhỏ đó lớn lên, lan đến ranh giới hạt tinh thể, rồi cuối cùng cắt đứt cấu trúc. Để hiểu quá trình này, phải quan sát cả thứ rất nhỏ và thứ khá lớn cùng nhau. Hiện tượng như vậy gọi là hiện tượng đa thang đo (Multi-scale).

Lý thuyết phiếm hàm mật độ (DFT) truyền thống khó làm việc này. Phép tính này chính xác nhưng tốn nhiều tính toán. Khó vượt quá vài trăm nguyên tử. Tổ chức thực tế nơi vết nứt lan rộng có hơn hàng chục tỷ nguyên tử. Vì vậy, chỉ bằng tính toán lượng tử chính xác thì không thể thấy toàn cảnh.

Thể năng kinh nghiệm có vấn đề ngược lại. Thể năng là quy tắc tính lực tác dụng giữa các nguyên tử. Thể năng kinh nghiệm nhẹ và nhanh. Nhưng trong hợp kim phức tạp có hơn năm nguyên tố, độ chính xác giảm. Mỗi nguyên tố có môi trường lân cận khác nhau, nhưng nó không chứa được sự tinh vi đó.

Vì vậy cần một công cụ kết hợp ưu điểm của hai phương pháp. Đó là công cụ có cả độ chính xác của tính toán lượng tử và tốc độ của quy tắc kinh nghiệm. Thứ lấp vào vị trí này là thể năng học máy. Trí tuệ

nhân tạo học dữ liệu tính toán lượng tử để dự đoán lực. Thế năng tiến hóa thần kinh là công cụ mạnh về tốc độ trong nhóm này.

Cũng cần hiểu động lực học phân tử là gì. Phép tính này áp dụng định luật chuyển động Newton cho tất cả nguyên tử. Nó tính lực tác dụng lên từng nguyên tử rồi di chuyển nguyên tử trong một khoảng thời gian rất ngắn. Quá trình này được lặp lại vô số lần. Khi đó, ta có thể xem chuyển động của các nguyên tử theo thời gian giống như một bộ phim.

#### 4.5.1 Thế năng tiến hóa thần kinh đạt cả độ chính xác và tốc độ

Thế năng tiến hóa thần kinh là công cụ kết hợp hai ý tưởng. Một là mạng nơ-ron. Mạng nơ-ron là trí tuệ nhân tạo học quy tắc từ dữ liệu. Ý tưởng còn lại là chiến lược tiến hóa. Chiến lược tiến hóa là phương pháp tạo nhiều ứng viên, chọn ứng viên tốt hơn rồi tinh chỉnh. Đây là cách học mô phỏng chọn lọc trong tự nhiên.

Thế năng này trước hết biến môi trường quanh nguyên tử thành số. Nó tạo đặc trưng từ việc một nguyên tử cách các láng giềng bao xa và nằm ở góc nào. Các đặc trưng này được đưa vào mạng nơ-ron để dự đoán năng lượng của một nguyên tử. Cộng năng lượng của tất cả nguyên tử sẽ có năng lượng toàn phần. Và từ sự thay đổi năng lượng, ta thu được lực tác dụng lên nguyên tử.

Ta cần hiểu vì sao phải xem cả khoảng cách và góc. Nếu chỉ nhìn khoảng cách, ta chỉ biết các nguyên tử ở gần nhau đến mức nào. Nếu nhìn cả góc, ta biết các nguyên tử được đặt theo hình dạng nào. Dù cùng khoảng cách, bố trí tam giác và bố trí thẳng hàng có tính chất khác nhau. Vì vậy phải đưa cả khoảng cách và góc vào thì mới chứa đúng môi trường.

Cách làm này chạm đến bản chất của liên kết hóa học. Nguyên tử liên kết khác nhau tùy theo môi trường xung quanh. Nếu láng giềng dày đặc, liên kết mạnh; nếu thưa, liên kết yếu. Thế năng tiến hóa thần kinh học quan hệ này từ dữ liệu. Với quan hệ đã học, nó nhanh chóng ước lượng năng lượng của cách sắp xếp mới.

Ưu điểm của cách này nằm ở việc đạt cả độ chính xác và tốc độ. Mạng nơ-ron học độ chính xác gần với tính toán lượng tử. Bản thân phép tính nhẹ như quy tắc kinh nghiệm. Vì vậy có thể thu được lực chính xác một cách nhanh chóng. Đây là chìa khóa mở ra tính toán quy mô lớn.

Bí quyết tốc độ nằm ở bộ xử lý đồ họa (Graphics Processing Unit, GPU). Thiết bị này ban đầu được tạo ra để vẽ màn hình trò chơi. Nó giải tô đồng thời vô số điểm trên màn hình. Khả năng xử lý song song này cũng rất hợp với tính toán nguyên tử. Vì nó có thể tính lực của rất nhiều nguyên tử cùng lúc.

Thế năng tiến hóa thần kinh được thiết kế để chạy tốt trên thiết bị này. Nó được dùng cùng chương trình dành cho bộ xử lý đồ họa tên là GPUMD. Sự kết hợp này đã được dùng rộng rãi trong nghiên cứu dẫn nhiệt, khuyết tật và biến dạng nhiệt độ cao. Gần đây cũng đã có nghiên cứu tiếp theo làm phương pháp học nhanh hơn. Những cải tiến như vậy giúp phép tính lớn hơn trở nên khả thi.

Cũng cần nêu quá trình tạo thế năng này. Trước hết, lý thuyết phiếm hàm mật độ (DFT) được dùng để tính năng lượng và lực của nhiều cách sắp xếp nguyên tử. Các giá trị này trở thành sách giáo khoa của trí tuệ nhân tạo. Trí tuệ nhân tạo nhìn vào sách giáo khoa này và tinh chỉnh dự đoán của mình. Chiến lược tiến hóa chọn dần các giá trị theo hướng dự đoán khớp tốt với dữ liệu. Khi quá trình học này kết thúc, một máy tính lực nhanh và chính xác được hoàn thành.

Học chủ động cũng được dùng trong quá trình học. Trí tuệ nhân tạo tự tìm ra những cách sắp xếp mà nó không tự tin. Chỉ các cách sắp xếp đó được chọn để tính mới bằng lý thuyết phiếm hàm mật độ (DFT). Dữ liệu thu được như vậy lại được đưa vào học. Đây là cách thông minh để lấp đầy phạm vi rộng bằng ít phép tính.

Tuy nhiên, câu trả lời nhanh không phải lúc nào cũng là câu trả lời đúng. Thế năng trí tuệ nhân tạo chính xác trong phạm vi nó đã học. Với thành phần hoặc nhiệt độ không có trong dữ liệu học, nó có thể sai. Hợp kim entropy cao có môi trường thay đổi lớn chỉ với một thay đổi nhỏ trong tổ hợp nguyên tố. Vì vậy tính đại diện của dữ liệu học quan trọng hơn mọi thứ khác. Cần có quy trình kiểm tra riêng độ bất định của dự đoán.

#### 4.5.2 Biến dạng và vết nứt nhìn bằng tính toán quy mô lớn

1 nghìn tỷ là một con số khó hình dung. Nếu xếp 1 nghìn tỷ nguyên tử, kích thước sẽ gần với tổ chức thực tế. Điều đó có nghĩa là có thể vẽ một mẫu ở đơn vị micromet đến từng nguyên tử. Trong phép tính nhỏ, chỉ thấy một vết nứt và vài ranh giới. Trong phép tính lớn, ta thấy nhiều ranh giới và khuyết tật ảnh hưởng lẫn nhau.

Một nghiên cứu sơ bộ được công bố năm 2026 đã vượt qua ngưỡng này. Nghiên cứu này tính toán 1,6 nghìn tỷ nguyên tử bằng thế năng tiến hóa thần kinh. Trên một siêu máy tính mới của Trung Quốc, nghiên cứu dùng đồng thời khoảng 45.000 thiết bị tính toán. Điều nghiên cứu này cho thấy là quy mô và tốc độ tính toán. Nhóm nghiên cứu báo cáo rằng họ chạy cùng phép tính nhanh hơn nhiều so với kỷ lục cao nhất trước đó. Đây không phải là nghiên cứu đã thực sự đưa ra kết quả phân tích lệch mạng và vết nứt của hợp kim entropy cao ở quy mô này. Trình diễn mở ra quy mô và việc dùng quy mô đó để làm sáng tỏ điều gì là hai vấn đề khác nhau. Nghiên cứu này vẫn là nghiên cứu giai đoạn đầu, chưa qua bình duyệt.

Ta hãy tạm so sánh con số này có nghĩa gì. Trước đây, tính toán 1 triệu nguyên tử đã là việc lớn. Vài năm trước, quy mô đã vượt 100 triệu nguyên tử và chẳng bao lâu đạt 10 tỷ nguyên tử. Nay nó đã vượt 1 nghìn tỷ nguyên tử. Kích thước có thể tính toán đang tăng rất nhanh. Tuy nhiên, tính toán 1 nghìn tỷ nguyên tử là một trình diễn mang tính biểu tượng, dùng đến tận cùng một siêu máy tính cực lớn. Trong thiết kế hợp kim thực tế, các phép tính dùng hàng ngày thường ở quy mô từ vài nghìn đến vài trăm nghìn nguyên tử. Thế năng trí tuệ nhân tạo và máy tính nhanh đang dẫn dắt sự tăng trưởng này.

Lợi ích của quy mô lớn nằm ở thống kê. Càng có nhiều nguyên tử thì càng bắt được cả những sự kiện hiếm. Mầm của vết nứt xuất hiện rất hiếm. Trong phép tính nhỏ, ta có thể không nhìn thấy mầm đó. Trong phép tính lớn, ta quan sát được nhiều mầm cùng lúc. Nhờ vậy dự đoán tiến gần thực tế hơn.

Trước hết cần nói rõ vì sao phép tính này hữu ích cho vật liệu tên lửa. Cấu trúc thực của lớp lót vòi phun có số nguyên tử nhiều đến mức không thể đếm xuể. Phải biết điều gì xảy ra trong cấu trúc lớn này thì mới biết tuổi thọ. Phép tính lớn cho ta một bức tranh gần với cấu trúc thực đó. Vì vậy, nó có thể được dùng để ước lượng thành phần nào sẽ chịu được lâu hơn.

Về nguyên lý, có ba sự kiện có thể quan sát bằng động lực học phân tử quy mô lớn như vậy. Những nội dung dưới đây là mục tiêu mà phương pháp này nhắm tới, không phải là kết quả mà trình diễn nói trên đã làm sáng tỏ. Thứ nhất là sự tạo mầm lệch mạng. Lệch mạng là khuyết tật nhỏ nơi biến dạng bắt đầu. Ta theo dõi theo thời gian thực lệch mạng hình thành và trượt như thế nào trong mạng lập phương tâm khối. Trong hợp kim entropy cao, dao động của thành phần cục bộ giữ lệch mạng lại. Sự ghim giữ này là nguyên lý làm tăng độ bền.

Thứ hai là trượt biên hạt. Biên hạt là ranh giới giữa các hạt tinh thể. Khi chịu lực ở nhiệt độ cao, ranh giới này trượt. Phép tính lớn đo ở cấp nguyên tử sự trượt này và sự tụ lại của khoảng rỗng. Hành vi này chi phối tuổi thọ creep.

Thứ ba là sự phát triển vết nứt. Đầu mút vết nứt là phần đi trước nhất của khe nứt. Ở đây, sắp xếp nguyên tử thay đổi và biến dạng xảy ra. Phép tính lớn quan sát vết nứt lớn lên và dừng lại như thế nào. Đây là manh mối để ước lượng thành phần nào chịu nứt tốt.

Ba sự kiện này nối tiếp nhau. Lệnh mạng tích tụ rồi dồn về biên hạt. Biên hạt trượt và tạo khoảng rỗng. Các khoảng rỗng đó nối với nhau thành vết nứt. Trong phép tính nhỏ, rất khó thấy mỗi nối này. Trong phép tính lớn, có thể nhìn trọn từ đầu đến cuối trong một khung hình.

Cần nói rõ vì sao việc nhìn thấy mỗi nối này hữu ích. Tuổi thọ của vật liệu được quyết định bởi một mắt xích yếu. Sợi xích đứt ở mắt xích yếu. Nếu biết giai đoạn nào sụp trước, ta có thể gia cố chỗ đó. Tính toán quy mô lớn giúp tìm trước mắt xích yếu nằm ở đâu.

Thông tin này hữu ích cho nghiên cứu lớp lót vòi phun tên lửa. Lớp lót vòi phun chịu đồng thời khí nóng và sốc nhiệt. Nếu ước lượng trước thí nghiệm thành phần nào chịu được lâu hơn, ta tiết kiệm thời gian. Tính toán quy mô lớn đóng vai trò như chiếc cân để thu hẹp các ứng viên.

Tuy nhiên, phép tính này cũng có điều kiện đi kèm. Độ chính xác của kết quả phụ thuộc vào chất lượng của thể năng trí tuệ nhân tạo. Trong tình huống mà thể năng chưa học được, phép tính cũng sai. Vì vậy, kết quả của phép tính lớn phải được kiểm chứng bằng tính toán chính xác ở quy mô nhỏ và bằng thí nghiệm. Quy mô lớn và đúng là hai vấn đề khác nhau.

## 4.6 Trí tuệ nhân tạo thiết kế kim loại không trái với định luật vật lý

Chỉ học từ dữ liệu thuần túy có thể tạo ra dự đoán nguy hiểm. Phương pháp ngăn rủi ro đó bằng định luật vật lý được gọi là mạng nơ-ron dựa trên vật lý (Physics-Informed Neural Network, PINN). Ta hãy xem phương pháp này được áp dụng thế nào vào luyện kim, tức ngành khoa học xử lý kim loại. Trước hết cần chỉ ra điểm yếu của học dữ liệu thuần túy. Trí tuệ nhân tạo trả lời tốt trong phạm vi dữ liệu đã học. Nhưng ở vùng không có dữ liệu, nó đưa ra câu trả lời lạc hướng. Những vùng này được gọi là ngoài phân bố học tập (Out-of-Distribution). Nó giống như một học sinh gặp bài nằm ngoài phạm vi ôn thi.

Trong thiết kế kim loại, câu trả lời lạc hướng này vi phạm định luật vật lý. Ví dụ, nhiệt độ tăng nhưng giá trị ổn định lại chuyển động ngược chiều. Nó cũng có thể dự đoán rằng thể tích tăng nhưng năng lượng mạng tinh thể giảm. Những câu trả lời như vậy không thể xảy ra trong tự nhiên. Nếu tin nguyên xi các dự đoán đó, thiết kế sẽ đi lệch.

Mạng nơ-ron dựa trên vật lý ngăn vấn đề này ngay trong giai đoạn học. Cách làm là đưa định luật vật lý vào hàm mất mát. Hàm mất mát là thước đo trí tuệ nhân tạo sai bao nhiêu. Thông thường, nó chỉ đo chênh lệch giữa dự đoán và dữ liệu thực. Ở đây, ta cộng thêm mức độ vi phạm định luật vật lý.

Làm vậy, trí tuệ nhân tạo cố gắng giữ đồng thời hai việc. Một là khớp với dữ liệu. Hai là tuân thủ định luật vật lý. Ngay cả trong vùng không có dữ liệu, định luật vẫn định hướng cho câu trả lời. Nó giống như một học sinh gặp bài ngoài phạm vi ôn thi nhưng vẫn thử giải bằng nguyên lý cơ bản.

Trong thiết kế kim loại, ta đưa vào những định luật nào. Một ví dụ là năng lượng tự do phải thay đổi thế nào theo nhiệt độ. Ta cũng đưa vào quan hệ rằng độ bền được quyết định bởi tổng của nhiều đóng góp tăng bền. Các điều kiện mà cân bằng pha phải tuân thủ cũng được đưa vào. Những định luật này làm giảm sự tự do của trí tuệ nhân tạo. Sự tự do bị giảm ấy lại tạo ra câu trả lời đáng tin hơn.

Ta hãy giải thích thêm cách đưa định luật vào hàm mất mát. Định luật vật lý thường được viết bằng phương trình. Khi đưa giá trị dự đoán vào phương trình đó, hai vế phải khớp nhau. Nếu không khớp, phần chênh lệch còn lại. Phần chênh lệch này được gọi là phần dư. Cách làm là phạt khi phần dư lớn.

Khi làm vậy, trí tuệ nhân tạo cố tránh bị phạt. Nó cố khớp với dữ liệu và đồng thời giữ đúng phương trình. Nó giống như một học sinh khi nộp bài phải đúng cả công thức tính và đơn vị. Học sinh đó mạnh hơn học sinh chỉ học thuộc đáp án khi gặp bài ứng dụng. Đây là lý do mạng nơ-ron dựa trên vật lý mạnh ở ngoài dữ liệu.

Giá trị của phương pháp này tăng lên khi dữ liệu ít. Vật liệu cực hạn dùng cho tên lửa có dữ liệu thí nghiệm rất quý. Mỗi lần thử nghiệm tốn nhiều chi phí và thời gian. Khi dữ liệu ít, học thuần túy dễ dao động. Định luật vật lý bù vào phần dữ liệu còn thiếu.

Cần phân biệt hai phương pháp này. Nghiên cứu dự đoán điểm nóng chảy đã xem ở phần trước không phải là mạng nơ-ron dựa trên vật lý. Nghiên cứu đó là mô hình hồi quy chọn và đưa vào các đặc trưng có ý nghĩa vật lý. Nó không đưa trực tiếp phương trình vật lý vào hàm mất mát. Hồi quy dùng đặc trưng tốt và học đưa định luật vật lý vào hàm mất mát là hai phương pháp khác nhau. Tuy vậy, cả hai hướng tới cùng một mục tiêu. Đó là làm cho dự đoán của trí tuệ nhân tạo không trái với luyện kim. Khi đặc trưng tốt đi cùng ràng buộc vật lý, ta tiến gần mục tiêu đó hơn.

#### 4.6.1 Tập hợp công cụ cho thiết kế lớp lót vôi phun siêu cao nhiệt

Tiểu mục này xem các công cụ nói trên được dùng cùng nhau thế nào trong thiết kế thực tế. Mục tiêu là hợp kim entropy cao chịu lửa dùng cho lớp lót vôi phun siêu cao nhiệt. Trọng tâm đặt vào quy trình thiết kế và kiểm chứng hơn là một thành quả hoàn chỉnh. Ta cũng xem nghiên cứu gần đây đã đi đến đâu và còn lại điều gì. Một hợp kim tốt cho lớp lót vôi phun phải đồng thời thỏa mãn nhiều tính chất. Điểm nóng chảy phải cao. Nếu mật độ quá lớn, tên lửa sẽ nặng. Độ bền phải được duy trì ở nhiệt độ cao. Bề mặt phải chịu được khi gặp oxy. Lớp phủ và giãn nở nhiệt phải khớp nhau. Chỉ một tính chất tốt thì không thể dùng làm linh kiện.

Việc nắm đồng thời nhiều mục tiêu như vậy là quá sức đối với con người. Nếu cho nhiều tungsten, điểm nóng chảy tăng. Nhưng mật độ cũng tăng và vật liệu có thể dễ gãy. Nếu cho nhôm hoặc chromium, vật liệu chống oxy hóa tốt hơn. Nhưng các tính chất khác có thể bị dao động. Việc xử lý các mục tiêu đan xen như vậy được gọi là tối ưu hóa đa mục tiêu.

Trong tên lửa, mật độ là vấn đề rất nhạy. Linh kiện nặng thì toàn bộ tên lửa nặng. Tên lửa nặng đi được ít xa hơn với cùng lượng nhiên liệu. Vì vậy, dù độ bền tốt, nếu quá nặng thì không dùng được. Phải đặt điểm nóng chảy, độ bền và trọng lượng lên cùng một chiếc cân.

Trong phép cân đo này không có một đáp án duy nhất. Nếu nhường một chút độ bền, có thể giảm trọng lượng. Nếu nhường một chút trọng lượng, có thể tăng độ bền. Như vậy sẽ xuất hiện một nhóm ứng viên đánh đổi lẫn nhau. Con người đặt nhóm này trước mặt rồi chọn một phương án phù hợp với mục đích. Trí tuệ nhân tạo vẽ nhanh nhóm đó.

Trí tuệ nhân tạo xử lý đồng thời các mục tiêu đan xen này. Nó biến thành phần thành con số bằng các descriptor đã xem ở trên. Nó dùng mô hình dự đoán để nhanh chóng sàng lọc điểm nóng chảy và độ bền. Nó dùng ràng buộc vật lý để loại các ứng viên vô lý. Nó dùng tính toán nguyên tử quy mô lớn để xem biến dạng và vết nứt. Dòng công việc này thu hẹp mạnh các ứng viên.

Các nghiên cứu gần đây cho thấy ích lợi của dòng công việc này. Một nghiên cứu đăng trên tạp chí năm 2024 đã thiết kế hợp kim entropy cao chịu lửa dùng cho siêu cao nhiệt bằng học máy. Nghiên cứu thu hẹp không gian thành phần rộng bằng trí tuệ nhân tạo và chọn ứng viên triển vọng. Đó là nỗ lực nhằm tăng đồng thời độ bền và độ dẻo. Những nghiên cứu như vậy cho thấy thiết kế bằng trí tuệ nhân tạo vận hành được ở giai đoạn phòng thí nghiệm.

Trí tuệ nhân tạo cũng đi vào vấn đề oxy hóa. Có một nghiên cứu sơ bộ được công bố cuối năm 2025. Nghiên cứu này dùng học chủ động để tìm kiếm hợp kim cô đặc phức hợp chịu lửa có khả năng chống oxy hóa ở nhiệt độ cao. Học chủ động là phương pháp để trí tuệ nhân tạo tự chọn thành phần cần thử tiếp theo. Đây là con đường tìm ứng viên tốt bằng ít thí nghiệm. Nghiên cứu này cũng vẫn là nghiên cứu giai đoạn đầu, chưa qua bình duyệt.

Tuy nhiên, còn quá sớm để nói như thể đó là thành quả đã được xác nhận. Các con số như một thành phần nào đó cho độ bền chính xác bao nhiêu ở một nhiệt độ nhất định cần được kiểm tra từ nguyên văn nguồn. Cuốn sách này không dùng các giá trị tính năng chưa được kiểm chứng. Thay vào đó, phần này nhấn mạnh quy trình kiểm chứng mà thiết kế phải đi qua.

Kiểm chứng gồm ba bước. Trước hết xem thành phần ứng viên có ổn định trong tính toán hay không. Xác nhận cân bằng pha bằng CALPHAD và lý thuyết phiếm hàm mật độ. Tiếp theo xem hợp kim có thực sự chế tạo được hay không. Dù tốt trong tính toán, nếu không thể chế tạo thì vô ích. Cuối cùng đo độ bền và khả năng chống oxy hóa ở nhiệt độ cao. Thử nghiệm gần với môi trường thực là cửa ải cuối cùng.

Không được quên rằng sai số chảy theo cùng các bước này. Sai số nhỏ trong tính toán nguyên tử lan sang dự đoán vi cấu trúc, rồi lại phóng lớn trong dự đoán độ bền linh kiện. Nếu bề mặt bị bào mòn do oxy hóa và ranh giới dịch chuyển, dự đoán sẽ dao động hơn nữa. Độ lệch phát sinh trong từng quy trình chế tạo cũng cộng thêm vào đó. Vì vậy, một thiết kế tốt không đưa ra một giá trị duy nhất mà đưa ra cả khoảng tuổi thọ. Sau cùng, nó phải trải qua thử nghiệm chứng nhận trong điều kiện giống linh kiện thực.

Ba bước này không chỉ đi theo thứ tự một chiều. Nếu thử nghiệm cho thấy điểm yếu, ta sửa lại thành phần. Thành phần đã sửa lại được tính toán lại và chế tạo lại. Quá trình xoay vòng như vậy được lặp đi lặp lại. Vòng tuần hoàn này được gọi là thiết kế vòng kín. Phòng thí nghiệm tự hành là nỗ lực tự động vận hành vòng tuần hoàn này.

Trí tuệ nhân tạo tham gia vào từng bước của vòng tuần hoàn này. Nó chọn thành phần cần thử tiếp theo, học kết quả, rồi lại chọn. Vòng tuần hoàn vẫn chạy dù con người không thức canh suốt đêm. Nhờ vậy, con đường đi tới thành phần tốt ngắn lại. Dòng công việc này đang thay đổi tương lai của phát triển hợp kim nhiệt độ cao.

Ở phía siêu hợp kim nền nickel, thiết kế bằng trí tuệ nhân tạo cũng đang diễn ra sôi nổi. Các nghiên cứu năm 2025 đã dự đoán tuổi thọ creep của siêu hợp kim đơn tinh thể bằng học máy. Chúng đưa vào thông tin thành phần và vi cấu trúc để chọn thành phần chịu được lâu. Tỷ phần thể tích và hình dạng của các hạt gamma prime được dùng làm đặc trưng quan trọng. Những nghiên cứu như vậy cho thấy trí tuệ nhân tạo đem lại công dụng mới cho cả kim loại lâu đời.

Ở đây cũng có nghiên cứu kết hợp tính toán trường pha với trí tuệ nhân tạo. Trường pha là phương pháp tính toán về sự thay đổi của tổ chức vật liệu dưới dạng một trường trơn. Một nghiên cứu năm 2025 đã nối phương pháp tính này với học máy. Nghiên cứu dự đoán đồng thời ảnh hưởng của thành phần và tổ chức lên biến dạng creep. Những cách nối tính toán với trí tuệ nhân tạo như vậy đang ngày càng tăng.

Các trường hợp này đưa ra một bài học chung. Trí tuệ nhân tạo phát huy sức mạnh khi có đặc trưng tốt và dữ liệu đủ. Nếu đặc trưng chạm tới vật lý, dự đoán sẽ vững. Nếu dữ liệu lệch về một vùng nhất định, dự đoán sẽ dao động ở ngoài vùng đó. Vì vậy, việc đầu tiên là thu thập dữ liệu rộng và chọn đặc trưng tốt.

Trong quy trình này, vị trí của trí tuệ nhân tạo rất rõ. Trí tuệ nhân tạo là người dẫn đường thu hẹp không gian thành phần. Nếu trí tuệ nhân tạo giảm 100 ứng viên xuống còn 3, con người chỉ thử sâu 3 ứng viên đó. Khi chia vai như vậy, phát triển nhanh hơn và thất bại ít hơn. Đây là hình ảnh thực tế của trí tuệ nhân tạo thiết kế kim loại không trái với định luật vật lý.

## Tổng kết

Chương này bàn về kim loại bảo vệ những vị trí đạt nhiệt độ cao nhất trong tên lửa. Những vị trí đó là vòi phun buồng đốt và bề mặt tái nhập. Ở đây, nhiệt, áp suất và lực cùng lúc hành hạ vật liệu.

Các siêu hợp kim nền nickel đã được dùng từ lâu sẽ chạm giới hạn ở khoảng 1.150 độ. Ứng viên mới nhắm đến nhiệt độ cao hơn mức đó là hợp kim entropy cao chịu lửa. Hợp kim này được tạo bằng cách trộn nhiều kim loại có điểm nóng chảy cao với tỷ lệ gần tương đương nhau.

Các nghiên cứu ban đầu giải thích tính chất của hợp kim này bằng bốn hiệu ứng. Đó là entropy cấu hình, biến dạng mạng tinh thể, khuếch tán chậm và hiệu ứng cocktail. Một số điểm trong đó về sau đã gặp phản biện, nên khó xem là quy luật chung. Dù vậy, khung này vẫn là điểm khởi đầu để suy nghĩ xem thành phần nào có thể ổn định và giữ nguyên một pha. Tham số omega và nồng độ electron hóa trị là những chỉ báo giản dị giúp cho phán đoán đó.

Muốn xử lý không gian thành phần rộng thì cần có trí tuệ nhân tạo. Trí tuệ nhân tạo đọc vật liệu bằng các con số gọi là descriptor. Mô hình ngôn ngữ lớn trích xuất các đặc trưng này từ văn bản bài báo. Mạng nơ-ron đồ thị học vật liệu dưới dạng các điểm và đường nối. Mạng nơ-ron đồ thị phân số tinh thể là một nỗ lực mới nhằm đưa tính vô trật tự của hợp kim entropy cao vào mô hình.

Hành vi ở cấp độ nguyên tử được quan sát bằng động lực học phân tử. Một dạng trí tuệ nhân tạo gọi là thể năng tiến hóa nơ-ron làm cho phép tính này nhanh hơn. Nghiên cứu gần đây đã đạt tới quy mô hơn 1 nghìn tỷ nguyên tử. Phép tính lớn này theo dõi lệch mạng, trượt biên hạt và vết nứt ở cấp độ nguyên tử.

Ý tưởng gắn tất cả các công cụ này lại với nhau là tính nhất quán vật lý. Học thuần túy từ dữ liệu có thể đưa ra câu trả lời vi phạm vật lý. Mạng nơ-ron dựa trên vật lý (PINN) đưa các định luật vào hàm mất mát để ngăn điều đó. Trí tuệ nhân tạo là người dẫn đường giúp thu hẹp ứng viên, còn phán đoán cuối cùng do thí nghiệm đưa ra.

Phần lớn các trường hợp trong chương này vẫn là nghiên cứu giai đoạn đầu. Các con số trong nghiên cứu sơ bộ chưa phải giá trị đã được xác định chắc chắn. Trường hợp của công ty chỉ là ví dụ công nghiệp, không phải kiểm chứng học thuật. Vì vậy, một thiết kế tốt luôn phải đi qua cả xác nhận bằng tính toán và thí nghiệm. Trí tuệ nhân tạo thu hẹp con đường, con người quyết định con đường. Đó là hình thức hợp tác hiện nay trong lĩnh vực này.

Nếu rút gọn chương này thành một câu thì như sau. Trí tuệ nhân tạo làm thay đổi tốc độ thiết kế kim loại nhiệt độ cao. Trước đây, thử nghiệm một ứng viên mất nhiều thời gian. Nay trí tuệ nhân tạo sàng lọc trước vô số ứng viên. Chỉ một số ít còn lại mới được xác nhận bằng tính toán chính xác và thử nghiệm.

Những nhiệm vụ phía trước cũng đã rõ. Cần thu thập dữ liệu thí nghiệm rộng hơn và đồng đều hơn. Cần thông báo kèm độ bất định của dự đoán trí tuệ nhân tạo. Cần đưa các định luật vật lý vào học máy dày hơn. Ba điều này càng được đáp ứng, thiết kế vật liệu tên lửa càng nhanh hơn.

## Tài liệu căn cứ liên quan

Yeh, J. W., Chen, S. K., Lin, S. J., Gan, J. Y., Chiu, 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 Nb<sub>25</sub>Mo<sub>25</sub>Ta<sub>25</sub>W<sub>25</sub> and V<sub>20</sub>Nb<sub>20</sub>Mo<sub>20</sub>Ta<sub>20</sub>W<sub>20</sub> 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 Bản tiền ấn phẩm (chưa qua bình duyệt), 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 Bản tiền ấn phẩm (chưa qua bình duyệt), 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 Bản tiền ấn phẩm (chưa qua bình duyệt), 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 Bản tiền ấn phẩm (chưa qua bình duyệt), arXiv:2605.08103. <https://arxiv.org/abs/2605.08103>
- Bejjipurapu, A., Karumuri, S., Flanagan, J. C., Tucker, V., Bilonis, I., Strachan, A., Sandhage, K. H., & Titus, M. S. (2025). Active Learning Discovery of High Temperature Oxidation Resistant Refractory Complex Concentrated Alloys. arXiv Bản tiền ấn phẩm (chưa qua bình duyệt), 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., Benazzouq, 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>

## Chương 5. Đánh giá kiểm soát vi cấu trúc và tuổi thọ chịu gia nhiệt của vật liệu composite dùng cho hàng không vũ trụ (CMC) và gốm siêu chịu nhiệt (UHTC)

### Mở đầu

Bát sứ không chảy dù đặt vào lò nóng. Nhưng nếu thả ngay chiếc bát nóng vào nước lạnh, nó sẽ nứt. Vì nó chịu nhiệt tốt nhưng yếu trước thay đổi nhiệt độ đột ngột và va đập. Vật liệu bảo vệ bề mặt của tên lửa và phương tiện bay siêu vượt âm cũng phải tìm lời giải giữa hai tính chất này. Chỉ có điều nhiệt độ và lực ở đó không thể so với nhà bếp.

Mặt trong buồng đốt của động cơ tên lửa đầy khí nóng ở nhiệt độ hàng nghìn độ. Phần mũi của phương tiện bay siêu vượt âm va vào không khí và tự nóng lên. Bề mặt ở những nơi như vậy cùng lúc chịu nhiệt, oxy hóa, mài mòn và lực lớn. Oxy hóa là hiện tượng oxy gặp vật liệu và ăn mòn vật liệu.

Kim loại bền chắc nhưng mềm đi ở nhiệt độ rất cao. Nó giống như thanh kẹo mạch nha cong trong tay. Vì vậy các nhà nghiên cứu chuyển hướng sang gốm. Gốm chịu nhiệt tốt như đồ sứ, nên vẫn bền ở nhiệt độ cao.

Hai loại vật liệu được bàn trong chương này là vật liệu composite nền gốm (CMC) và gốm siêu chịu nhiệt (UHTC). Vật liệu composite nền gốm được tạo ra bằng cách đưa sợi mảnh vào trong nền gốm. Gốm siêu chịu nhiệt là loại gốm có điểm nóng chảy rất cao. Cả hai đều hướng tới mục tiêu giữ được hình dạng ở nơi nóng.

Vấn đề là khó biết trước khi nào các vật liệu này sẽ hỏng. Khó tái tạo nguyên vẹn môi trường bay thật trong phòng thí nghiệm. Chi phí cho một lần thử nghiệm cũng lớn. Vì vậy trí tuệ nhân tạo (AI) xuất hiện như một công cụ giúp giảm bớt khó khăn này.

Trí tuệ nhân tạo được dùng để đọc cấu trúc nhỏ bên trong vật liệu và dự đoán tuổi thọ. Cấu trúc nhỏ là hình dạng của hạt, sợi, lỗ rỗng và ranh giới nhìn thấy dưới kính hiển vi. Chương này lần lượt giải thích cách biến cấu trúc nhỏ đó thành con số và cách nhìn trước tương lai. Không dùng công thức, chỉ giải thích bằng lời.

Chương này có một dòng chảy lớn. Đó là biến thể giới nhỏ của vật liệu thành con số. Ảnh hiển vi, hướng tinh thể, dao động của nguyên tử đều là đối tượng. Thông tin này được chuyển thành dạng mà trí tuệ nhân tạo có thể đọc. Khi đó trí tuệ nhân tạo nhìn trước tương lai và tạo hình cấu trúc.

Trình tự của chương này như sau. Trước tiên xem vật liệu hư hại ra sao trong môi trường khắc nghiệt. Tiếp theo xem cách biến các hạt tinh thể thành đồ thị để dự đoán tuổi thọ. Sau đó xem cách thiết kế ngược cấu trúc dệt của sợi. Cuối cùng xem cách nối dòng nhiệt ở thể giới rất nhỏ với dòng nhiệt của toàn bộ linh kiện.

Cần nêu trước một nguyên tắc. Trí tuệ nhân tạo trong chương này không thay thế thử nghiệm. Chỉ số hiệu năng do trí tuệ nhân tạo đưa ra phải được kiểm tra bằng vật thật. Chỉ số chưa được kiểm chứng chỉ có giá trị tham khảo. Cuốn sách này chỉ đưa dữ liệu đã được kiểm chứng vào phần chính. Nguyên tắc này là nền tảng bảo vệ an toàn trong phát triển vật liệu.

## 5.1 Môi trường khắc nghiệt của bay siêu vượt âm và động cơ đẩy tên lửa thế hệ mới, cùng cơ chế suy thoái của kết cấu nhiệt

### Nội dung của mục này

Mục này trước hết phác họa môi trường mà vật liệu phải chịu khắc nghiệt đến mức nào. Ví dụ được lấy từ phương tiện bay siêu vượt âm và bên trong động cơ tên lửa áp suất cao. Mục này làm rõ vì sao kim loại thông thường không đủ và vì sao cần gốm.

Sau đó xem vật liệu gốm thực sự hư hại như thế nào. Lần lượt xem xét oxy hóa của composite sợi carbon, sự nóng chảy ăn mòn và biến đổi pha của gốm siêu chịu nhiệt. Nóng chảy ăn mòn là hiện tượng bề mặt bị khí nóng làm chảy hoặc bào mòn. Phải hiểu quá trình hư hại này mới hiểu được mô hình dự đoán bằng trí tuệ nhân tạo ở mục tiếp theo đang học điều gì.

### 5.1.1 Gia nhiệt khí động không cân bằng và môi trường khí cháy enthalpy cao

Gia nhiệt khí động là hiện tượng không khí làm nóng bề mặt do bay nhanh. Ma sát với bề mặt cũng góp thêm nhiệt. Nhưng ở tốc độ siêu vượt âm, phần nhiệt do không khí bị nén mạnh và nóng lên lớn hơn. Tốc độ càng cao, nhiệt này tăng càng nhanh và nung nóng bề mặt.

Siêu vượt âm là tốc độ lớn hơn năm lần tốc độ âm thanh. Mach 7 là tốc độ nhanh hơn âm thanh bảy lần. Ở tốc độ này, không khí không chảy bình thường. Sóng xung kích hình thành phía trước phương tiện bay. Sóng xung kích là bức tường áp suất mỏng do không khí bị nén đột ngột tạo ra.

Không khí đi qua sóng xung kích chậm lại và nóng lên trong khoảng khắc. Khi nhiệt độ tăng rất cao, phân tử không khí bị tách ra. Phân tử oxy và phân tử nitơ tách thành từng nguyên tử riêng lẻ. Oxy nguyên tử được tách ra như vậy có tính phản ứng cao hơn oxy thông thường, nên tấn công vật liệu mạnh hơn.

Những nơi chịu nhiệt nghiêm trọng hơn bất kỳ chỗ nào khác trên phương tiện bay là mép trước và mũi. Mép trước là cạnh của cánh gặp không khí đầu tiên. Mũi là phần đầu nhọn của phương tiện bay. Bề mặt ở đó tăng lên tới hàng nghìn độ. Lực cắt đẩy bề mặt sang bên cũng tác động chòng lên đó.

Bên trong động cơ tên lửa cũng khắc nghiệt như vậy. Khi đốt methane và oxy lỏng, khí nóng được tạo ra. Trong khí này có nước, carbon dioxide và carbon monoxide trộn lẫn. Cũng còn có oxy dư. Áp suất buồng đốt đạt tới hàng chục atm. Khí áp suất cao này liên tục va đập vào thành buồng đốt.

Môi trường như vậy được gọi là môi trường enthalpy cao. Enthalpy là lượng nhiệt năng mà khí chứa đựng. Enthalpy cao có nghĩa là khí mang rất nhiều nhiệt. Khí này đổ nhiệt vào thành. Tuy nhiên nhiệt độ thành thấp hơn nhiệt độ khí. Vì nhiệt lan vào bên trong vật liệu và quá trình làm mát cũng diễn ra.

Thời gian cũng trở thành vấn đề. Tên lửa cháy trong hàng trăm giây từ lúc phóng đến khi vào quỹ đạo. Phương tiện bay siêu vượt âm lướt lâu trong khí quyển. Điều đó có nghĩa là nó chịu nhiệt trong thời gian dài, không phải chỉ trong khoảng khắc ngắn. Vì vậy vật liệu chỉ chịu được chốc lát là chưa đủ.

Hãy thử cảm nhận mức độ nóng. Điểm đình trệ ở mép trước siêu vượt âm gần với giới hạn của vật liệu. Điểm đình trệ là vị trí dòng chảy dừng lại và nhiệt tập trung cao nhất. Tải nhiệt ở vị trí này không thể so với ngọn lửa gas trong bếp. Lực cắt đẩy bề mặt sang bên cũng cùng lúc đặt lên vị trí này.

Một vật liệu phải chịu tất cả các điều kiện này. Nhiệt, oxy hóa, lực cắt và thời gian phơi nhiễm dài chòng lên nhau. Một bài tổng quan năm 2024 về vật liệu cho bay siêu vượt âm đã xem môi trường tổng hợp này là bài toán then chốt trong lựa chọn vật liệu. Bài toán này là điểm xuất phát của nghiên cứu vật liệu gốm.

Việc trí tuệ nhân tạo cần làm ở đây rất rõ. Thử nghiệm thật đắt và hiếm. Vì vậy phải nhìn trước một phạm vi điều kiện rộng từ dữ liệu của vài lần thử nghiệm. Trí tuệ nhân tạo gởi rút ra quy luật từ ít dữ liệu. Tuy nhiên nếu không có sự trợ giúp của định luật vật lý, nó sẽ đưa ra câu trả lời sai lệch. Đây là lý do nhiều phương pháp trong chương này dùng vật lý cùng với trí tuệ nhân tạo.

### 5.1.2 Quá trình composite sợi carbon bị oxy làm hư hại (C/SiC)

Silicon carbide gia cường bằng sợi carbon (C/SiC) là vật liệu kết cấu nhiệt tiêu biểu. Sợi carbon mảnh, nhẹ và khó đứt. Silicon carbide (SiC) chịu nhiệt độ cao và oxy hóa tương đối tốt. Kết hợp hai vật liệu này tạo ra composite nhẹ và bền. Vật liệu này dưới đây được viết là C/SiC.

Có hai phương pháp chính để tạo vật liệu này. Một là thẩm pha hơi hóa học. Khí được cho thẩm vào để lấp gồm giữa các sợi. Phương pháp kia là thẩm nhựa lỏng rồi carbon hóa. Người ta đưa nhựa vào và nung để biến nó thành gốm. Cả hai phương pháp đều lấp đầy khoảng giữa các bó sợi bằng gốm.

Vấn đề bắt đầu ngay trong quá trình chế tạo. Sợi carbon và nền silicon carbide giãn nở khác nhau khi nhận nhiệt. Điều này được gọi là chênh lệch hệ số giãn nở nhiệt. Khi nguội từ trạng thái nóng xuống nhiệt độ phòng, chúng co lại với mức khác nhau. Kết quả là một mạng vết nứt mảnh xuất hiện trong nền. Các vết nứt vi mô này không thể tránh trong quá trình chế tạo.

Vết nứt trở thành đường cho oxy đi vào bên trong. Bên ngoài trông vẫn nguyên vẹn nhưng oxy luồn vào bên trong. Nó giống như nước thấm qua khe nhỏ trên giấy dán tường. Oxy đi vào tấn công sợi carbon và lớp giao diện. Lớp giao diện là ranh giới mỏng giữa sợi và nền.

Nền silicon carbide phản ứng với oxy theo hai cách. Khi nhiệt độ không quá cao và oxy dồi dào, một màng thủy tinh hình thành trên bề mặt. Màng này là lớp bảo vệ mỏng làm bằng silicon dioxide. Điều này được gọi là oxy hóa thụ động. Màng thủy tinh này chặn oxy thâm nhập và bảo vệ bên trong vật liệu.

Ngược lại, khi nhiệt độ rất cao hoặc áp suất oxy thấp, tình hình thay đổi. Màng thủy tinh bảo vệ không thể hình thành. Thay vào đó, silicon carbide bay hơi thăng thành khí. Điều này được gọi là oxy hóa chủ động. Trong oxy hóa chủ động, bề mặt tiếp tục bị bào mòn mà không có màng bảo vệ. Nguy cơ này tăng lên trong bay siêu vượt âm ở độ cao lớn.

Khi lớp giao diện biến mất, ưu điểm của composite sụp đổ. Ban đầu, khi vết nứt phát triển, sợi giữ vết nứt lại như một cây cầu. Điều này được gọi là bắc cầu bằng sợi. Khi giao diện mất đi do oxy hóa, lực giữ này biến mất. Khi đó vật liệu dễ rơi vào trạng thái vỡ đột ngột. Hư hại bóc tách, trong đó các lớp tách khỏi nhau, cũng xảy ra cùng lúc.

Hư hại này không dễ lộ ra bên ngoài. Bề mặt có thể vẫn nguyên vẹn nhưng giao diện bên trong hỏng trước. Vì vậy chỉ kiểm tra bằng mắt là chưa đủ. Phải xem xét oxy thâm nhập bên trong vật liệu và trạng thái giao diện. Mô hình tuổi thọ bằng trí tuệ nhân tạo đọc được hư hại ẩn bên trong này, chứ không chỉ nhìn bề mặt.

Một nghiên cứu công bố năm 2026 cho thấy hư hại này bằng con số. Nhóm của Wu đã dự đoán hư hại oxy hóa của C/SiC đạt 2.5 chiều bằng mô hình vật lý. Trong nghiên cứu này, độ bền kéo ở nhiệt độ phòng khoảng 264 megapascal. Khi tăng đến 1100 độ, độ bền giảm xuống khoảng 184 megapascal. Mô đun đàn hồi cũng giảm khoảng 23 phần trăm trong cùng khoảng đó. Nghiên cứu này dùng mô hình cấu trúc dựa trên vật lý, không phải mô hình trí tuệ nhân tạo.

Cũng có con đường khác để chặn oxy. Đó là phủ lớp bảo vệ lên bề mặt. Lớp này được gọi là lớp phủ chắn môi trường. Lớp phủ này ngăn oxy và hơi nước đi vào bên trong. Trong nghiên cứu siêu vượt âm, mục tiêu là lớp phủ chịu được trên 2000 độ C. Trí tuệ nhân tạo được dùng để quyết định độ dày và thành phần của lớp phủ.

Kết quả này gợi ra điều gì. Nếu chỉ nhìn tính chất trung bình của vật liệu, ta sẽ bỏ sót hư hại giao diện. Phải xem oxy đi vào đâu và bao nhiêu, theo từng vị trí. Lịch sử nhiệt độ và lịch sử tải cũng phải được xem cùng nhau. Mô hình tuổi thọ bằng trí tuệ nhân tạo ở phần sau đảm nhận đúng việc này.

Cũng cần nêu giới hạn ở đây. Mô hình của nhóm Wu là tính toán dựa trên vật lý. Nó dự đoán hư hại ngay cả khi không có trí tuệ nhân tạo. Trí tuệ nhân tạo được dùng theo hướng học dữ liệu từ các mô hình vật lý như vậy để tăng tốc. Mô hình vật lý và trí tuệ nhân tạo không phải đối thủ cạnh tranh, mà là cặp đôi bù đắp điểm thiếu của nhau.

### 5.1.3 Quá trình gốm siêu chịu nhiệt nóng chảy và bong ra (UHTC)

Gốm siêu chịu nhiệt là gốm có điểm nóng chảy rất cao. Vật liệu tiêu biểu là zirconium diboride và hafnium diboride. Zirconium diboride là hợp chất của zirconium và boron. Điểm nóng chảy của các vật liệu này vượt quá 3000 độ C. Vì vậy chúng được nghiên cứu làm ứng viên cho mép trước siêu vượt âm và linh kiện tên lửa.

Không thể yên tâm chỉ vì điểm nóng chảy cao. Vấn đề thật sự là oxy hóa. Khi zirconium diboride gặp oxy, hai loại oxide hình thành. Một là oxide cứng gọi là zirconia. Loại kia là boron oxide gọi là diboron trioxide.

Diboron trioxide có tính chất đáng chú ý ở nhiệt độ thấp. Ở khoảng 450 độ C, nó nóng chảy thành chất lỏng. Chất lỏng này lấp các khe nhỏ trên bề mặt. Nó hoạt động như thuốc mỡ bôi lên vết thương để lấp khe hở. Đây được gọi là màng tự chữa lành. Trong giai đoạn oxy hóa ban đầu, màng này bảo vệ vật liệu.

Nhưng khi nhiệt độ tăng thêm, lớp phòng vệ này sụp đổ. Khi vượt khoảng 1100 độ C, diboron trioxide bắt đầu bay hơi từng chút một. Khi đạt quanh 1500 độ C, màng này gần như biến mất. Nhiệt độ giới hạn chính xác thay đổi theo áp suất oxy, tốc độ khí và áp suất. Vì vậy khó nói bằng một giá trị cố định.

Khi thêm silicon carbide, tình hình tốt hơn. Thủy tinh lỏng do silicon carbide bị oxy hóa trộn với diboron trioxide. Hai thành phần gặp nhau và tạo ra pha thủy tinh gọi là borosilicate. Thủy tinh này phủ bề mặt đến nhiệt độ cao hơn.

Khi nhiệt độ cao hơn nữa, ngay cả thủy tinh này cũng không chịu nổi. Khi vượt khoảng 1700 độ C, độ nhớt của thủy tinh giảm. Thủy tinh bay hơi và bị dòng khí nhanh giật ra. Khi màng bảo vệ biến mất như vậy, bề mặt bị bào mòn nhanh. Cuối cùng trên bề mặt chỉ còn bộ khung zirconia đầy lỗ rỗng. Bộ khung đầy lỗ này cho oxy đi sâu vào bên trong dễ hơn.

Zirconia còn có một cái bẫy khác. Cấu trúc tinh thể của zirconia thay đổi theo nhiệt độ. Trong quá trình nguội, nó chuyển từ hệ tứ phương sang hệ đơn nghiêng. Hai tên gọi này chỉ sự khác nhau trong cách nguyên tử xếp chồng. Khi thay đổi này xảy ra, thể tích tăng khoảng 4.5 phần trăm.

Khi thể tích tăng đột ngột, lực lớn xuất hiện trong màng oxide. Nó giống như nước đóng băng vào mùa đông và làm vỡ chai. Lực này tạo ra vết nứt trong màng oxide. Khi các vết nứt nối với nhau, màng oxide rơi ra từng mảnh. Điều này được gọi là bong tách. Ở vị trí bong ra, bề mặt mới lộ ra và quá trình oxy hóa bắt đầu lại.

Boride entropy cao là một cách tiếp cận mới đối với vấn đề này. Gốm entropy cao là loại gốm trộn nhiều nguyên tố kim loại với lượng gần tương đương nhau. Có khi người ta đưa năm kim loại vào cùng một tinh thể. Vì dải thành phần rộng, nó phù hợp với việc tìm kiếm bằng trí tuệ nhân tạo. Máy tính có thể lọc trước tổ hợp nào chống oxy hóa tốt.

Một bài tổng quan năm 2026 về gốm siêu nhiệt độ cao cho thấy rõ xu hướng này. Bài tổng quan này bàn cùng lúc về thí nghiệm, mô phỏng đa thang đo và thiết kế dựa trên dữ liệu. Đa thang đo là cách nối nhiều

tầng từ kích thước nguyên tử đến kích thước linh kiện. Thiết kế dựa trên dữ liệu là cách để trí tuệ nhân tạo chọn thành phần và vi cấu trúc.

Trí tuệ nhân tạo cũng được dùng để dự đoán độ dày màng oxy hóa. Một nhóm nghiên cứu đã dự đoán hư hại oxy hóa của boride siêu nhiệt độ cao bằng học máy. Họ dự báo độ dày màng oxy hóa bằng phương pháp gọi là rừng ngẫu nhiên. Rừng ngẫu nhiên là phương pháp gom nhiều quy tắc quyết định để đưa ra câu trả lời. Tuy nhiên, hành vi thực tế của màng oxy hóa phải được xác nhận bằng thử nghiệm trong hầm gió và thử nghiệm arc-jet. Arc-jet là thiết bị phun khí siêu nóng để mô phỏng môi trường bay.

thể năng liên nguyên tử học máy (MLIP) cũng được dùng để dự đoán điểm nóng chảy. Đây là phương pháp để trí tuệ nhân tạo tính nhanh lực giữa các nguyên tử. Năm 2025, một nhóm nghiên cứu đã dùng MLIP để dự báo điểm nóng chảy của gốm entropy cao. Họ lọc trước bằng máy tính các tổ hợp có điểm nóng chảy cao. Phương pháp này thu hẹp danh sách ứng viên trước khi làm thí nghiệm. Nhờ vậy, số lần tổng hợp và thử nghiệm tốn kém giảm đi. Khi tính toán lọc trước như vậy, toàn bộ quá trình tìm kiếm vật liệu sẽ nhanh hơn.

Dù vậy vẫn còn một câu hỏi. Tổ hợp mà máy tính chọn có thật sự chống oxy hóa tốt không. Tính toán giả định các điều kiện lý tưởng. Trên bề mặt thực tế có tạp chất và khuyết tật. Dòng chảy và áp suất của khí cũng khác với tính toán. Vì vậy, phán đoán cuối cùng do thử nghiệm trên vật thật đưa ra. Trí tuệ nhân tạo chỉ dừng ở việc chọn thông minh các ứng viên cần thử nghiệm.

## 5.2 Cách biến các hạt tinh thể thành đồ thị để dự báo tuổi thọ

### Nội dung của mục này

Mục này bàn về cách biến bên trong vật liệu thành dạng mà trí tuệ nhân tạo có thể đọc được. Công cụ đó là mạng nơ-ron đồ thị không-thời gian. Đồ thị là hình vẽ thể hiện quan hệ bằng điểm và đường. Sơ đồ tuyến tàu điện ngầm là một ví dụ tốt. Nhà ga là điểm, còn đoạn tuyến là đường.

Ta biến các hạt tinh thể bên trong vật liệu thành điểm, và biến ranh giới giữa các hạt thành đường. Sau đó thêm dòng thời gian vào đó. Vì khi vật liệu chịu nhiệt và lực lặp lại, hư hại sẽ tích tụ. Mục này giải thích cách dựng đồ thị đó, hư hại lớn lên ra sao, và đưa các định luật vật lý vào học máy như thế nào. Vì mô hình nhìn cả thời gian và không gian, nó được gọi là mô hình không-thời gian.

### 5.2.1 Biến vi cấu trúc thành đồ thị

Hạt tinh thể là những hạt tinh thể nhỏ bên trong vật liệu. Kim loại hoặc gốm không phải là một tinh thể lớn duy nhất. Chúng được tạo thành từ vô số hạt nhỏ tập hợp lại. Hướng nguyên tử trong từng hạt hơi khác nhau. Hướng này được gọi là định hướng tinh thể.

Thiết bị tiêu biểu để đọc định hướng tinh thể là nhiễu xạ tán xạ ngược electron (EBSD). Trong kính hiển vi electron, người ta bắn electron để đọc hướng của từng hạt. Thiết bị này quét bề mặt và ghi lại hướng của từng hạt. Kết quả hiện ra dưới dạng bản đồ định hướng nhiều màu.

Ranh giới hạt là ranh giới nơi hai hạt gặp nhau. Ranh giới này dễ trở thành mắt xích yếu của vật liệu. Oxy thấm vào theo con đường này. Vết nứt cũng lớn lên dọc theo ranh giới này. Vì vậy, biết tính chất của ranh giới là cốt lõi của dự đoán tuổi thọ.

Cách biến thành đồ thị xuất phát từ một ý tưởng dễ hiểu. Ta đặt một hạt làm một điểm, tức một nút. Nếu hai hạt dính nhau, ta nối giữa chúng bằng một đường, tức một cạnh. Làm như vậy, ảnh hiển vi phức tạp trở thành một sơ đồ quan hệ gọn gàng. Trí tuệ nhân tạo đọc sơ đồ quan hệ này tốt hơn ảnh.

Nút chứa thông tin của hạt. Ta đưa vào kích thước, hình dạng và hướng của hạt. Cũng đưa vào nhiệt độ và ứng suất tại vị trí đó. Ứng suất là độ lớn của lực nén hoặc kéo một vật thể. Như vậy, một điểm tóm tắt trạng thái của hạt đó.

Cạnh chứa thông tin của ranh giới. Ta đưa vào chênh lệch hướng giữa hai hạt. Cũng đưa vào diện tích và hướng của ranh giới. Mức độ hư hại do oxy hóa cũng được đưa vào. Ranh giới có chênh lệch hướng càng lớn thường có xu hướng để oxy xâm nhập dễ hơn.

Cách này đã được thử nghiệm trên vật liệu khác. Năm 2023, một nhóm nghiên cứu đã áp dụng cách đồ thị này cho thép. Họ phân loại những nơi có hư hại mỗi nặng trong từng hạt của thép. Mỗi là hư hại sinh ra khi vật liệu chịu lực lặp lại. Họ biến bản đồ EBSD thành đồ thị rồi cho mô hình học. Trong nghiên cứu này, cách đồ thị cho thấy có ích trong việc nhận ra vị trí hư hại theo từng hạt. Tuy nhiên, nghiên cứu này xử lý phân loại mỗi của thép. Nó không phải là nghiên cứu dự đoán tuổi thọ của gôm, đứt gãy do creep, hay biến đổi theo thời gian.

Đưa cách đồ thị này sang dự đoán tuổi thọ của gôm là hướng mà cuốn sách này đề xuất. Việc kiểm chứng trên gôm vẫn chưa được thực hiện. Phần giải thích dưới đây không phải là sự thật đã được kiểm chứng, mà là bức tranh về con đường nghiên cứu có thể đi tới.

Cũng xuất hiện những phương pháp mới để xử lý thông tin ranh giới. Năm 2026, một nhóm nghiên cứu đã dùng học đồ thị để lấp các vùng trống trong bản đồ EBSD. Đây là kỹ thuật điền vào các vùng bị thiếu trong phép đo sao cho phù hợp với quy tắc tính thể học. Phương pháp này đưa quy tắc tính thể vào trí tuệ nhân tạo để không điền những hướng vô lý. Vì dữ liệu vật liệu không phải lúc nào cũng đầy đủ, các kỹ thuật bổ sung như vậy có ích.

Cách đồ thị có điểm hơn ảnh. Ảnh xử lý từng pixel một. Vì vậy, nó không tự biết hạt bắt đầu và kết thúc ở đâu. Đồ thị gom một hạt thành một đơn vị. Cách này gần với cách con người nhìn vật liệu. Vì vậy, nó dễ học quan hệ giữa các hạt.

Đồ thị cũng dễ xử lý dù kích thước khác nhau. Ảnh phải có kích thước cố định. Đồ thị xử lý theo cùng một cách dù số hạt tăng hay giảm. Vì mỗi vật liệu có số hạt khác nhau, tính chất này có ích. Ta có thể thử chuyển quy tắc học được từ mẫu nhỏ sang linh kiện lớn. Tuy nhiên, khi kích thước lớn hơn, phân bố khuyết tật và gradient nhiệt cũng thay đổi. Vì vậy, cần có bước kiểm chứng lại trên linh kiện lớn.

Đồ thị được tạo như vậy cũng có thể chuyển sang dùng cho gôm. Vì gôm cũng được tạo thành từ hạt tinh thể và ranh giới. Oxy hóa và creep của gôm siêu nhiệt độ cao cũng bắt đầu từ ranh giới. Creep là hiện tượng vật liệu từ từ giãn ra khi chịu lực trong thời gian dài. Vì vậy, cách đồ thị này trở thành nền tảng cho dự đoán tuổi thọ gôm.

### 5.2.2 Học đồ thị nhìn đồng thời hàng xóm và thời gian

Bây giờ ta thêm dòng thời gian vào đồ thị. Hư hại không kết thúc trong một khoảnh khắc. Nhiệt và lực lặp lại, rồi tích tụ từng chút một. Vì vậy, ngay cả cùng một vật liệu, đồ thị cũng thay đổi theo thời gian. Theo dõi sự thay đổi này là việc của mô hình không-thời gian.

Trước hết nhìn phía không gian. Trạng thái của một hạt chịu ảnh hưởng của các hạt lân cận. Nếu hạt bên cạnh chịu lực lớn, lực cũng truyền sang hạt của mình. Phép tính thu thập thông tin hàng xóm này là tích chập đồ thị. Mỗi nút nhận thông tin từ các nút lân cận và cập nhật trạng thái của chính nó.

Không phải mọi hàng xóm đều được đối xử như nhau. Có ranh giới đóng góp lớn vào hư hại, có ranh giới đóng góp ít hơn. Vì vậy, mô hình đặt trọng số khác nhau cho từng hàng xóm. Cách xác định trọng số này là cơ chế chú ý. Đó là cách dành nhiều chú ý hơn cho hàng xóm quan trọng.

Điểm cốt lõi là đưa tri thức vật lý vào cơ chế chú ý. Nếu chênh lệch hướng của ranh giới lớn, năng lượng ranh giới cao. Nếu năng lượng ranh giới cao, oxy đi vào dễ hơn. Các quy tắc vật lý như vậy được phản ánh trong phép tính trọng số. Khi đó, dù dữ liệu ít, mô hình vẫn tránh được kết luận vô lý.

Phía thời gian cũng được xử lý tương tự. Trạng thái trước tạo ra trạng thái sau. Hư hại của hôm nay được cộng lên trên hư hại đã tích tụ đến hôm qua. Mô hình học chuỗi thời gian này. Để làm việc đó, nó dùng cấu trúc mạng nơ-ron xử lý thứ tự thời gian.

Hư hại quan trọng được bàn trong mục này có hướng. Trong composite gốm, trước hết xuất hiện các vết nứt nhỏ trong nền. Dọc theo các vết nứt này, mặt phân cách giữa sợi và nền bị bong ra. Oxy đi vào chỗ đó và lớp oxy hóa lớn lên. Hư hại thường lan theo hướng của lực và khí.

Hư hại có hướng này ảnh hưởng lớn đến tuổi thọ vật liệu. Nếu vết nứt và bong tách mặt phân cách nối tiếp nhau theo một hướng, hướng đó yếu đi. Nếu các lỗ rỗng nối với nhau, đường đi của oxy rộng ra. Pha thủy tinh kẹt ở ranh giới cũng có thể di chuyển và truyền hư hại. Những thay đổi này làm đổi tốc độ creep và khiến đứt gãy đến sớm hơn. Đứt gãy là khoảnh khắc vật liệu không chịu nổi nữa và bị đứt.

Trượt ranh giới hạt cũng là hư hại quan trọng trong mục này. Ở nhiệt độ cao, các hạt trượt lên nhau dọc theo ranh giới. Nó giống cảnh những viên gạch ướt bị đẩy trượt khỏi nhau. Sự trượt này làm tăng biến dạng creep. Hướng và tính chất của ranh giới quyết định độ lớn của trượt. Vì vậy, cạnh chứa thông tin ranh giới là quan trọng.

Khí nóng thường chảy theo một hướng và đẩy bề mặt. Vì vậy, hư hại của gốm cũng lớn lên theo hướng. Mô hình đồ thị không-thời gian cố đưa cả tính có hướng này vào học máy. Khi đó có thể dự báo hư hại sẽ lớn lên theo hướng nào. Dự đoán có hướng như vậy vẫn còn thiếu dữ liệu để kiểm chứng.

Có hai cách xử lý thời gian. Một là đọc thứ tự thời gian lần lượt. Cách này giống như đọc câu từ đầu đến cuối. Cách kia là xử lý trực tiếp tốc độ thay đổi. Nó học xem hư hại lớn lên nhanh đến mức nào. Cả hai cách đều chứa quá trình hư hại tích tụ.

Cơ chế chú ý là công cụ trung tâm của trí tuệ nhân tạo hiện nay. Trí tuệ nhân tạo dịch văn bản cũng dùng cách này. Nguyên lý là tập trung hơn vào những từ quan trọng trong câu. Trong đồ thị vật liệu, nó tập trung vào các hàng xóm quan trọng đối với hư hại. Cùng một nguyên lý chú ý được dùng cả trong dịch ngôn ngữ và phân tích vật liệu.

Cách đồ thị không-thời gian này vẫn là kỹ thuật đang phát triển. Chưa có nhiều kiểm chứng quy mô lớn trên composite gốm. Muốn dự đoán tốt thì cần nhiều dữ liệu tốt. Bản thân việc thu thập dữ liệu thử nghiệm siêu nhiệt độ cao đã khó. Vì vậy, cần sự trợ giúp của các định luật vật lý sẽ được bàn ở tiểu mục tiếp theo.

### 5.2.3 Đứt gãy do creep xét theo hướng và học máy đưa quy tắc vật lý vào

Creep là hiện tượng vật liệu từ từ giãn ra khi chịu lực trong thời gian dài. Nó giống như dây phơi đồ treo lâu ngày dần chùng xuống. Ở nhiệt độ cao, hiện tượng này diễn ra nhanh hơn. Linh kiện tên lửa chịu lực trong thời gian dài khi đang nóng. Vì vậy, creep trở thành biến số lớn quyết định tuổi thọ linh kiện tên lửa.

Hư hại creep tích tụ khác nhau theo từng hướng. Điều này được gọi là tính dị hướng. Tính dị hướng là đặc tính mà tính chất khác nhau theo hướng. Nó giống như gỗ dễ tách theo thớ. Vật liệu chịu lực đồng thời từ nhiều hướng. Điều này được gọi là tải đa trục.

Các nhà nghiên cứu tóm tắt hư hại này bằng một biến. Biến đó được gọi là biến hư hại. Nếu biến hư hại bằng 0, vật liệu còn nguyên vẹn. Nếu bằng 1, vị trí đó đã đứt. Hư hại càng lớn, phần còn lại càng phải

chịu lực nhiều hơn. Khi đó hư hại lớn nhanh hơn. Vì vòng phản hồi này, hư hại tăng vượt kiểm soát và kéo đứt gãy đến sớm hơn.

Khi nhiệt độ tăng, creep nhanh hơn. Đó là vì nguyên tử chuyển động tích cực hơn. Sự phụ thuộc vào nhiệt độ này được đưa vào bằng hàm mũ. Chỉ cần nhiệt độ tăng một chút, tốc độ hư hại cũng nhảy vọt. Vì vậy, hạ nhiệt độ cao nhất trên bề mặt là việc quan trọng. Lớp phủ ở mức trước và dòng nhiệt ở mức này gặp nhau tại đây.

Oxy hóa làm hư hại này nhanh hơn. Màng oxy hóa càng dày, ranh giới càng yếu đi. Vì vậy, độ dày oxy hóa cũng được đưa vào quy luật hư hại. Điều đó có nghĩa là không xem creep và oxy hóa riêng rẽ, mà xử lý chúng cùng nhau. Vì trong môi trường tên lửa, hai hiện tượng này luôn xảy ra cùng lúc.

Ở đây có một quy tắc vật lý quan trọng. Hư hại chỉ tăng theo một hướng. Vết nứt đã hình thành không tự lành lại. Dù thời gian trôi qua, biến hư hại không giảm. Quy tắc này được gọi là tính không thuận nghịch của hư hại. Nó cùng mạch với định luật thứ hai của nhiệt động lực học.

Nếu trí tuệ nhân tạo vi phạm quy tắc này, dự đoán sẽ sai. Nếu nó đưa ra câu trả lời rằng hư hại giảm, điều đó trái với thực tế. Vì vậy, ta đưa hình phạt vào quá trình học. Khi xuất hiện câu trả lời vi phạm quy tắc, hàm mất mát bị tính lớn. Cách đưa hình phạt này vào hàm mất mát chính là học máy dựa trên vật lý.

Trong học máy dựa trên vật lý, các quy tắc khác cũng được đưa vào cùng. Cân bằng lực là một trong số đó. Tổng lực đẩy và kéo tại một vị trí phải khớp. Quy tắc rằng khối lượng không biến mất và cũng không tự sinh ra cũng được đưa vào. Các quy tắc như vậy được cộng vào hàm mất mát. Khi đó, mô hình tránh được câu trả lời trái vật lý.

Tiêu chuẩn đứt gãy cũng xét cả hướng. Vật liệu yếu trước lực kéo. Vì vậy, hướng có lực kéo lớn nhất được xem xét riêng. Áp suất làm phòng từ bên trong cũng được xét cùng. Lực theo nhiều hướng được gộp thành một mức độ nguy hiểm. Mức độ nguy hiểm này quyết định tốc độ hư hại lớn lên.

Gốc rễ của cách làm này là mạng nơ-ron dựa trên vật lý. Năm 2019, nhóm nghiên cứu của Raissi đã hệ thống hóa khái niệm này. Họ đưa trực tiếp các phương trình vật lý vào quá trình học của mạng nơ-ron. Ý tưởng là dù dữ liệu ít, các quy luật vật lý sẽ lấp vào chỗ trống. Ý tưởng đó được nối thẳng sang dự đoán tuổi thọ gốm.

Giá trị thật của quy luật vật lý lộ rõ khi mở rộng phạm vi dự đoán. Dữ liệu thử nghiệm thường tập trung ở vùng nhiệt độ thấp. Điều ta thật sự muốn biết là hành vi gần 2000 độ C. Dữ liệu ở nhiệt độ cao như vậy hầu như không có. Nếu có quy luật vật lý, ta vẫn có thể dự báo hợp lý cả những vùng không có dữ liệu.

Tuy nhiên cách làm này cũng có giới hạn. Nếu đưa sai quy luật vật lý vào mô hình, dự đoán có thể tệ hơn. Việc đặt trọng số giữa các quy luật cũng khó. Không thể bỏ qua bước kiểm chứng bằng dữ liệu phá hủy thực tế. Vì vậy, dự đoán bằng trí tuệ nhân tạo không thay thế thử nghiệm, mà được dùng như công cụ để giảm số lần thử nghiệm.

### 5.3 Mô hình sinh thiết kế cấu trúc dệt sợi và lớp phủ

Phần trước dự báo tuổi thọ của vật liệu đã được chế tạo. Phần này đổi hướng. Ta xác định trước tính năng mong muốn, rồi truy ngược để tìm cấu trúc tạo ra tính năng đó. Việc này gọi là thiết kế ngược. Nếu ví với nấu ăn, đó là xác định hương vị mong muốn rồi tìm cách phối nguyên liệu.

Đối tượng của phần này là cấu trúc sợi dệt ba chiều và lớp phủ giao diện. Vấn đề là dệt sợi theo góc nào và phủ dày bao nhiêu. Để làm việc đó, ta dùng trí tuệ nhân tạo sinh gọi là mô hình khuếch tán. Mô hình khuếch tán là phương pháp tạo ra cấu trúc rõ nét từ nhiễu mờ. Ta còn thêm vào đó tính chất để kết quả không dao động khi hướng thay đổi.

### 5.3.1 Cấu trúc sợi dệt không gian và học không dao động theo hướng

Vật liệu composite dệt ba chiều là vật liệu dệt các bó sợi theo không gian. Nó giống hình ảnh một chiếc giỏ được đan dày theo ba hướng. Composite thông thường được tạo bằng cách xếp các lớp mỏng. Khi đó vùng giữa các lớp yếu và dễ tách ra. Hiện tượng này gọi là tách lớp.

Dệt ba chiều loại bỏ điểm yếu này. Sợi đan xen theo nhiều hướng nên không có ranh giới lớp. Vì vậy hư hỏng do lớp bị tách ít xảy ra. Điều này có lợi ở những nơi chịu sốc nhiệt lớn, như mép trước của phương tiện siêu vượt âm. Sốc nhiệt là hư hỏng sinh ra khi vật liệu nóng lên hoặc nguội đi đột ngột.

Có nhiều kiểu dệt. Có thể dệt sợi theo bốn hướng, hoặc theo nhiều hướng hơn nữa. Càng nhiều hướng, vật liệu càng chịu đều nhiều loại tải. Đổi lại, việc dệt trở nên khó hơn. Vì vậy phải chọn kiểu dệt phù hợp với công dụng của bộ phận. Việc chọn kiểu dệt này cũng là đối tượng của thiết kế ngược.

Vấn đề là cấu trúc dệt không phải lúc nào cũng hoàn hảo. Trong chế tạo thực tế, sợi bị xoắn. Mặt cắt cũng có thể bị ép méo. Góc dệt cũng dao động từng chút một. Độ dày lớp phủ cũng không đều giữa các vị trí. Những dao động này làm lực tập trung vào một số chỗ.

Dao động này được xử lý bằng xác suất. Lý do là mọi sản phẩm không giống hệt nhau mà khác nhau từng chút một. Vì vậy ta không xử lý một cấu trúc đáp án duy nhất, mà xử lý phân bố của các cấu trúc. Mô hình khuếch tán phù hợp để xử lý phân bố này. Mô hình khuếch tán có thể tạo ra nhiều cấu trúc có vẻ hợp lý.

Ở đây, tính đẳng biến rất quan trọng. Đẳng biến là tính chất trong đó khi xoay vật thể, kết quả cũng xoay theo cùng cách. Nếu xoay bộ phận sang trái, lực dự đoán cũng phải xoay sang trái. Nếu không có tính chất này, cùng một cấu trúc sẽ bị đánh giá khác nhau tùy theo hướng đặt, trái với vật lý.

Trong không gian ba chiều, vật thể có thể xoay và cũng có thể tịnh tiến. Có một hệ toán học xử lý đồng thời hai chuyển động này. Hệ đó gọi là nhóm Euclid đặc biệt. Tên gọi khó, nhưng ý nghĩa thì rõ. Đó là tập hợp các quy tắc xử lý chung phép quay và phép tịnh tiến. Trí tuệ nhân tạo được thiết kế để luôn tuân thủ các quy tắc quay và tịnh tiến này.

Các mô hình khuếch tán đẳng biến như vậy đã phát triển trước trong thiết kế phân tử. Năm 2025, nhóm nghiên cứu của Chen đã đưa ra một mô hình khuếch tán đẳng biến kết hợp các đồ thị ở nhiều thang đo. Họ dùng mô hình này để thiết kế phân tử ba chiều. Phân tử cũng là đối tượng xoay và tịnh tiến trong không gian ba chiều. Vì vậy cùng một hệ toán học có thể được chuyển sang cấu trúc sợi. Tuy nhiên nghiên cứu đó xử lý phân tử, không phải cấu trúc dệt sợi gồm. Bó sợi, ràng buộc dệt, lỗ rỗng và khả năng chế tạo vẫn chưa được kiểm chứng bằng mô hình này. Chuyển nó sang dệt gồm là hướng đi mà cuốn sách này phác ra.

Hãy xem ví dụ để hiểu vì sao tính đẳng biến quan trọng. Ta đặt cùng một con ốc vít trên bàn theo nhiều hướng khác nhau. Độ bền của con ốc vít không phụ thuộc vào hướng đặt. Trí tuệ nhân tạo cũng phải giữ đúng sự thật này. Nếu không có tính đẳng biến, nó sẽ đánh giá cùng một con ốc vít khác nhau theo từng hướng. Khi đó phải thu thập dữ liệu riêng cho từng hướng, làm lãng phí lớn trong học máy.

Nếu giữ được tính đẳng biến, ta có thể tiết kiệm dữ liệu. Chỉ với dữ liệu của một hướng, mô hình xử lý được mọi hướng. Trong tình huống dữ liệu vật liệu khan hiếm, sự tiết kiệm này rất lớn. Vì vậy nó phù hợp với những lĩnh vực ít dữ liệu như gốm siêu cao nhiệt. Có thể xem đây là cách khắc quy luật vật lý vào cấu trúc mô hình.

Có thể nói ba yếu tố gặp nhau ở đây. Cấu trúc ba chiều, dao động xác suất và tính chất không dao động theo hướng. Ba yếu tố này được đưa vào một mô hình. Chỉ khi đó mô hình mới phản ánh nguyên vẹn sự đa dạng của chế tạo thực tế. Mục tiếp theo xem mô hình này tạo cấu trúc như thế nào.

### 5.3.2 Mô hình khuếch tán khôi phục cấu trúc từ nhiều

Ý tưởng cơ bản của mô hình khuếch tán chia thành hai bước. Bước trước là quá trình làm mờ cấu trúc thành nhiều. Bước sau là quá trình khôi phục cấu trúc từ nhiều. Trí tuệ nhân tạo học quá trình khôi phục này. Có thể hiểu như quay ngược cảnh một giọt mực loang trong nước.

Ta xem bước trước. Nhiều được đưa dần vào cấu trúc sợi rõ nét. Thời gian càng trôi qua, cấu trúc càng mờ. Cuối cùng chỉ còn lại nhiều hoàn toàn. Quá trình này diễn ra từ từ theo một quy tắc. Quy tắc đó gọi là lịch nhiều.

Bước sau đi ngược lại quá trình này. Nó bắt đầu từ nhiều và dần khôi phục cấu trúc. Ở mỗi bước, nó loại bỏ một phần nhiều. Khi đó có một giá trị cho biết phải loại bỏ theo hướng nào. Giá trị này gọi là điểm số. Điểm số chỉ hướng mà cấu trúc trở nên hợp lý hơn.

Trí tuệ nhân tạo học chính điểm số này. Nó nhìn cấu trúc mờ và đoán hướng đi về phía rõ nét. Khi học xong, nó có thể tạo cấu trúc ngay cả từ nhiều thuần túy. Cách làm này gọi là khuếch tán dựa trên điểm số. Các trí tuệ nhân tạo tạo ảnh ngày nay cũng vẽ tranh theo cùng nguyên lý.

Cách làm này phù hợp với thiết kế vật liệu vì một lý do. Cấu trúc vật liệu không chỉ có một đáp án. Có nhiều cấu trúc tạo ra cùng một tính năng. Mô hình khuếch tán tạo ra nhiều ứng viên như vậy. Sau đó con người chọn trong số đó cấu trúc dễ chế tạo. Giá trị của nó nằm ở chỗ không đưa ra một đáp án, mà đưa ra nhiều lựa chọn.

Việc gắn điều kiện vào đây là cốt lõi của thiết kế ngược. Nếu chỉ tạo ra cấu trúc bất kỳ thì không có ích. Phải tạo ra cấu trúc có tính năng mong muốn. Vì vậy ta đưa tính năng mục tiêu vào làm điều kiện. Mục tiêu gồm độ dẫn nhiệt, độ bền kéo, độ dai phá hủy, tốc độ ăn mòn xói mòn và các giá trị khác.

Độ dai phá hủy là khả năng chống lại sự phát triển của vết nứt. Tốc độ ăn mòn xói mòn là tốc độ bề mặt bị bào mòn. Ta đặt các giá trị này ở mức mong muốn. Khi đó mô hình chọn và tạo cấu trúc phù hợp với điều kiện đó. Cũng có thể điều chỉnh mức phản ánh mục tiêu mạnh hay yếu.

Kỹ thuật phản ánh điều kiện này đã được dùng trong nhiều loại vật liệu. Năm 2025, một nhóm nghiên cứu đã đưa ra mô hình khuếch tán ba chiều có đưa ràng buộc vật lý vào. Họ thiết kế composite polymer gia cường sợi theo tính năng mục tiêu. Họ xác định đường cong ứng suất-biến dạng mong muốn rồi truy ngược để tìm cấu trúc. Phương pháp này buộc mô hình tuân thủ trực tiếp quy luật vật lý trong quá trình sinh.

Nếu đưa điều kiện quá mạnh, một vấn đề khác xuất hiện. Nếu chỉ chạy theo tính năng mục tiêu, mô hình có thể tạo ra cấu trúc không thể chế tạo. Trên giấy thì đẹp, nhưng trong thực tế không thể dệt. Vì vậy phải để mô hình tìm đáp án trong phạm vi có thể chế tạo. Đó là cách đưa cả quy tắc chế tạo vào điều kiện. Việc giữ cân bằng giữa tính năng và khả năng chế tạo là điểm khó của thiết kế ngược.

Muốn dùng cách này cho composite gốm, còn có trở ngại phải vượt qua. Phải thu thập nhiều dữ liệu cấu trúc ba chiều của gốm. Dữ liệu này thu được bằng chụp cắt lớp tia X. Chụp cắt lớp tia X có nguyên lý giống chụp CT trong bệnh viện. Ta có thể nhìn bên trong theo không gian ba chiều mà không cần cắt bộ phận ra. Cấu trúc thu được như vậy trở thành nền dữ liệu của mô hình.

Việc thu thập dữ liệu tốn thời gian. Chụp một mẫu thử mất lâu. Việc tách dữ liệu đã chụp thành hạt và sợi cũng cần nhiều công sức. Vì vậy dữ liệu luôn thiếu. Tính đẳng biến đã nói ở trên giúp giảm phần thiếu hụt này. Lý do là nó giúp mô hình học được nhiều trường hợp hơn từ ít dữ liệu hơn.

### 5.3.3 Xác định đồng thời góc sợi và độ dày lớp phủ

Mô hình thiết kế ngược tạo ra hai quyết định. Một là góc dệt sợi. Hai là độ dày của lớp phủ giao diện. Hai yếu tố này chi phối lớn độ bền và khả năng chịu nhiệt của vật liệu.

Trước hết xem góc sợi. Lực thường tác dụng lớn theo một hướng. Vòi phun tên lửa chịu lực đẩy từ bên trong. Mép trước siêu vượt âm chịu lực cắt theo hướng dòng chảy. Nếu sợi nằm theo hướng của lực, vật liệu chịu lực tốt. Vì vậy mô hình xác định góc dựa trên lực.

Độ dày lớp phủ cũng quan trọng không kém. Ở đây phải phân biệt hai loại màng. Một là lớp giao diện mỏng đặt giữa sợi và nền. Lớp giao diện này được làm từ boron nitride hoặc carbon nhiệt phân. Màng này ngăn vết nứt cắt đứt sợi ngay lập tức. Vết nứt trượt theo màng và phân tán lực. Nếu lớp giao diện quá mỏng, hiệu ứng này yếu; nếu quá dày, vấn đề khác xuất hiện.

Loại còn lại là lớp phủ chắn môi trường phủ bên ngoài vật liệu. Lớp phủ này ngăn oxy và hơi nước đi vào bên trong. Disilicate đất hiếm là ứng viên cho lớp phủ bên ngoài này. Vật liệu này chặn oxy tốt ở nhiệt độ cao. Lớp giao diện và lớp phủ bên ngoài khác nhau về vị trí và vai trò. Đôi khi người ta xếp nhiều lớp màng chồng lên nhau và giao cho mỗi lớp một vai trò khác nhau.

Mô hình thiết kế ngược xử lý đồng thời hai quyết định này. Nó tìm cách tăng độ bền đồng thời giảm ứng suất nhiệt. Hai mục tiêu này thường va chạm nhau. Được một bên thì mất một ít bên kia. Vấn đề như vậy gọi là tối ưu hóa đa mục tiêu. Mô hình tìm điểm cân bằng tốt giữa độ bền và ứng suất nhiệt.

Việc lực dồn vào một chỗ gọi là tập trung ứng suất. Nó giống như chiếc áo bắt đầu rách từ một sợi chỉ. Vết nứt xuất hiện trước ở nơi có tập trung ứng suất lớn. Vì vậy thiết kế tìm cách chia lực cho đều. Đây là lý do phải chỉnh góc sợi và độ dày lớp phủ.

Khi tìm cân bằng, mô hình cũng xem xét độ trơn mượt. Nếu sợi đột ngột gấp hướng, lực sẽ dồn vào chỗ đó. Độ dày lớp phủ thay đổi gấp cũng vậy. Vì vậy mô hình dẫn thiết kế sao cho góc và độ dày nối tiếp nhau êm hơn. Cách làm là phạt những thiết kế có góc hoặc độ dày thay đổi đột ngột.

Thiết kế ngược như vậy hiện vẫn gần với giai đoạn chứng minh khái niệm. Cần thận trọng với những tuyên bố rằng một giá trị tính năng cụ thể được cải thiện bao nhiêu. Phải kiểm tra xem cấu trúc do mô hình tạo ra có thật sự đạt tính năng đó hay không. Việc kiểm tra lại được thực hiện bằng chế tạo và thử nghiệm. Trí tuệ nhân tạo đảm nhiệm vai trò thu hẹp danh sách ứng viên.

Năm 2026 cũng xuất hiện nghiên cứu sinh cấu trúc bằng điều kiện ngôn ngữ. Một nhóm nghiên cứu đã tạo vi cấu trúc từ câu mô tả tính chất mục tiêu. Đây vẫn là nghiên cứu ban đầu, chưa qua bình duyệt. Nhưng hướng đi thì rõ. Con người nói ra tính năng mong muốn, trí tuệ nhân tạo tạo ra cấu trúc. Dòng chảy này có thể nối sang thiết kế composite gồm.

## 5.4 Nối thế giới cực nhỏ với dòng nhiệt của toàn bộ bộ phận

Phần này bàn về dòng nhiệt bên trong vật liệu. Nhiệt là chìa khóa quyết định tuổi thọ của vật liệu. Phải biết nhiệt chảy đi đâu và nhanh đến mức nào thì mới dự đoán được nhiệt độ. Phải biết nhiệt độ thì mới dự đoán được tốc độ oxy hóa và creep tiếp theo.

Vấn đề là dòng nhiệt được quyết định trong thế giới cực nhỏ. Dao động của nguyên tử mang nhiệt đi. Nhưng điều ta muốn biết là nhiệt độ của toàn bộ bộ phận. Phải nối thế giới cực nhỏ với thế giới rất lớn. Sự nối này gọi là khớp đa thang đo. Phần này giải thích cách thực hiện sự nối đó bằng mạng nơ-ron toán tử sâu.

### 5.4.1 Tính độ dẫn nhiệt từ dao động nguyên tử (động lực học phân tử Green-Kubo)

Trong gốm, nhiệt được truyền bằng dao động của nguyên tử. Hạt của dao động này gọi là phonon. Có thể hiểu phonon là hạt của âm thanh. Dao động lan dọc theo mạng tinh thể và mang nhiệt đi. Dao động này lan tốt đến mức nào sẽ quyết định độ dẫn nhiệt.

Ở nhiệt độ cao, các dao động này va chạm với nhau. Khi các phonon va chạm, dòng nhiệt bị nhiễu loạn. Ranh giới hạt, khuyết tật và nguyên tử pha trộn cũng làm tán xạ dao động. Càng có nhiều tán xạ như vậy, nhiệt càng chảy chậm. Muốn hiểu độ dẫn nhiệt của gốm siêu cao nhiệt, phải hiểu sự tán xạ này.

Gốm entropy cao khác thường ở điểm này. Nhiều nguyên tử trộn với nhau làm mạng tinh thể không đồng đều. Sự bất quy tắc này tán xạ dao động rất mạnh. Kết quả là nhiệt không lan tốt. Nếu mục tiêu là chặn nhiệt, tính chất này có ích. Lý do là nhiệt ở bề mặt đi vào bên trong ít hơn.

Phương pháp tính toán thế giới nhỏ bé này là động lực học phân tử. Ta thử di chuyển từng nguyên tử trong máy tính. Ta theo dõi các nguyên tử rung và va chạm như thế nào. Tính toán này xử lý những khoảng thời gian ngắn ở cấp nano giây. Nano giây là 1 phần tỷ giây. Đó là một thế giới nhỏ và nhanh đến mức ấy.

Công thức rút ra độ dẫn nhiệt là công thức Green-Kubo. Công thức này theo dõi dao động của dòng nhiệt trong một thời gian dài. Nó giống như quan sát gợn sóng trên mặt nước lặng để biết tính chất của nước. Ta thu thập các dao động của dòng nhiệt do chính các nguyên tử tạo ra. Từ các dao động đó, ta tính độ dẫn nhiệt. Công thức này được Green và Kubo xây dựng vào thập niên 1950.

Độ chính xác của tính toán phụ thuộc vào việc ta biết lực giữa các nguyên tử tốt đến đâu. Quy tắc xử lý lực này được gọi là thế năng liên nguyên tử. Trước đây, người ta dùng các quy tắc giản lược do con người viết bằng tay. Các quy tắc này nhanh nhưng kém chính xác. Ngược lại, tính toán lượng tử chính xác nhưng quá chậm.

Ở đây, thế năng liên nguyên tử học máy (MLIP) bắc một cây cầu. Trí tuệ nhân tạo học kết quả tính toán lượng tử rồi dự đoán lực. Nó cho độ chính xác gần với tính toán lượng tử với tốc độ nhanh. Phương pháp này được trình bày kỹ trong các chương trước của cuốn sách này. Trong tính toán độ dẫn nhiệt ở nhiệt độ siêu cao, MLIP cũng được dùng.

Các nghiên cứu gần đây cho thấy sức mạnh của sự kết hợp này. Năm 2021, một nhóm nghiên cứu đã xử lý zirconia bằng MLIP. Họ tính cùng lúc chuyển pha theo nhiệt độ và dẫn nhiệt. Họ rút ra dòng nhiệt bằng phương pháp Green-Kubo. Phương pháp này chứa đầy đủ các tác động phức tạp của dao động nguyên tử.

Tính toán này có một giới hạn cần nêu rõ. Phương pháp Green-Kubo tìm độ dẫn nhiệt khi vật liệu ở trạng thái cân bằng. Giá trị này cho biết tính chất cơ bản của phần bên trong vật liệu. Nhưng bề mặt nơi xảy ra mài mòn bóc tách không ở trạng thái cân bằng. Ở đó, bức xạ và phản ứng hóa học bề mặt chi phối mạnh dòng nhiệt. Vì vậy, chỉ riêng giá trị Green-Kubo không thể giải thích toàn bộ nhiệt ở bề mặt. Giá trị này phù hợp để dùng làm đầu vào cho dẫn nhiệt bên trong linh kiện.

Vì sao phải đi xuống tận thế giới nhỏ bé này. Chẳng phải chỉ cần đo trực tiếp độ dẫn nhiệt là được sao. Ở nhiệt độ siêu cao, đo đạc rất khó. Không dễ đo bên trong gốm ở 2000 độ C. Quá trình oxy hóa diễn ra và chính vật liệu cũng thay đổi. Vì vậy, cần một con đường lấy giá trị bằng tính toán.

Tính toán còn có một ưu điểm khác. Ta có thể xem trước cả những thành phần chưa chế tạo. Gốm entropy cao có vô số tổ hợp. Không thể chế tạo và đo tất cả chúng. Tính toán cho biết sớm khả năng dẫn nhiệt của các ứng viên. Khi đó, ta chỉ chọn các tổ hợp có triển vọng để chế tạo thật.

Dòng chảy này vẫn tiếp tục. Một nghiên cứu năm 2026 đã dùng Green-Kubo và MLIP để dự đoán độ dẫn nhiệt rất thấp. Sự kết hợp này đang tìm được vị trí trong nhiều vật liệu. Tuy vậy, tính toán vẫn cần máy tính lớn. Cũng phải xử lý sai số như cắt thời gian hoặc dao động thống kê. Vì vậy, người ta tính nhiều lần rồi lấy trung bình.

Nguyên tắc kiểm chứng ở đây cũng giống vậy. Giá trị tính toán phải được đối chiếu với giá trị thực nghiệm. Ta thử tính toán trên một vài vật liệu đã biết. Nếu tính toán khớp tốt với thực nghiệm, nó tạo

niềm tin cho vật liệu mới. Giá trị chưa qua quá trình đối chiếu này chỉ dừng ở mức tham khảo. Tính toán và thực nghiệm là hai con mắt kiểm tra lẫn nhau.

### 5.4.2 Phương trình dòng nhiệt của toàn bộ linh kiện

Tiểu mục trước đã xử lý một thế giới rất nhỏ. Bây giờ ta mở rộng tầm nhìn ra toàn bộ linh kiện. Nhiệt độ của linh kiện khác nhau theo vị trí và thay đổi theo thời gian. Quy tắc xử lý sự thay đổi này là phương trình dẫn nhiệt. Phương trình được viết bằng công thức, nhưng ý nghĩa có thể diễn giải bằng lời.

Ý nghĩa của phương trình dẫn nhiệt rất dễ hiểu. Sự thay đổi nhiệt độ tại một vị trí được quyết định bởi chênh lệch giữa nhiệt đi vào và nhiệt đi ra. Nhiệt chảy từ phía nóng sang phía lạnh. Tốc độ của dòng chảy này là độ dẫn nhiệt. Giá trị đã tính ở trên đi vào đây.

Phương trình này khó vì vài lý do. Trước hết, độ dẫn nhiệt thay đổi theo nhiệt độ. Khi nhiệt độ tăng, độ dẫn nhiệt cũng thay đổi. Điều này được gọi là phi tuyến. Phi tuyến là quan hệ trong đó nguyên nhân và kết quả không tỷ lệ thẳng với nhau. Bên cạnh đó, vật liệu có tính chất khác nhau theo từng hướng. Nhiệt chảy tốt hơn theo hướng sợi được đặt.

Vấn đề bản thân vật liệu thay đổi cũng được cộng thêm vào đây. Khi oxy hóa tiến triển, độ dẫn nhiệt thay đổi. Khi bề mặt bị bào mòn, độ dày giảm. Vì vậy, các giá trị trong phương trình thay đổi theo thời gian. Oxy hóa và creep ở mục trước đan vào dòng nhiệt của mục này. Do đó, phải giải các hiện tượng này cùng nhau, không tách riêng.

Những gì xảy ra ở biên cũng phức tạp. Trên bề mặt siêu vượt âm, nhiều loại nhiệt chồng lên nhau. Nhiệt do không khí làm nóng đi vào. Cũng có bức xạ, trong đó bề mặt phát nhiệt ra ngoài bằng ánh sáng. Đó là hiện tượng một vật nóng đỏ lên và mất nhiệt. Cộng thêm vào đó là phát nhiệt hóa học.

Phát nhiệt hóa học xuất hiện khi các nguyên tử tụ lại trên bề mặt. Ở tốc độ siêu vượt âm, phân tử không khí bị tách thành nguyên tử. Khi các nguyên tử này ghép cặp trở lại trên bề mặt, chúng phát nhiệt. Hiện tượng này được gọi là tái hợp xúc tác. Độ lớn của nhiệt này thay đổi theo vật liệu bề mặt. Vì vậy, chọn vật liệu nào sẽ chi phối mạnh nhiệt độ bề mặt.

Tất cả các loại nhiệt này được đưa vào điều kiện biên của một phương trình. Nhiệt đi vào, bức xạ đi ra và phát nhiệt hóa học được đặt cùng nhau. Thời gian cũng cùng trôi. Nhiệt liên tục thay đổi theo quỹ đạo bay. Điều này được gọi là trạng thái không ổn định. Nó có nghĩa là xử lý một quá trình đang trôi, không phải một khoảnh khắc đứng yên.

Phương pháp truyền thống để giải phương trình này là phương pháp phần tử hữu hạn. Ta chia linh kiện thành các mảnh nhỏ. Ta giải phương trình ở từng mảnh rồi nối chúng lại. Phương pháp này chính xác nhưng mất nhiều thời gian tính toán. Cần tính lặp lại ở từng thời điểm. Khó giải toàn bộ quỹ đạo Mach 7 theo thời gian thực.

Phát triển tên lửa và phương tiện siêu vượt âm cần câu trả lời nhanh. Mỗi lần thay đổi thiết kế, phải tính lại nhiệt độ. Trong quá trình thử nghiệm, cần dự báo gần với thời gian thực. Tính toán chậm không đáp ứng được yêu cầu này. Vì vậy, toán tử trí tuệ nhân tạo ở tiểu mục tiếp theo xuất hiện.

### 5.4.3 Mạng nơ-ron biến hàm thành hàm (DeepONet)

Mạng nơ-ron toán tử sâu là trí tuệ nhân tạo học quan hệ biến một hàm thành một hàm. Nó được gọi là DeepONet. Mạng nơ-ron thông thường nhận số và cho ra số. DeepONet nhận hàm và cho ra hàm. Ở đây, hàm là những thứ như đầu vào nhiệt thay đổi theo thời gian.

Ý tưởng này được nhóm Lu trình bày có hệ thống vào năm 2021. Họ xây dựng lý thuyết và cấu trúc của mạng nơ-ron học toán tử. Nghiên cứu này bắt rễ từ một định lý toán học. Đó là định lý cho rằng có thể xấp xỉ toán tử bằng mạng nơ-ron. Định lý xấp xỉ này trở thành nền tảng lý thuyết nâng đỡ DeepONet.

DeepONet gồm hai mạng lưới. Một mạng đọc hàm đầu vào. Mạng còn lại đọc vị trí và thời điểm cần xem đáp án. Kết quả của hai mạng được nhân với nhau để tạo ra đáp án. Mạng phía trước được gọi là nhánh, mạng phía sau được gọi là thân. Nhánh và thân gặp nhau rồi tạo ra nhiệt độ.

Vì sao việc biến hàm thành hàm lại quan trọng. Đầu vào của bài toán nhiệt không phải một con số. Đầu vào là gia nhiệt khí động thay đổi theo thời gian. Đầu ra cũng là trường nhiệt độ thay đổi theo vị trí và thời gian. Bài toán như vậy là quan hệ nối một hàm với một hàm. DeepONet học toàn bộ quan hệ này.

Điểm mạnh của DeepONet là tốc độ. Sau khi học một lần, nó lập tức đưa ra đáp án cho đầu vào mới. Nó không giải lại từ đầu mỗi lần như phương pháp phần tử hữu hạn. Có thể quét nhanh nhiều điều kiện bay. Nó phù hợp để xem nhiệt độ khi thay đổi thiết kế.

Vấn đề nằm ở dữ liệu học. Muốn học đáp án tốt, cần nhiều dữ liệu chính xác. Nhưng dữ liệu chính xác rất đắt để tạo ra. Vì tính toán phần tử hữu hạn và tính toán động lực học phân tử mất nhiều thời gian. Nếu dữ liệu ít, trí tuệ nhân tạo không học tốt. Phương pháp đa độ trung thực giải quyết vấn đề này.

Độ trung thực có nghĩa là phép tính chính xác đến mức nào. Độ trung thực cao là tính toán chính xác nhưng đắt. Độ trung thực thấp là tính toán thô hơn nhưng rẻ. Đa độ trung thực dùng cả hai. Ta gom nhiều dữ liệu rẻ và chỉ gom một ít dữ liệu đắt.

Ý tưởng của phương pháp này gần với phác thảo và tô phủ. Ta vẽ bức tranh lớn trước bằng tính toán rẻ. Sau đó dùng tính toán đắt để tô phủ chi tiết lên trên. Trí tuệ nhân tạo học chênh lệch giữa đáp án độ trung thực thấp và đáp án độ trung thực cao. Chênh lệch này được gọi là phần dư. Nếu học phần dư, ta có thể tăng độ chính xác ngay cả với ít dữ liệu đắt.

Năm 2022, nhóm Lu đã áp dụng DeepONet đa độ trung thực này vào truyền nhiệt nano. Thứ họ xử lý là bài toán dòng nhiệt trong một thế giới rất nhỏ. Phương pháp đa độ trung thực đã giảm sai số mạnh với cùng lượng dữ liệu đắt. Nó cũng làm giảm mạnh lượng dữ liệu đắt cần có. Tuy vậy, nghiên cứu này xử lý truyền nhiệt nano. Nó không phải là thực chứng nối oxy hóa, mài mòn bóc tách của gốm siêu cao nhiệt hoặc tải nhiệt Mach 7.

Cũng cần giải thích từ đồng nhất hóa. Vật liệu gồm composite trộn lẫn sợi và lỗ rỗng. Nếu xử lý từng phần của cấu trúc phức tạp này, tính toán sẽ quá nặng. Vì vậy, người ta đổi một vùng nhỏ thành các tính chất trung bình. Điều này được gọi là đồng nhất hóa. Tính toán độ trung thực thấp vẽ bức tranh lớn bằng các tính chất trung bình này.

Cũng cần chỉ ra vì sao việc nối hai thế giới lại khó. Thế giới nguyên tử chuyển động ở cấp nano giây. Thế giới linh kiện xử lý hàng trăm giây. Quy mô thời gian khác nhau hơn mười bậc độ lớn. Quy mô kích thước cũng khác nhiều như vậy. Không thể tính khoảng cách này cùng một lúc.

Việc nối hai thế giới bằng phương pháp này trong kết cấu nhiệt gốm là con đường mà cuốn sách này đề xuất. Độ trung thực thấp trở thành mô hình đồng nhất hóa giản lược. Độ trung thực cao trở thành phân tích chính xác nối động lực học phân tử và phần tử hữu hạn. Trí tuệ nhân tạo là bức tranh kết nối hai bên. Khi đó, dẫn nhiệt trong thế giới rất nhỏ có thể nối sang nhiệt độ linh kiện. Kết nối này vẫn chưa được kiểm chứng rộng rãi trong kết cấu nhiệt gốm.

#### 5.4.4 Học có đưa quy tắc vật lý vào và tăng tốc

Trí tuệ nhân tạo nhanh không có nghĩa là được đưa ra bất kỳ đáp án nào. Đáp án phải tuân thủ định luật vật lý. Vì vậy, người ta đưa quy tắc vật lý vào quá trình học. Phương pháp này có cùng gốc với dự đoán tuổi thọ ở mục trước.

Trước hết, ta học để khớp dữ liệu. Ta khớp cả dữ liệu độ trung thực thấp và dữ liệu độ trung thực cao. Ta dẫn dắt để đáp án của trí tuệ nhân tạo tiến gần dữ liệu. Ta phạt chênh lệch giữa đáp án và dữ liệu. Mức phạt này được gọi là hàm mất mát.

Tiếp đó, ta thêm việc học để tuân thủ phương trình. Ta đưa đáp án của trí tuệ nhân tạo vào phương trình dẫn nhiệt. Nếu đáp án thỏa phương trình tốt, mức phạt nhỏ. Nếu lệch, mức phạt lớn. Làm như vậy, nó tạo ra đáp án phù hợp vật lý cả ở những vị trí không có dữ liệu.

Khi đưa quy tắc vật lý vào, mô hình vẫn chịu được cả khi dữ liệu ít. Học thuần dữ liệu cần nhiều dữ liệu. Nếu dữ liệu ít, nó lấp khoảng giữa bằng những giá trị sai lệch. Quy tắc vật lý lấp khoảng trống này bằng phương trình. Nó phù hợp với bài toán gốm, nơi dữ liệu nhiệt độ siêu cao rất hiếm. Khả năng chịu được dữ liệu ít là lợi ích lớn của học dựa trên vật lý.

Điều kiện biên cũng được đưa vào theo cùng cách. Ta buộc mô hình giữ cân bằng nhiệt trên bề mặt. Nhiệt đi vào và bức xạ đi ra phải khớp với nhau. Nếu vi phạm quy tắc này, ta đặt mức phạt. Nhiệt độ bề mặt ảnh hưởng lớn đến tuổi thọ vật liệu, nên điều kiện này quan trọng hơn hết.

Các mức phạt này được gom thành một. Ta cộng mức phạt dữ liệu, mức phạt phương trình và mức phạt biên. Mỗi mức phạt được gán trọng số. Trí tuệ nhân tạo học để làm mức phạt tổng thể này nhỏ nhất có thể. Khi đó, nó đưa ra dự báo vừa nhanh vừa phù hợp vật lý.

Muốn đưa quy tắc vật lý vào, cần tính vi phân. Phải đo nhiệt độ thay đổi bao nhiêu theo vị trí và thời gian. Có một kỹ thuật tự động tính vi phân này. Nó được gọi là vi phân tự động. Tuy vậy, phép tính này nặng và làm giảm tốc độ.

Vì vậy, người ta dùng nhiều cách làm nhẹ phép tính. Ta chia tính toán để giảm phần trùng lặp. Ta đổi tọa độ bằng hàm trơn để làm vi phân dễ hơn. Nhờ các tối ưu này, ta có được suy luận nhanh hơn nhiều so với tính toán truyền thống. Tuy vậy, hệ số tăng tốc hoặc giá trị sai số cụ thể khác nhau theo từng bài toán. Những giá trị đó chỉ nên được trích dẫn trong các trường hợp đã được kiểm chứng.

Giá trị của phương pháp này lộ rõ ở giai đoạn thiết kế. Khi thay đổi hình dạng vòi phun, phân bố nhiệt độ thay đổi. Với tính toán truyền thống, mỗi lần thay đổi đều phải chờ lâu. Toán tử trí tuệ nhân tạo lập tức cho thấy nhiệt độ mới. Nhờ đó, trong một ngày có thể so sánh hàng chục hình dạng. Người thiết kế đến được đáp án tốt hơn nhanh hơn.

Nó cũng hữu ích tại hiện trường thử nghiệm. Trong thử nghiệm động cơ, cần dự báo gần với thời gian thực. Phải biết trước vị trí nào sẽ đạt đến nhiệt độ nguy hiểm. Toán tử trí tuệ nhân tạo đưa ra dự báo này nhanh. Tuy vậy, phải liên tục hiệu chỉnh bằng dữ liệu thử nghiệm thì mới đáng tin. Việc giảm chênh lệch giữa dự báo và đo thực tế sẽ tiếp tục.

Giới hạn của phương pháp này cũng rõ ràng. Nếu đưa quy tắc vật lý sai, đáp án sẽ xấu đi. Việc đặt trọng số giữa các quy tắc cũng tốn nhiều công. Dữ liệu thực tế ở nhiệt độ siêu cao ít, nên kiểm chứng cũng khó. Vì vậy, công nghệ này vẫn đang ở giai đoạn vừa nghiên cứu vừa thử nghiệm. Nó được dùng như công cụ giảm số lần thử nghiệm, chứ không thay thế thử nghiệm.

## Tóm lược

Chương này đã bàn về bề mặt nóng của tên lửa và phương tiện bay siêu vượt âm. Vật liệu ở khu vực này phải chịu cùng lúc nhiệt, oxy hóa, lực và thời gian dài. Kim loại là chưa đủ, nên cần vật liệu gốm. Composite sợi carbon và gốm siêu nhiệt độ cao là các ứng viên cho vai trò đó. Biết trước tuổi thọ của những vật liệu này là chìa khóa của phát triển.

Trước hết, ta đã xem vật liệu hư hỏng như thế nào. Composite sợi carbon hư hỏng khi oxy đi vào theo các vết nứt. Silicon carbide có thể có hoặc mất lớp màng bảo vệ tùy theo nhiệt độ và áp suất oxy. Gốm

siêu nhiệt độ cao bị bong bề mặt do sự bay hơi của màng oxy hóa và biến đổi pha. Quá trình hư hỏng này là đối tượng của dự đoán tuổi thọ. Hư hỏng thường bắt đầu âm thầm bên trong vật liệu, vì vậy phải đọc phần bên trong chứ không phải chỉ nhìn bề ngoài.

Tiếp theo, ta đã xem trí tuệ nhân tạo xử lý hư hỏng này như thế nào. Các hạt tinh thể được đổi thành điểm, còn ranh giới được đổi thành đường để tạo thành đồ thị. Khi thêm thời gian vào đó, ta theo dõi quá trình hư hỏng tích tụ. Bằng cách đưa quy luật vật lý vào học máy, ta cũng nhìn trước được vùng siêu nhiệt độ cao nơi dữ liệu còn ít. Quy tắc tiêu biểu là hư hỏng không quay ngược lại.

Sau đó, ta đã xem thiết kế ngược. Ta đặt ra hiệu năng mong muốn trước, rồi tìm ngược lại cấu trúc tương ứng. Mô hình khuếch tán tạo cấu trúc từ nhiều mờ. Nó đưa vào tính chất không dao động theo phương hướng để phù hợp với vật lý. Góc sợi và độ dày lớp phủ giao diện được xác định cùng nhau. Mô hình khuếch tán không đưa ra một đáp án duy nhất, mà đưa ra nhiều ứng viên. Khi đó, con người chọn trong các ứng viên ấy cấu trúc dễ chế tạo.

Cuối cùng, ta đã xem sự khớp nối đa thang đo của dòng nhiệt. Cần nối thế giới dao động nguyên tử với thế giới của toàn bộ chi tiết. Động lực học phân tử Green-Kubo tính độ dẫn nhiệt của thế giới nhỏ. Mạng nơ-ron toán tử sâu chuyển giá trị đó thành nhiệt độ của chi tiết. Phương pháp đa độ trung thực nối dữ liệu rẻ với dữ liệu đắt. Nhờ sự khớp nối này, ta có thể dự đoán nhanh nhiệt độ chi tiết. Dự đoán nhanh này hữu ích cả ở giai đoạn thiết kế lẫn tại hiện trường thử nghiệm.

Tất cả các công cụ này có một điểm chung. Đó là trí tuệ nhân tạo không thay thế thử nghiệm. Trí tuệ nhân tạo thu hẹp danh sách ứng viên và đẩy nhanh dự đoán. Việc xác nhận cuối cùng thuộc về thử nghiệm hầm gió, arcjet và thử nghiệm đốt. Chỉ khi trí tuệ nhân tạo và thử nghiệm thực tế đi cùng nhau, phát triển mới thật sự nhanh hơn.

Một điểm chung nữa là sự kết hợp với vật lý. Trí tuệ nhân tạo trong chương này không học chỉ bằng dữ liệu. Nó đưa các định luật nhiệt động lực học và cân bằng lực vào quá trình học. Vì vậy, nó xử lý được cả vùng siêu nhiệt độ cao nơi dữ liệu còn ít. Đây là bức tường khó vượt qua bằng học thuần dữ liệu. Sự gặp nhau giữa vật lý và trí tuệ nhân tạo là chìa khóa cốt lõi của chương này.

Các phương pháp trong chương này nối với nhau. Dự đoán dòng nhiệt cho ta nhiệt độ bề mặt. Nhiệt độ bề mặt quyết định tốc độ oxy hóa và creep. Oxy hóa và creep dẫn đến dự đoán tuổi thọ. Dự đoán tuổi thọ trở thành mục tiêu của thiết kế ngược. Có thể xem đó là một chuỗi liên kết.

Vẫn còn nhiều bài toán phải giải. Dữ liệu thực ở siêu nhiệt độ cao vẫn thiếu. Các tuyên bố về hiệu năng của mô hình thiết kế ngược cần được kiểm chứng bằng vật thật. Cách đưa quy tắc vật lý vào mô hình cũng cần được chỉnh kỹ hơn. Khi các bài toán này được giải dần, thiết kế hệ thống bảo vệ nhiệt thế hệ tiếp theo sẽ tiến về phía trước. Hệ thống bảo vệ nhiệt là toàn bộ cấu trúc bảo vệ phương tiện bay khỏi bề mặt nóng.

Hàn Quốc cũng đang bước vào lĩnh vực này. Nuri và phương tiện phóng thế hệ tiếp theo xử lý các chi tiết nóng. Tên lửa tái sử dụng phải dùng các chi tiết nhiều lần. Vì vậy, dự đoán tuổi thọ càng trở nên quan trọng. Những phương pháp dùng chung trí tuệ nhân tạo và vật lý này sẽ là nền tảng. Con mắt đọc được thế giới nhỏ của vật liệu sẽ bảo vệ an toàn cho những chuyến bay lớn.

## Tài liệu căn cứ liên quan

- 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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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. Bản tiền ấn phẩm, chưa qua bình duyệt. <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- $\kappa$  materials using machine learning interatomic potentials. *Small*, 22(8), 2513476. <https://doi.org/10.1002/sml.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>

## Chương 6. Mô hình hóa ứng suất nhiệt trong quy trình chế tạo bồi đắp kim loại (LPBF/DED) và nhận biết khuyết tật tại chỗ (In-situ)

### Mở đầu

Hãy tưởng tượng ta đang nướng bánh trong bếp. Ta đổ bột mì đã nhào vào khuôn, cho vào lò, rồi lấy ra một khối bánh. Cách đổ vật liệu vào khuôn rồi làm ra sản phẩm trong một lần như vậy là phương pháp chế tạo cũ. Linh kiện tên lửa trong thời gian dài cũng được làm theo cách này. Người ta gọt bớt từ một khối kim loại lớn, hoặc làm riêng nhiều mảnh rồi ghép lại.

Nhưng bên trong buồng đốt và vòi phun của động cơ tên lửa có hàng trăm kênh làm mát rất nhỏ chạy qua. Muốn bảo vệ thành trước ngọn lửa nóng, phải cho nhiên liệu lạnh chảy qua các kênh đó. Cấu trúc bên trong phức tạp như vậy rất khó tạo ra bằng cách gia công cắt gọt. Nếu ghép các mảnh lại với nhau, mối nối sẽ trở thành điểm yếu.

Chế tạo bồi đắp kim loại mở ra một con đường khác. Phương pháp này tạo linh kiện bằng cách xếp kim loại từng lớp một. Giống như máy in 3D ép nhựa thành từng đường rồi xếp chồng lên nhau, nó nung chảy bột kim loại hoặc dây kim loại rồi bồi đắp lên. Các kênh rỗng phức tạp bên trong cũng có thể được tạo ra trong một quy trình. Vì không có mối nối, số điểm yếu cũng giảm.

Phương pháp này đem lại tự do, nhưng cũng đem lại nỗi lo mới. Trong lúc kim loại được nung chảy và xếp thành từng lớp, các lỗ rất nhỏ có thể xuất hiện. Khi mỗi lớp nóng lên rồi nguội đi, linh kiện cũng có thể bị cong vênh. Hình dạng tinh thể có thể khác nhau theo từng hướng, khiến độ bền lệch về một phía. Những khuyết tật này khó nhìn thấy bằng mắt. Nhưng trong linh kiện tên lửa, một khuyết tật nhỏ cũng có thể dẫn đến tai nạn lớn.

Chương này xem xét trí tuệ nhân tạo giúp giải quyết vấn đề này như thế nào. Trí tuệ nhân tạo làm ba việc. Nó phát hiện nhanh khuyết tật trong khi chế tạo. Nó thay thế các phép tính chậm để dự đoán nhanh ứng suất nhiệt và biến dạng. Nó giúp thay đổi điều kiện quy trình theo từng thời điểm. Khi lần lượt đi theo ba việc này, ta sẽ thấy phác thảo của một nhà máy có thể tự chế tạo linh kiện tên lửa gần như không có khuyết tật.

Nền tảng của các công nghệ được bàn trong chương này là một ý nghĩ. Chỉ trí tuệ nhân tạo thì chưa đủ, và chỉ vật lý cũng chưa đủ. Trí tuệ nhân tạo chỉ học từ dữ liệu để bị nhiễu làm lung lay. Nếu chỉ dùng tính toán vật lý thì quá chậm. Khi kết hợp hai bên, chúng bù cho điểm yếu của nhau. Sự kết hợp này là cốt lõi của học máy dựa trên vật lý, và nhiều công nghệ trong chương này là ví dụ của điều đó.

### 6.1 Hai phương pháp chế tạo linh kiện tên lửa bằng cách bồi đắp

#### Nội dung của mục này

Có hai phương pháp tiêu biểu để bồi đắp linh kiện động cơ tên lửa bằng kim loại. Một phương pháp là chiếu laser lên lớp bột kim loại được trải mỏng. Phương pháp còn lại là đưa bột kim loại hoặc dây kim loại vào tiêu điểm laser, đồng thời nung chảy và bồi đắp. Hai phương pháp này khác nhau về kích thước và độ chính xác của linh kiện được chế tạo.

Mục này cũng giới thiệu ba kim loại dùng trong tên lửa. Ba kim loại này có tính chất khác nhau. Những điểm cần thận trọng khi chế tạo cũng khác nhau. Ta phải biết điều gì xảy ra bên trong khi kim loại nóng chảy rồi đông đặc. Khi đó mới hiểu vì sao các công nghệ trí tuệ nhân tạo xuất hiện sau đây lại cần thiết. Mục này giới thiệu tính chất của ba kim loại và đặt nền cho phần sau.

### 6.1.1 Cơ quan đẩy tên lửa thể hệ mới và chế tạo bồi đắp kim loại (Additive Manufacturing)

Động cơ tên lửa làm việc trong điều kiện cực hạn. Áp suất trong buồng đốt đạt đến hàng trăm lần áp suất lớp ô tô. Nhiệt độ ở nơi nhiên liệu cháy vượt quá 3 nghìn độ. Ngược lại, trong các kênh làm mát có nhiên liệu lỏng gần âm 200 độ chảy qua. Một linh kiện phải chịu đồng thời cả cái nóng và cái lạnh này.

Nếu chế tạo những linh kiện như vậy theo cách cũ, công việc rất mất công. Phải gia công từng kênh làm mát, rồi ghép nhiều mảnh lại bằng phương pháp gọi là brazing. Brazing là phương pháp nối bằng cách nung chảy một kim loại khác vào giữa các kim loại. Càng nhiều mảnh thì càng nhiều mối nối, và việc chế tạo mất vài tháng. Mối nối về sau dễ trở thành nơi rò rỉ hoặc nứt vỡ.

Chế tạo bồi đắp kim loại rút ngắn mạnh quá trình này. Laser Powder Bed Fusion (LPBF) trải bột kim loại thành một lớp mỏng. Sau đó chiếu laser lên trên. Chỉ những vị trí cần thiết được nung chảy và gắn lại, rồi lại trải một lớp bột mỏng và nung chảy tiếp. Lặp lại quá trình này hàng nghìn lần thì linh kiện hoàn thành. Phương pháp này phù hợp với các linh kiện nhỏ và chính xác, như liner buồng đốt có nhiều kênh dày đặc.

Directed Energy Deposition (DED) có cách làm hơi khác. Phương pháp này đưa bột kim loại hoặc dây kim loại vào tiêu điểm laser, đồng thời nung chảy và bồi đắp. Có thể hình dung gần giống việc dùng cọ quét thêm từng lớp sơn, nhưng vật liệu ở đây là kim loại. Tốc độ bồi đắp nhanh và phù hợp với linh kiện lớn. Khi chế tạo skirt vòi phun có đường kính hơn 1 mét, người ta dùng phương pháp này.

Hai phương pháp có ưu và nhược điểm riêng. Laser Powder Bed Fusion rất chính xác nhưng chậm. Vì phải trải lớp mỏng hàng nghìn lần. Kích thước linh kiện có thể chế tạo cũng bị giới hạn bởi kích thước buồng máy. Ngược lại, Directed Energy Deposition nhanh và có thể tạo linh kiện lớn, nhưng bề mặt thô. Sau khi chế tạo, cần thêm công đoạn gia công cắt gọt để hoàn thiện.

Chất lượng bột cũng quyết định kết quả. Hạt bột phải tròn đều thì mới trải mỏng tốt. Nếu kích thước hạt không đều, khoảng trống sẽ xuất hiện. Nếu bột hút ẩm, bột khí sẽ sinh ra trong lúc chế tạo. Vì vậy bột được bảo quản và quản lý trong trạng thái khô. Linh kiện tốt bắt đầu từ bột tốt.

Cả hai phương pháp đều có sức mạnh tạo linh kiện thành một khối. Liner buồng đốt, đầu phun, cánh bơm của turbopump, và vòi phun lớn có thể được tạo hình không có mối nối. Chương trình công nghệ đẩy tên lửa RAMPT của Cơ quan Hàng không Vũ trụ Hoa Kỳ (NASA) đã chế tạo vòi phun cỡ lớn bằng phương pháp này. Theo tài liệu công khai, họ đã bồi đắp một vòi phun có kênh làm mát với đường kính khoảng 1,5 mét và chiều dài khoảng 1,8 mét trong 90 ngày. Vòi phun này có kích thước bằng 65 phần trăm vòi phun động cơ lớn của phương tiện phóng vũ trụ.

Chương trình này cũng công bố kết quả thử nghiệm. Trong các thử nghiệm phát triển kết hợp nhiều injector, buồng đốt và vòi phun, linh kiện đã chịu được quá trình đốt trong thời gian dài. Tổng thời gian đốt tích lũy vượt xa 10 nghìn giây, và việc đánh lửa cũng được lặp lại hàng trăm lần. Điều đó có nghĩa là linh kiện được tạo bằng cách bồi đắp vật liệu từng lớp có thể chịu được quá trình đốt thật của tên lửa. Kết quả này cho thấy chế tạo bồi đắp đang vượt ra khỏi phòng thí nghiệm để đi vào linh kiện dùng cho bay.

Lợi ích mà phương pháp này đem lại nằm ở thời gian và số lượng linh kiện. Cách cũ phải chế tạo riêng hàng trăm mảnh rồi ghép lại. Chế tạo bồi đắp giảm hàng trăm mảnh đó xuống còn vài mảnh. Khi số mảnh giảm, mối nối giảm và thời gian chế tạo cũng giảm. Khi muốn thay đổi thiết kế, người ta cũng có thể nhanh chóng chế tạo lại để thử. Tốc độ này đẩy nhanh việc phát triển động cơ mới.

### 6.1.2 Đặc tính luyện kim theo từng hợp kim và hành vi chuyển pha

Kim loại dùng trong linh kiện tên lửa có nhiệm vụ riêng. Mục này xem xét ba kim loại tiêu biểu. Đó là Inconel 718, GRCo-42 và C-103. Luyện kim học là ngành nghiên cứu cấu trúc bên trong kim loại và sự biến đổi của cấu trúc đó. Cùng một kim loại nhưng cấu trúc bên trong sẽ khác nhau tùy theo cách nó nguội đi.

Ba kim loại giữ ba vị trí khác nhau. Inconel 718 đảm nhiệm các linh kiện bền, chịu lực. GRCo-42 đảm nhiệm thành buồng đốt cần thoát nhiệt nhanh. C-103 đảm nhiệm phần cuối vòi phun tiếp xúc trực tiếp với ngọn lửa. Trong một động cơ, ba kim loại chia nhau nhiệm vụ. Vì vậy cần có luyện kim học để chế tạo tốt từng kim loại.

Inconel 718 là siêu hợp kim nền nickel. Siêu hợp kim là hợp kim được tạo ra để giữ độ bền ngay cả ở nhiệt độ cao. Kim loại này được dùng rộng rãi trong turbopump, van và đường ống áp suất cao. Nó bền, dễ hàn gắp, và chịu nóng tốt. Vì vậy nó phù hợp với các linh kiện chịu lực của động cơ.

Vấn đề xuất hiện khi làm nguội nhanh. Kim loại được laser nung chảy đông đặc trong chớp mắt. Khi đó các nguyên tố như niobium và molybdenum dồn về ranh giới giữa các tinh thể. Sự dồn tụ này được gọi là vi thiên, tức microsegregation. Nếu thiên nghiêm trọng, một hợp chất yếu gọi là pha Laves (Laves phase) sẽ hình thành tại đó. Hợp chất này lấy mất các nguyên tố làm kim loại bền lên, khiến độ bền giảm.

Vì vậy Inconel 718 cần xử lý nhiệt sau khi chế tạo. Linh kiện được nung lại ở nhiệt độ cao để hòa tan và loại bỏ hợp chất yếu. Sau đó hạ nhiệt độ để dẫn dắt sự hình thành các hạt tăng bền rất nhỏ và đồng đều. Phải trải qua quá trình này thì linh kiện mới bền trước mỗi. Mỗi là hiện tượng vật liệu dần nứt khi chịu lực lặp đi lặp lại.

Các hạt tăng bền này có tên riêng. Đó là một pha rất nhỏ do nickel và niobium tạo ra, gọi là gamma double-prime. Theo công thức hóa học, ba nguyên tử nickel gắn với một nguyên tử niobium. Phần lớn độ bền của Inconel 718 đến từ pha này. Đây là pha khác với gamma-prime thường gặp trong các siêu hợp kim nickel khác. Vì vậy kiểm soát tốt niobium là chìa khóa của kim loại này.

GRCo-42 là hợp kim đồng do NASA phát triển cho buồng đốt tên lửa. Đồng dẫn nhiệt rất tốt. Muốn nhanh chóng rút nhiệt nóng ra khỏi thành buồng đốt, cần tính chất này. Nhưng đồng nguyên chất mềm đi khi nóng lên. Vì vậy người ta cho một lượng nhỏ chromium và niobium vào đồng. Hai nguyên tố này tạo ra các hạt rất nhỏ để giữ đồng lại.

Khi chế tạo GRCo-42 bằng laser, có một khó khăn. Đồng phản xạ ánh sáng tốt. Nó hắt lại một phần lớn ánh sáng laser như gương. Vì vậy muốn nung chảy đủ, cần laser mạnh hoặc laser ánh sáng xanh lục. Sau khi chế tạo, linh kiện đi qua công đoạn ép bằng áp suất và nhiệt độ cao để loại bỏ lỗ rỗng bên trong. Sau công đoạn này, nó vẫn dẫn nhiệt tốt và giữ độ bền trong môi trường nóng hơn 800 độ.

Hãy xem vì sao thành buồng đốt phải thoát nhiệt tốt như vậy. Một phía của thành tiếp xúc với ngọn lửa hơn 3 nghìn độ. Phía còn lại có nhiên liệu lạnh chảy qua. Bức thành mỏng này tách hai mức nhiệt đó ra. Nếu không thể rút nhiệt nhanh, thành sẽ nóng chảy hoặc nứt. Vì vậy hợp kim đồng dẫn nhiệt tốt được dùng làm vật liệu thành.

Hợp kim này đi cùng với các kênh làm mát. Người ta tạo nhiều kênh nhỏ dày đặc chạy bên trong thành. Nhiên liệu chảy qua các kênh đó và mang nhiệt của thành đi. Kênh càng phức tạp thì nhiệt càng được rút ra đồng đều. Sức mạnh tạo ra các kênh phức tạp như vậy trong một lần nằm ở chế tạo bồi đắp. Vật liệu và phương pháp chế tạo làm nhau phát huy tác dụng.

C-103 là hợp kim nhiệt độ cao nền niobium. Niobium có điểm nóng chảy vượt xa 2 nghìn độ. Vì vậy nó được dùng cho đầu vòi phun tiếp xúc trực tiếp với ngọn lửa hoặc thruster điều khiển tư thế. Người ta

pha thêm một lượng nhỏ hafnium, titanium và zirconium để điều chỉnh tính chất. Khi chế tạo vòi phun tấm mỏng cỡ lớn, người ta dùng phương pháp DED.

Khi chế tạo C-103, phải cẩn thận với không khí. Niobium nóng chảy nhanh chóng hút oxygen và nitrogen trong không khí. Nếu lượng oxygen vượt quá một mức nhất định, kim loại trở nên dễ gãy. Nó chuyển sang tính chất vỡ giòn như thủy tinh. Vì vậy phòng chế tạo được lấp đầy bằng khí bảo vệ như argon để giữ oxygen ở mức rất thấp.

Trong quá trình sử dụng, oxy hóa cũng là vấn đề lớn. Oxygen trong khí thải nóng liên tục tấn công niobium. Vì vậy bề mặt được phủ một lớp bảo vệ silicide có chứa silicon. Lớp này ngăn oxygen chạm vào kim loại. Cần cả quản lý khí quyển khi chế tạo và lớp bảo vệ khi sử dụng. Hai việc quản lý này quyết định thành bại của linh kiện C-103.

Nghiên cứu gần đây đi sâu vào quy trình chế tạo kim loại này. Một nghiên cứu năm 2025 đã chế tạo tấm mỏng C-103 bằng lắng đọng dây laser và xem xét tính chất của nó. Một nghiên cứu khác đã vẽ bản đồ khuyết tật bằng mô hình bể nóng chảy giải đồng thời nhiều hiện tượng vật lý. Tức là vẽ trước điều kiện nào sẽ sinh khuyết tật. Khi có bản đồ như vậy, ta có thể chọn điều kiện tốt để dùng. Trước khi chế tạo, ta tìm đường tránh khuyết tật.

## 6.2 Cơ chế nhận biết khuyết tật tại chỗ (In-situ) và mô hình hình ảnh dựa trên vật lý

### Nội dung của mục này

In-situ có nghĩa là quan sát ngay tại chỗ đang chế tạo. Nếu chờ linh kiện chế tạo xong rồi mới tìm khuyết tật thì đã muộn. Nếu phát hiện khuyết tật trong lúc chế tạo, ta có thể sửa ngay tại chỗ hoặc dừng lại. Mục này bàn về phương pháp nhìn thấy khuyết tật trong lúc chế tạo.

Trước hết, ta xem khuyết tật vì sao và bằng cách nào xuất hiện. Sau đó giới thiệu thiết bị quan sát vũng nhỏ do laser tạo ra. Cuối cùng, ta xem trí tuệ nhân tạo đọc hình ảnh quan sát được như thế nào. Ở đây xuất hiện phương pháp dùng đồng thời định luật vật lý và trí tuệ nhân tạo. Phương pháp này là một ví dụ của học máy dựa trên vật lý xuyên suốt cuốn sách này.

### 6.2.1 Cơ chế nhiệt lưu động lực học của sự hình thành khuyết tật trong bể nóng chảy

Khi laser nung chảy bột kim loại, một vũng nhỏ xuất hiện. Vũng này được gọi là bể nóng chảy (Melt pool). Đường kính của nó nhỏ cỡ vài sợi tóc. Khi laser đi qua, vũng này cũng di chuyển theo. Chất lượng linh kiện phụ thuộc rất nhiều vào trạng thái của vũng nhỏ này.

Bên trong vũng này, kim loại chuyển động không ngừng. Sức căng bề mặt ở nơi nóng và nơi ít nóng khác nhau. Sự khác biệt đó đẩy kim loại nóng chảy qua lại. Dòng chảy này được gọi là đối lưu Marangoni. Nó giống cảnh bề mặt xoáy trong một nồi canh. Xoáy này có khi tạo ra khuyết tật, cũng có khi xóa khuyết tật.

Một khuyết tật là rỗ khí keyhole (Keyhole porosity). Nó xuất hiện khi laser quá mạnh hoặc đi qua quá chậm. Ánh sáng mạnh khoét sâu vào kim loại, tạo ra một hố hẹp và sâu như lỗ chìa khóa. Hố này không ổn định nên đột ngột sụp xuống. Khi sụp, khí bên trong bị mắc kẹt trong kim loại và trở thành bọt khí tròn. Bọt khí bị mắc kẹt như vậy chính là rỗ khí keyhole.

Keyhole không phải chỉ có hại. Keyhole có độ sâu vừa phải làm kim loại nóng chảy sâu và gắn kết tốt. Vấn đề nảy sinh khi keyhole quá sâu và không ổn định. Vì vậy, quy trình cố giữ độ sâu của keyhole ở mức phù hợp. Người ta không tăng hay giảm công suất laser quá mức, mà tìm một cửa sổ phù hợp. Tìm được cửa sổ phù hợp là cốt lõi của quản lý quy trình.

Một khuyết tật khác là thiếu nóng chảy kết hợp (Lack-of-Fusion; LoF). Trái với keyhole, nó xuất hiện khi laser yếu hoặc khoảng cách giữa các đường quét quá rộng. Đường laser bên cạnh hoặc lớp bên dưới không nóng chảy và gắn với nhau đúng mức. Giữa chúng còn lại khoảng rỗng. Khoảng rỗng này có hình dạng sắc nhọn và bên trong có bột chưa nóng chảy.

Hai khuyết tật này có tính chất nguy hiểm khác nhau. Rỗ khí keyhole có dạng tròn nên tương đối ít nguy hiểm hơn. Thiếu nóng chảy kết hợp có dạng sắc nhọn nên lực tập trung ở đầu nhọn. Nguyên lý giống như đầu của mảnh kính nhọn dễ nứt ra. Vì vậy, thiếu nóng chảy kết hợp dễ trở thành điểm khởi đầu của vết nứt mỏi. Linh kiện tên lửa chịu vô số rung động, nên khuyết tật này trở thành rủi ro lớn.

Các nghiên cứu gần đây chia khuyết tật thành các nhóm nhỏ hơn. Một nghiên cứu năm 2026 chia lỗ rỗng thành ba loại. Không có lỗ, lỗ nhỏ hơn 15 micromet, và lỗ lớn hơn mức đó. Mức 15 micromet này là ranh giới phân loại do nghiên cứu đó đặt ra. Trong nghiên cứu đó, lỗ keyhole lớn gây hại nhiều hơn. Lỗ khí nhỏ ít gây hại hơn trong điều kiện đó. Nghiên cứu này dùng đồng thời tín hiệu photodiode và ảnh X-quang chính xác để huấn luyện cách phân loại này.

Tuy nhiên, không được lấy ngay ranh giới này làm tiêu chuẩn đạt hay không đạt. Mức nguy hiểm của lỗ rỗng không chỉ do kích thước quyết định. Vị trí, hình dạng, lực tác dụng, vật liệu và tuổi thọ yêu cầu cùng quyết định điều đó. Cùng một kích thước, nếu lỗ nằm gần bề mặt thì nguy hiểm hơn. Vì vậy, cần xem ranh giới phân loại và tiêu chuẩn an toàn kết cấu là hai việc khác nhau.

Có lý do để chia khuyết tật nhỏ như vậy. Nếu xử lý mọi lỗ rỗng như nhau, cả linh kiện còn tốt cũng bị loại thành hàng lỗi. Ngược lại, nếu bỏ qua tất cả thì sẽ bỏ sót lỗ rỗng nguy hiểm. Phải phân biệt loại và kích thước của khuyết tật mới có thể phản ứng phù hợp. Trí tuệ nhân tạo đảm nhận việc phân biệt tinh vi này.

Khuyết tật để lại dấu hiệu báo trước có thể quan sát được. Ngay trước khi keyhole hình thành, hình dạng của vũng nóng chảy dao động. Xuất hiện hoa văn cho thấy vũng sâu hơn và hẹp hơn. Mắt người không theo kịp thay đổi nhanh này. Camera chụp hàng chục nghìn khung hình mỗi giây và trí tuệ nhân tạo đọc các dấu hiệu báo trước đó. Mục tiêu là nhận ra trước khi khuyết tật hình thành.

Bắn tóe bột cũng liên quan đến khuyết tật. Bắn tóe là hiện tượng kim loại nóng chảy văng ra khi laser tác động mạnh. Giọt văng rơi lên lớp bột chưa nóng chảy. Khi laser tiếp theo đi qua vị trí đó, vật liệu không nóng chảy đúng mức. Vì vậy, bắn tóe xảy ra thường xuyên thì thiếu nóng chảy kết hợp tăng lên. Mức độ bắn tóe cũng trở thành tín hiệu báo trước khuyết tật.

### 6.2.2 Con mắt đo nhiệt độ vũng nóng chảy và đường chuẩn vật lý

Muốn nhìn vũng nóng chảy thì cần một con mắt đặc biệt. Vũng này rất sáng và thay đổi rất nhanh. Pyrometry là phương pháp đo nhiệt độ bằng ánh sáng do vật nóng phát ra. Nó dùng nguyên lý vật càng nóng thì phát ra ánh sáng khác nhau. Phương pháp này đo nhiệt độ mà không chạm vào vật.

Pyrometry đồng trục (Coaxial) quan sát ánh sáng từ cùng hướng với đường laser đi qua. Có thể hiểu là laser và camera cùng dùng chung một thấu kính. Vì vậy, dù laser đi qua đâu, camera vẫn luôn nhìn chính diện vũng nóng chảy. Camera tốc độ cao chụp hàng chục nghìn khung hình mỗi giây ghi lại độ sáng và biến đổi hình dạng của vũng. Khi so sánh ánh sáng ở hai bước sóng, ta có thể thu được nhiệt độ ít bị dao động bởi trạng thái bề mặt hơn.

Phương pháp hai bước sóng này cũng không hoàn hảo. Độ phát xạ chỉ mức độ một vật phát ra ánh sáng. Phương pháp này giả định tỷ lệ độ phát xạ ở hai bước sóng là cố định. Nếu giả định này sai, nhiệt độ đo được sẽ có sai số. Hơi kim loại bắn ra hoặc thấu kính bị bắn cũng làm giá trị dao động. Vì vậy, cần hiệu chỉnh giá trị đo theo định kỳ.

Chỉ camera là chưa đủ. Cần có một chuẩn để so sánh đâu là bình thường và đâu là bất thường. Một chuẩn như vậy là mô hình dẫn nhiệt Rosenthal. Mô hình này là một lý thuyết vật lý cũ ra đời năm 1946. Đó là công thức giải bằng tay phân bố nhiệt độ khi một nguồn nhiệt chuyển động nung nóng kim loại.

Mô hình Rosenthal chỉ là bản phác thảo thô. Nó xem kim loại như một khối bán vô hạn và đặt nguồn nhiệt thành một điểm. Nó giả định tính chất của kim loại không đổi, không phụ thuộc vào nhiệt độ. Nó cũng bỏ qua nhiệt ẩn đi vào và đi ra khi kim loại nóng chảy và sôi. Nó không chứa được chuyển động dòng chảy của kim loại nóng chảy và sự bay hơi.

Vì vậy, công thức này chỉ cho nhiệt độ chuẩn gần đúng. Chuyển động nhanh khi keyhole sụp xuống không thể giải thích bằng công thức này. Khi cần chuẩn chính xác hơn, người ta dùng phân tích phần tử hữu hạn hoặc tính toán chất lưu. Trường hợp dùng trực tiếp giá trị Rosenthal làm đường chuẩn cho giám sát thời gian thực hiện vẫn còn hiếm. Ở đây, nó được nêu như một ví dụ cho ý tưởng đặt ra một chuẩn vật lý.

Trí tuệ nhân tạo có thể nhận các sai lệch như vậy làm đầu vào. Nó học chênh lệch giữa trường nhiệt độ thực tế và chuẩn vật lý để dự đoán khuyết tật. Khi có chuẩn vật lý, gánh nặng mà trí tuệ nhân tạo phải học sẽ nhẹ hơn. Dù dữ liệu ít, vật lý vẫn làm người dẫn đường. Chính chuẩn vật lý giúp giảm gánh nặng học tập của trí tuệ nhân tạo.

Vì sao thời gian thực lại quan trọng đến vậy. Laser đi qua với tốc độ vài mét mỗi giây. Thời gian để tạo một lớp rất ngắn. Vị trí có khuyết tật sẽ sớm bị lớp tiếp theo phủ lên. Phải nhận ra trước khi bị phủ thì mới có thể sửa vị trí đó. Vì vậy, phán đoán phải nhanh ở cấp mili giây.

Thiết bị quan sát cũng liên tục phát triển. Nghiên cứu gần đây đưa ra phương pháp dùng một camera để thu đồng thời hai bước sóng. Phương pháp này ghi lại biến đổi nhiệt độ của vùng nóng chảy tới tối đa 100.000 khung hình mỗi giây. Ảnh nhiệt chính xác thu được như vậy trở thành nền tảng để phân biệt lỗ rỗng theo nhiều cách. Phải có con mắt chính xác như vậy mới có thể phân biệt lỗ rỗng đúng mức.

### 6.2.3 Nguyên lý thiết kế mô hình hình ảnh dựa trên thông tin vật lý

Bây giờ đến lượt đọc các hình ảnh đã quan sát. Mục này đề xuất nguyên lý thiết kế mô hình hình ảnh có đưa vật lý vào. Trước hết, ta xem xét mạng nơ-ron hình ảnh được dùng rộng rãi. Vision Transformer là mạng nơ-ron chia ảnh thành các ô nhỏ rồi xem xét quan hệ giữa chúng. Nó cắt ảnh như một tấm chăn ghép mảnh và xem từng mảnh nối với nhau như thế nào. Trong hình ảnh vùng nóng chảy, phần sáng ở trung tâm, chiếc đuôi kéo về phía sau và ánh sáng của bột bắn tóe đều khác nhau. Mô hình đọc các quan hệ này để phân biệt khuyết tật.

Sức mạnh của mạng nơ-ron này nằm ở việc nối các phần ở xa nhau. Nhiệt độ phía trước của vùng nóng chảy và chiếc đuôi phía sau có liên hệ với nhau. Chỉ nhìn một mảnh thì không biết nhiệt chảy như thế nào. Vision Transformer so sánh tất cả các mảnh cùng lúc. Vì vậy, nó đọc được dòng nhiệt của toàn bộ vùng. Tầm nhìn rộng này giúp bắt dấu hiệu báo trước của khuyết tật.

Cuốn sách này đề xuất đưa thêm định luật dẫn nhiệt vào đây. Ta khiến trí tuệ nhân tạo tự kiểm tra xem câu trả lời của nó có phù hợp với vật lý hay không. Hàm mất mát (Loss function) là điểm số đo trí tuệ nhân tạo sai bao nhiêu. Ý tưởng là đưa đồng thời sai số phân loại khuyết tật, sai số trường nhiệt độ và phần dư vật lý vào điểm số này. Phần dư là giá trị chỉ phương trình sai lệch bao nhiêu.

Đưa vật lý vào như vậy có lợi ích rõ ràng. Ảnh camera luôn lẫn nhiễu. Bột bắn tóe cũng có thể che ánh sáng. Mô hình chỉ học từ dữ liệu thuần túy dễ bị nhiễu như vậy làm dao động. Mô hình học cùng với vật lý có chỗ dựa là định luật dẫn nhiệt. Dù hình ảnh hơi mờ, nó vẫn giữ được phán đoán phù hợp với vật lý.

Nghiên cứu thực tế cũng củng cố sức mạnh của đặc trưng vật lý. Một nghiên cứu đã rút ra các đặc trưng dựa trên vật lý từ tín hiệu camera hồng ngoại. Bằng các đặc trưng này, nghiên cứu phát hiện lỗ rỗng

trong linh kiện Inconel 718. Bởi vì chính hình dạng biến đổi của nhiệt độ chứa manh mối về khuyết tật. Đặc trưng rút ra từ vật lý chống nhiễu tốt và có ý nghĩa rõ ràng. Vì vậy, cũng dễ giải thích căn cứ của phán đoán.

Tốc độ cũng quan trọng. Muốn phán đoán trong khi chế tạo, câu trả lời phải xuất hiện trong chớp mắt. Một nghiên cứu năm 2026 đã kết hợp mạng nơ-ron học sâu với bộ phân loại truyền thống. Phương pháp này phân biệt hình ảnh vũng nóng chảy của hợp kim nickel với độ chính xác khoảng 0,95. Thời gian phán đoán một khung hình là 1,15 mili giây. Nghiên cứu này vẫn là nghiên cứu ban đầu, chưa qua bình duyệt. Dù vậy, nó cho thấy phán đoán đủ nhanh để dùng tại hiện trường nhà máy là có thể đạt được.

Cách đưa vật lý vào đi kèm với điều chỉnh trọng số. Nếu ép phần dư vật lý quá mạnh, mô hình học dữ liệu kém hơn. Nếu chỉ bám theo dữ liệu quá mức, mô hình rời xa vật lý. Phải cân bằng tốt hai lực này thì mới có mô hình tốt. Việc tìm điểm cân bằng này là phần tinh chỉnh của học dựa trên thông tin vật lý. Trọng số phù hợp khác nhau tùy linh kiện và thiết bị.

Ở đây cần tách điều đã được kiểm chứng và điều được đề xuất. Việc tìm lỗ rỗng bằng đặc trưng hồng ngoại và mô hình lai tốc độ cao là kết quả mà nghiên cứu thực tế đã cho thấy. Cấu trúc đưa phần dư dẫn nhiệt vào Vision Transformer là hướng đi mà cuốn sách này phác thảo. Đây chưa phải phương pháp đã được kiểm chứng trên linh kiện thực tế. Phán đoán do trí tuệ nhân tạo đưa ra phải được xác nhận lại bằng mặt cắt đã cắt ra và chụp cắt lớp vi tính.

## 6.3 Giảm biến dạng của vòi phun lớn

Váy vòi phun là linh kiện lớn hình loa gắn ở phía sau động cơ. Nó được tạo bằng cách đập tấm mỏng thành dạng tròn như hình nón. Linh kiện lớn và mỏng như vậy dễ bị biến dạng trong khi chế tạo. Mục này bàn về nguyên nhân của biến dạng đó và cách ngăn chặn.

Trước hết, ta xem vì sao lực tích tụ bên trong linh kiện. Sau đó, ta giới thiệu mô hình trí tuệ nhân tạo nhanh thay cho phép tính chậm. Cuối cùng, ta xem cách dùng mô hình đó để điều chỉnh chuyển động của laser và giảm biến dạng. Linh kiện càng lớn thì quản lý biến dạng này càng khó và càng quan trọng.

### 6.3.1 Tích tụ ứng suất nhiệt và biến dạng mất ổn định khi đập lớp tấm mỏng cỡ lớn

Ứng suất dư (Residual stress) là lực kéo và lực nén còn lại bên trong linh kiện. Mắt không nhìn thấy, nhưng nó luôn tác động từ bên trong. Hãy vẽ ra cách lực này hình thành. Khi laser nung nóng một đường kim loại, phần đó giãn ra. Vì kim loại được nung nóng sẽ phồng lên.

Khi laser đi qua, đường đó lại nguội đi. Khi nguội, nó muốn co về như ban đầu. Nhưng lớp bên dưới đã đông cứng nên giữ chuyển động này lại. Lực muốn co và lực giữ lại đối kháng nhau, khiến ứng suất còn lại bên trong. Lực kéo tích tụ ở phía trên, lực nén tích tụ ở phía dưới.

Trong váy vòi phun mỏng và lớn, lực này gây rắc rối lớn. Nó giống như khi kéo một bên tờ giấy thì giấy nhăn và cong lại. Linh kiện biến dạng theo hướng bán kính hoặc tấm mỏng đột ngột cong. Sự cong đột ngột này gọi là mất ổn định (Buckling). Nếu nghiêm trọng, linh kiện còn có thể tách khỏi bệ đỡ.

Linh kiện càng lớn thì quản lý càng khó. Vì khi số lớp lên tới hàng trăm, nhiệt liên tục tích tụ. Trạng thái nhiệt khi đập vài lớp đầu khác với khi đập vài lớp về sau. Dù đặt cùng điều kiện laser, kết quả vẫn khác. Vì vậy, với linh kiện lớn, phải quan sát trạng thái nhiệt và thay đổi điều kiện theo đó.

Với hợp kim niobium như C-103, vấn đề này còn khó hơn. Vì điểm nóng chảy cao nên phải đưa vào lượng nhiệt lớn mới làm nó nóng chảy. Nhiệt lớn dẫn tới ứng suất lớn. Khi chế tạo vòi phun cỡ lớn bằng kim loại này, quản lý biến dạng quyết định thành bại. Phải thiết kế phân tích nhiệt và đường đi ngay từ đầu cùng nhau.

Biến dạng rất khó xử lý sau khi đã chế tạo xong. Nếu cố ép thẳng linh kiện đã đông cứng, vết nứt mới sẽ xuất hiện. Vì vậy, ngăn ứng suất tích tụ trong khi chế tạo là cách tốt hơn. Điều này giống như phòng bệnh trước thì tốt hơn chữa bệnh sau. Phần phòng ngừa này là nhiệm vụ của dự đoán bằng trí tuệ nhân tạo và điều khiển chủ động. Mô hình thay thế sẽ trình bày sau giúp thực hiện việc phòng ngừa đó.

### 6.3.2 Mô hình thay thế dựa trên nhiệt động lực học (Surrogate Model)

Muốn biết trước ứng suất bên trong linh kiện thì cần tính toán. Phân tích phần tử hữu hạn (Finite Element Analysis) là phương pháp chia nhỏ linh kiện để tính nhiệt độ và lực. Nó rất chính xác, nhưng với linh kiện lớn thì mất nhiều thời gian. Vì phải giải từng mảnh nhỏ trong hàng triệu mảnh. Một lần phân tích vòi phun cỡ lớn đôi khi mất vài ngày.

Mô hình thay thế cho phép tính chậm này là mô hình thay thế (Surrogate model). Mô hình thay thế là mô hình nhanh đã học trước kết quả của phép tính chậm. Mạng nơ-ron đa lớp (Multilayer Perceptron; MLP) thường được dùng. Đây là mạng nơ-ron cơ bản biến đầu vào thành đầu ra qua nhiều lớp. Sau khi học nhiều kết quả phân tích chính xác, nó ước lượng kết quả của điều kiện mới.

Mô hình này nhận gì và đưa ra gì. Đầu vào là công suất laser, tốc độ di chuyển, đường đi mà laser đã quét qua, và thời gian chờ giữa các lớp. Đầu ra là phân bố ứng suất dư trong toàn bộ chi tiết và hình dạng biến dạng cuối cùng. Việc phân tích chính xác vốn mất vài ngày được mô hình này ước lượng trong thời gian ngắn. Nhờ vậy ta có thể nhanh chóng so sánh nhiều điều kiện với nhau.

Tốc độ mở ra con đường mới. Nếu mỗi lần đổi một điều kiện lại phải chờ vài ngày thì rất khó tìm được thiết kế tốt. Nếu ước lượng nhanh, dù chỉ là gần đúng, ta có thể sàng lọc trước hàng nghìn điều kiện. Chỉ chọn các ứng viên tốt rồi kiểm tra lại bằng phân tích chính xác. Mô hình thay thế đóng vai trò như một tấm lưới thu hẹp một vùng ứng viên rộng.

Để mô hình thay thế học tốt, cần có nhiều dữ liệu tốt. Ta chạy trước phân tích chính xác với nhiều điều kiện để gom dữ liệu. Các giá trị đo được khi chế tạo chi tiết thật cũng được đưa vào cùng. Dữ liệu càng rộng và càng đều thì ước lượng càng chính xác. Với những điều kiện lạ mà mô hình chưa từng học, ước lượng sẽ dao động. Vì vậy mô hình thay thế chỉ được tin dùng trong phạm vi đã học.

Các nghiên cứu gần đây đã đạt kết quả theo hướng này. Một nghiên cứu năm 2025 đã dùng học máy để ước lượng nhanh ứng suất dư và giảm cong vênh. Nghiên cứu này báo cáo rằng độ uốn của mẫu thử dạng cầu đã giảm rõ rệt. Sai khác giữa ứng suất dự đoán và phân tích chính xác cũng nhỏ. Điều đó có nghĩa là mô hình nhanh đã tiến gần đến phân tích chính xác.

Khi đưa thông tin vật lý vào, kết quả còn tốt hơn. Một nghiên cứu khác năm 2025 đã đưa các định luật vật lý vào mạng nơ-ron để dự đoán độ cong vênh của chi tiết lớn gần với thời gian thực. Nghiên cứu dùng cấu trúc gọi là toán tử vật lý (Neural Operator) để học sao cho phù hợp rộng rãi với nhiều điều kiện. Nghiên cứu này vẫn là nghiên cứu giai đoạn đầu, chưa qua bình duyệt. Dù vậy, nó cho thấy một con đường để có được cả dự đoán nhanh lẫn tính nhất quán vật lý.

### 6.3.3 Tối ưu hóa chủ động quỹ đạo quét và thời gian làm nguội giữa các lớp

Giờ là lúc thay đổi quy trình thật dựa trên dự đoán. Quỹ đạo quét (Scan path) là hướng và thứ tự di chuyển của laser. Nếu cứ cho laser đi mãi theo một hướng, nhiệt sẽ dồn về một phía. Lượng nhiệt bị dồn đó để lại ứng suất lớn. Vì vậy ta đổi hướng để nhiệt lan đều.

Đã có nhiều quỹ đạo thông minh được đưa ra. Có cách chia một mặt rộng thành các vùng nhỏ rồi lần lượt lấp đầy xen kẽ. Cũng có cách xoay nhẹ hướng laser ở từng vùng. Người ta còn dùng các đường cong đi qua không gian thật dày. Mục tiêu chỉ có một. Đó là phân tán dòng nhiệt đều ra mọi phía để ngăn nhiệt dồn về một bên.

Thời gian chờ giữa các lớp cũng quan trọng. Nếu lớp trước còn nóng mà đã đắp lớp tiếp theo, nhiệt sẽ chồng chất từng lớp. Nhiệt chồng lên nhau gây cong vênh lớn. Vì vậy ta chờ cho lớp trước nguội đến một mức nào đó. Khoảng chờ này gọi là thời gian làm nguội giữa các lớp.

Tối ưu hóa chủ động (Active optimization) là phương pháp thay đổi các điều kiện này ngay trong quá trình gia công. Pyrometry cho biết chi tiết hiện nóng đến mức nào. Mô hình thay thế nhanh chóng dự đoán nguy cơ ứng suất của đường quét tiếp theo. Dựa vào hai thông tin đó, ta điều chỉnh công suất laser, tốc độ và thời gian chờ theo từng thời điểm. Đây là quá trình chuyển kinh nghiệm lâu năm của con người thành con số và quy tắc.

Cách điều chỉnh cũng trở nên thông minh hơn. Tối ưu hóa Bayes (Bayesian Optimization) là phương pháp tìm điều kiện tốt với ít lần thử. Dựa trên các kết quả đã thử, nó chọn khôn ngoan điều kiện cần thử tiếp theo. Thuật toán di truyền (Genetic Algorithm) là phương pháp trộn các điều kiện tốt với nhau để tìm điều kiện tốt hơn. Bằng các phương pháp này, thời gian chờ được thay đổi cho phù hợp với trạng thái nhiệt của lớp trước.

Việc điều chỉnh này diễn ra khác nhau theo từng lớp. Lớp dưới mất nhiệt vào để đỡ nên nguội nhanh. Càng lên lớp trên, nhiệt đôi khi càng khó thoát. Khi đó, ta tăng thời gian chờ ở lớp trên. Tuy nhiên đây chỉ là một ví dụ. Mức độ tích tụ nhiệt khác nhau tùy theo hình dạng chi tiết, để đỡ, diện tích lớp và thứ tự đắp. Vì vậy cảm biến được dùng để xem trạng thái nhiệt ở từng lớp rồi quyết định thời gian chờ.

Nếu cách này vận hành tốt, độ cong vênh giảm rõ rệt. Ứng suất phân bố đều hơn và ứng suất cực đại thấp hơn. Chi tiết nằm trong dung sai cho phép. Tuy nhiên việc điều chỉnh này luôn phải nằm trong vùng quy trình an toàn. Cần xử lý cùng lúc độ trễ của cảm biến, sai số của mô hình và giới hạn phản ứng của thiết bị. Chỉ sau khi kiểm tra trước bằng mẫu thử mới có thể áp dụng cho chi tiết thật.

Ý tưởng bản sao số cũng lớn lên từ đây. Bản sao số là đặt một cặp song sinh của chi tiết thật bên trong máy tính. Các giá trị cảm biến của quy trình thật liên tục được đưa vào cặp song sinh đó. Cặp song sinh phản chiếu những gì đang xảy ra bên trong chi tiết ở thời điểm hiện tại. Nó cũng cho thấy trước điều gì có thể nguy hiểm ở lớp tiếp theo. Mô hình thay thế là bộ não giúp cặp song sinh này suy nghĩ nhanh.

## 6.4 Điều khiển hình dạng của tinh thể

Khi kim loại đông đặc, tinh thể lớn lên bên trong. Hình dạng của các tinh thể này quyết định độ bền của chi tiết. Nếu tinh thể chỉ mọc dài theo một hướng, độ bền sẽ khác nhau theo từng hướng. Mục này bàn về cách điều khiển hình dạng tinh thể theo ý muốn.

Trước hết ta xem hai hình dạng tinh thể và sự khác nhau giữa chúng. Sau đó giới thiệu mô hình sinh tạo vẽ ra hình dạng tinh thể từ điều kiện quy trình. Cuối cùng, ta xem điều khiển bằng học tăng cường để tự điều khiển hình dạng tinh thể trong khi chế tạo. Thế giới tinh thể không nhìn thấy bằng mắt thường quyết định số phận của chi tiết.

### 6.4.1 Chuyển tiếp giữa tinh thể cột và tinh thể đẳng trục (CET) và dị hướng cơ học

Khi kim loại đông đặc, tinh thể lớn lên theo hai dạng. Tinh thể cột (Columnar grain) là tinh thể mọc dài theo một hướng như cây cột. Nó gần giống hình ảnh nhiều cột gỗ dựng cạnh nhau. Tinh thể đẳng trục (Equiaxed grain) là tinh thể phát triển với kích thước gần giống nhau theo mọi phía. Nó gần giống hình ảnh rải đều sỏi trên mặt đất.

Sự khác biệt này chia tách tính chất của chi tiết. Khi có nhiều tinh thể cột, chi tiết mạnh theo hướng cột nhưng yếu theo phương ngang. Độ bền và tuổi thọ mỗi thay đổi theo hướng. Tính chất này gọi là dị hướng (Anisotropy). Nguyên lý giống như gỗ dễ tách theo thớ nhưng khó tách theo chiều ngang.

Trong chi tiết tên lửa, dị hướng đôi khi trở thành mối lo. Lý do là chi tiết chịu lực từ nhiều hướng. Nếu chỉ mạnh ở một hướng, hướng yếu sẽ nứt trước. Tinh thể đẳng trục làm giảm vấn đề này. Nó bền đều theo mọi phía và cũng chống nứt tốt khi đông đặc. Vì vậy trong các chi tiết chịu lực từ nhiều hướng, tinh thể đẳng trục có lợi.

Tuy nhiên tinh thể đẳng trục không phải lúc nào cũng tốt. Với chi tiết gặp vấn đề rão hoặc mỏi nhiệt, tinh thể dạng cột đôi khi tốt hơn. Rão là hiện tượng vật liệu bị kéo dài chậm khi chịu lực trong thời gian dài. Trong chi tiết mà nhiệt chỉ chảy theo một hướng, tinh thể có định hướng cũng có thể có lợi. Hình dạng tinh thể mong muốn thay đổi theo lực mà chi tiết phải chịu.

Trong chế tạo bồi đắp, tinh thể cột dễ hình thành. Đó là vì nhiệt thoát theo hướng đắp của laser. Tinh thể mọc như cột theo hướng thoát nhiệt. Kết quả là tinh thể xếp thành hàng theo hướng đắp. Nhóm tinh thể có hướng dồn về một phía như vậy được gọi là kết cấu tinh thể. Mục tiêu của mục này là thay đổi sự thiên lệch đó theo nhu cầu.

Hình dạng tinh thể phụ thuộc lớn vào hai giá trị. Gradient nhiệt độ (Temperature gradient;  $G$ ) chỉ mức độ dốc của nhiệt độ. Nó biểu thị nhiệt độ thay đổi nhanh đến đâu theo vị trí. Tốc độ đông đặc (Solidification rate;  $R$ ) là tốc độ di chuyển của ranh giới giữa lỏng và rắn. Hai giá trị này làm những việc khác nhau.

Tỷ số  $G$  chia cho  $R$  quyết định hình dạng tinh thể. Tỷ số này lớn thì nghiêng về dạng cột, nhỏ thì nghiêng về dạng đẳng trục. Tích  $G$  nhân  $R$  trở thành tốc độ làm nguội. Giá trị này càng lớn thì tinh thể càng bị chia nhỏ hơn. Vì vậy có thể điều khiển riêng hình dạng và kích thước. Khi thay đổi công suất và tốc độ laser, hai giá trị này cùng thay đổi.

Không chỉ hai giá trị này quyết định hình dạng. Các hạt nhỏ làm mầm, dòng chảy của kim loại nóng chảy, và mức độ nóng chảy lại cũng ảnh hưởng. Hướng của tinh thể đã mọc ở bên dưới cũng được kế thừa. Vì vậy gradient nhiệt độ càng thấp và đông đặc càng nhanh thì tinh thể đẳng trục càng dễ hình thành, nhưng không phải lúc nào cũng như vậy. Phải xem nhiều giá trị cùng lúc mới điều khiển đúng hình dạng tinh thể.

Có một điểm mà tinh thể cột chuyển sang tinh thể đẳng trục. Sự thay đổi này gọi là chuyển tiếp từ tinh thể cột sang tinh thể đẳng trục (Columnar-to-Equiaxed Transition; CET). Khi gradient nhiệt độ và tốc độ đông đặc vượt qua một điều kiện nào đó, sự thay đổi này xảy ra. Nếu biết điều kiện đó, ta có thể tiến gần đến hình dạng tinh thể mong muốn. Các lý thuyết cũ đã vẽ ranh giới này bằng công thức.

Các nghiên cứu gần đây đang tinh chỉnh ranh giới này chính xác hơn. Một nghiên cứu năm 2025 đã lập tiêu chí dự đoán CET dựa trên thành phần hợp kim. Nghiên cứu dùng học máy để làm rõ nguyên tố nào ảnh hưởng lớn đến hình dạng tinh thể. Điều đó có nghĩa là chỉ cần thay đổi nhẹ thành phần cũng có thể đổi hình dạng tinh thể. Nếu xử lý đồng thời quy trình và thành phần, phạm vi điều khiển tinh thể sẽ rộng hơn.

Học máy xem xét thành phần cũng tiếp tục phát triển. Nó phân biệt bằng dữ liệu nguyên tố nào ảnh hưởng lớn đến hình dạng tinh thể. Có kết quả cho thấy lượng nguyên tố như silicon hoặc carbon làm thay đổi hình dạng tinh thể. Những nghiên cứu như vậy đang lấp dần bản đồ điều khiển tinh thể.

#### 6.4.2 Mô hình sinh tạo vẽ hình dạng tinh thể từ điều kiện quy trình

Ta muốn nhìn trước hình dạng tinh thể từ điều kiện quy trình. Muốn vậy cần một công cụ tạo hình ảnh vi cấu trúc khi nhập điều kiện. Một công cụ như vậy là mạng nơ-ron sinh tạo có điều kiện (Conditional GAN). Mạng nơ-ron này là mô hình sinh tạo tạo ra hình ảnh mong muốn khi đưa điều kiện vào. Ở đây điều kiện là các giá trị như công suất laser, tốc độ, gradient nhiệt độ và tốc độ đông đặc. Việc dùng mô hình này cho điều khiển vi cấu trúc vẫn đang ở giai đoạn đề xuất.

Mô hình này học bằng cách để hai phần cạnh tranh với nhau. Bộ sinh (Generator) tạo ra hình ảnh vi cấu trúc. Bộ phân biệt (Discriminator) phân biệt hình đó là thật hay giả. Bộ sinh cố đánh lừa bộ phân biệt, còn bộ phân biệt cố không bị lừa. Khi lặp lại cuộc cạnh tranh này, bộ sinh dần vẽ ra hình ảnh giống thật hơn. Nó gần giống cảnh kẻ làm giả và người thẩm định cùng nâng cao tay nghề của nhau.

Ở đây hình ảnh thật là ảnh nhiễu xạ tán xạ ngược điện tử (EBSD). Phương pháp này đọc hướng tinh thể bằng kính hiển vi điện tử. Nó tô màu hướng của từng tinh thể và cho thấy như một tấm bản đồ. Bộ sinh học cách bắt chước bản đồ này. Khi đưa điều kiện vào, nó vẽ bản đồ tinh thể có thể xuất hiện trong điều kiện đó.

Nếu mô hình này học tốt, nó có ích. Khi nhà thiết kế thay đổi điều kiện quy trình, họ có thể nhìn trước tinh thể sẽ thay đổi ra sao. Họ có thể ước đoán mà không cần thật sự chế tạo chi tiết rồi cắt ra xem. Thời gian và vật liệu được tiết kiệm nhiều. Nhiều điều kiện được so sánh trên màn hình để chọn phía tốt hơn.

Để thật sự có được bản đồ tinh thể, cần rất nhiều công sức. Phải cắt chi tiết và mài bề mặt thật mịn. Phải dùng kính hiển vi điện tử để đọc hướng tinh thể từng vị trí một. Nếu muốn đọc hết một diện tích rộng, sẽ mất nhiều thời gian. Mô hình sinh tạo trở thành lối tắt rút ngắn quá trình chậm này. Tất nhiên lối tắt đó đôi khi phải được kiểm tra lại bằng đo đạc thật.

Cũng có điểm cần thận trọng. Mô hình sinh tạo rất giỏi tạo ra hình ảnh trông có vẻ hợp lý. Nhưng vẻ hợp lý không đồng nghĩa với sự thật vật lý. Hình ảnh có đẹp cũng có thể khác thực tế. Vì vậy cần kiểm tra cùng các chỉ số vật lý như kích thước hạt tinh thể, phân bố hướng và tỷ lệ pha. Chỉ khi các chỉ số này khớp với thực tế mới có thể tin hình ảnh.

Nếu đưa vật lý vào hàm mất mát, độ tin cậy tăng lên. Khi kích thước tinh thể trong hình được sinh ra lệch khỏi phân bố thực tế, ta phạt mô hình. Làm như vậy dẫn bộ sinh đến việc vẽ hình phù hợp với vật lý. Nó thu hẹp khoảng cách giữa hình đẹp và hình đúng. Mô hình sinh tạo cũng phải được đo bằng thước vật lý thì mới có thể dùng cho chi tiết tên lửa.

#### 6.4.3 Học tăng cường điều khiển hình dạng tinh thể trong khi chế tạo

Giờ là lúc để hệ thống tự điều khiển hình dạng tinh thể trong khi chế tạo. Học tăng cường (Reinforcement Learning) là phương pháp học hành động tốt thông qua thử nghiệm và phần thưởng. Nó giống quá trình em bé học đi bằng cách té ngã. Làm tốt thì thưởng, làm không tốt thì phạt. Các hành động nhận được nhiều phần thưởng dần tăng lên.

Trong chế tạo bồi đắp, hành động là điều chỉnh laser. Tăng hay giảm công suất laser bao nhiêu, thay đổi tốc độ bao nhiêu chính là hành động. Tuy nhiên không thể nhìn thấy ngay hình dạng tinh thể trong khi chế tạo. Vì vậy ta nhìn nhiệt độ và hình dạng của vũng nóng chảy thay thế. Từ các giá trị này, ta suy đoán hình dạng tinh thể. Pyrometry cho biết trạng thái này theo thời gian thực.

Phần thưởng được quyết định theo mức độ gần với mục tiêu. Càng gần hình dạng tinh thể mong muốn thì phần thưởng càng lớn. Khi tỷ lệ tinh thể đẳng trục đạt mục tiêu, phần thưởng tăng. Ngược lại, nếu xuất hiện keyhole hoặc lỗi thiếu liên kết nóng chảy, hệ thống bị phạt. Trí tuệ nhân tạo lần theo phần thưởng này để học cách điều chỉnh laser.

Cách thiết kế phần thưởng quyết định thành bại. Nếu thiết kế phần thưởng quá sơ sài, hệ thống sẽ đi chệch sang hướng không mong muốn. Nó cũng có thể chỉ đuổi theo hình dạng tinh thể mà lãng phí năng lượng. Vì vậy, phần thưởng phải chứa đồng thời hình dạng tinh thể, khuyết tật và năng lượng. Tức là đặt nhiều mục tiêu lên bàn cân và giữ cân bằng. Thiết kế tốt chiếc cân này là trọng tâm của thiết kế học tăng cường.

Vì sao cần điều khiển thời gian thực như vậy. Vì hình dạng của chi tiết khác nhau ở từng vị trí. Góc cạnh, vách mỏng và phần nhô ra có tốc độ thoát nhiệt khác nhau. Dù đặt cùng điều kiện laser, hình dạng tinh

thể vẫn thay đổi theo từng nơi. Điều khiển thời gian thực cố gắng thay đổi điều kiện theo từng vị trí để giữ hình dạng tinh thể đồng đều.

Hướng này vẫn chưa phải là tiêu chuẩn công nghiệp. Hiện có nhiều đề xuất ở mức mô phỏng và mẫu thử. Các trường hợp điều khiển hình dạng tinh thể bằng vòng kín theo thời gian thực còn hiếm. Lý do là khó đo trực tiếp hình dạng tinh thể ngay trong quá trình. Độ trễ của cảm biến, sai số của mô hình và giới hạn phản ứng của thiết bị cũng đều phải vượt qua.

Học tăng cường mạnh ở chỗ có thể trải qua trước các thử nghiệm nguy hiểm. Nếu thử điều kiện xấu trên thiết bị thật, chi tiết sẽ bị loại bỏ. Trong thí nghiệm mô phỏng trên máy tính, thất bại bao nhiêu cũng được. Hệ thống học trước cách điều chỉnh tốt từ vô số thất bại. Sau khi học xong, nó được chuyển cần thận sang thiết bị thật trong phạm vi an toàn. Có thể xem như máy tính trả thay cái giá của thất bại.

Vì vậy, thứ tự rất quan trọng. Trước hết học điều khiển trong mô phỏng. Sau đó xác nhận trên mẫu thử có giới hạn. Tiếp theo dùng nó như công cụ hỗ trợ trong phạm vi quy trình đã được kiểm chứng. Phải đi qua các bước này mới có thể đưa điều khiển học tăng cường vào chi tiết thật một cách thận trọng. Sau khi chế tạo, hình dạng tinh thể được kiểm tra lại bằng EBSD để xem điều khiển có đúng hay không.

## 6.5 Giám sát khuyết tật bằng cách nghe bên trong qua âm thanh

Cho đến đây, chúng ta đã tìm khuyết tật bằng ánh sáng. Nhưng ánh sáng chỉ nhìn thấy bề mặt. Các vết nứt sinh ra sâu bên trong chi tiết khó được camera bắt được. Mục này trình bày phương pháp nghe bên trong bằng âm thanh.

Trước hết xem khuyết tật phát ra âm thanh như thế nào. Sau đó xem cách biến âm thanh ấy thành hình ảnh. Cuối cùng xem cách dùng chung âm thanh và ánh sáng để phán đoán tốt hơn. Khi kết hợp các giác quan khác nhau, ta có thể bù vào phần mà một giác quan bỏ sót.

### 6.5.1 Cơ chế vật lý của tín hiệu phát xạ âm bằng rộng (AE)

Phát xạ âm (Acoustic Emission; AE) là âm thanh phát ra khi một biến đổi đột ngột xảy ra trong vật liệu. Chính xác hơn, đó là một sóng đàn hồi rất nhỏ. Nguyên lý giống tiếng tách khi ly thủy tinh bị nứt. Nó cũng giống tiếng răng rắc khi gỗ bị uốn rồi gãy. Sự kiện bên trong vật liệu rò rỉ ra ngoài dưới dạng âm thanh.

Trong chế tạo bồi đắp, nhiều sự kiện khác nhau tạo ra các âm thanh khác nhau. Nứt nóng phóng ra năng lượng lớn trong khoảnh khắc. Vì vậy, nó tạo ra âm thanh sắc ở tần số cao. Nó gần với một tiếng bật ngắn và khô. Âm thanh này có biên độ lớn và vọt lên đột ngột.

Khi keyhole sụp xuống cũng có âm thanh. Khoang rỗng trong vùng nóng chảy bất ngờ sụp xuống và tạo va đập. Âm thanh này dao động liên tiếp ở tần số thấp hơn một chút so với vết nứt. Nó giống dao động khi bọt khí vỡ trong nước. Kết cấu âm thanh khác với nứt.

Khi bột bắn tóe hoặc thiết bị cọ quẹt cũng có âm thanh. Âm thanh này nằm ở tần số thấp hơn. Tiếng ồn bình thường không phải khuyết tật cũng trộn vào tín hiệu âm thanh. Vì vậy phải nghe và lọc âm thật kỹ. Trọng tâm là phân biệt âm thanh của khuyết tật với tiếng ồn thông thường.

Mỗi âm thanh có kết cấu khác nhau là có lý do. Khi năng lượng lớn được giải phóng đột ngột, thành phần tần số cao xuất hiện. Khi giải phóng chậm, thành phần tần số thấp xuất hiện. Vết nứt mở ra trong chớp mắt nên tạo âm cao. Bọt bắn tóe tương đối chậm nên tạo âm thấp. Tuy nhiên, tần số thực tế thu được thay đổi theo cảm biến, đường truyền và cách lọc. Vì vậy, trật tự này nên được hiểu là xu hướng quan sát trong một thử nghiệm cụ thể.

Phương pháp nghe bằng âm thanh có một ưu điểm lớn. Ánh sáng chỉ nhìn bề mặt, nhưng âm thanh đi xuyên qua bên trong. Ngay cả vết nứt sâu trong khối bồi đắp cũng có thể được bắt bằng âm thanh. Nếu nhận ra khoảnh khắc vết nứt vừa sinh ra, ta có thể phản ứng nhanh. Quy trình có thể dừng ngay tại chỗ, đánh dấu vùng đó rồi kiểm tra lại. Cũng có thể nghĩ đến sửa chữa bằng cách bắn lại laser để nung chảy. Nhưng tái nóng chảy không phải lúc nào cũng loại bỏ được vết nứt. Nó cũng có thể gây ứng suất nhiệt mới và vết nứt mới. Cần nhìn quy trình sửa chữa đã được kiểm chứng khác với việc chỉ nung chảy lại.

Có hai cách nghe âm thanh. Một là nghe bằng cảm biến áp điện tiếp xúc với chi tiết. Cảm biến áp điện là phần tử tạo ra điện khi chịu lực. Nó được gắn vào tấm đỡ chi tiết để nghe sóng đàn hồi đi qua bên trong. Cách còn lại là nghe âm thanh trong không khí bằng micro. Hai cách này khác nhau về đường truyền âm và tần số.

### 6.5.2 Biến đổi wavelet thời gian-tần số

Nếu nhìn tín hiệu âm thanh nguyên dạng, rất khó phán đoán. Chỉ dựa vào dao động theo thời gian thì khó biết đó là khuyết tật gì. Lý do là nhiều âm ở các cao độ khác nhau chồng lên nhau trong cùng một khoảnh khắc. Phải xem đồng thời thời điểm và cao độ của âm thanh. Khi đó mới đọc được dấu vân tay của khuyết tật.

Biến đổi wavelet (Wavelet Transform) làm việc này. Đây là phương pháp vẽ thời gian và tần số cùng lúc. Nó biến tín hiệu âm thanh thành một hình có trục ngang là thời gian, trục dọc là cao độ. Hình này được gọi là scalogram. Nó giống bản nhạc vì cùng cho thấy thời gian và cao độ.

Hình này tạo ra hoa văn khác nhau cho từng khuyết tật. Âm sắc của vết nứt hiện thành điểm sáng ở vùng cao. Dao động của keyhole hiện thành hoa văn kéo dài ở vùng giữa. Tiếng ồn trải mờ ở vùng thấp. Khi âm thanh trở thành hình ảnh như vậy, mắt người dễ so sánh hơn.

Khi đọc hoa văn, phải chọn tốt wavelet mẹ. Wavelet mẹ là dạng sóng ngắn dùng làm thước đo để so âm thanh. Phải dùng thước đo giống với kết cấu của âm thanh thì hoa văn mới rõ. Nếu dùng thước đo có kết cấu khác, hoa văn sẽ mờ. Chọn thước đo phù hợp cũng là một phần của tiền xử lý. Phải chọn thước đo phù hợp mới thu được hình ảnh có hoa văn rõ.

Trí tuệ nhân tạo học hoa văn từ hình ảnh này. Mạng nơ-ron đọc ảnh sẽ đọc scalogram. Nó học từ nhiều ví dụ để biết hoa văn nào tương ứng với khuyết tật nào. Sau khi học xong, nó nhìn hình mới và phân loại khuyết tật. Nhờ biến âm thanh thành hình ảnh, ta có thể mượn sức mạnh của trí tuệ nhân tạo thị giác.

Nghiên cứu gần đây đã đạt kết quả tốt bằng phương pháp này. Một nghiên cứu ước lượng lỗ keyhole chỉ bằng âm thanh trong không khí và tốc độ laser. Âm thanh thu bằng micro được chuyển thành scalogram để học. Giá trị biểu thị độ tin cậy của kết quả vượt 0.8. Đây là kết quả đã đăng trên tạp chí học thuật. Nghiên cứu cho thấy chỉ với âm thanh cũng có thể ước lượng khuyết tật bên trong ở thang mili giây.

Nghe âm thanh còn có ưu điểm là chi phí thấp. Camera tốc độ cao và thiết bị tia X chính xác đắt tiền và khó vận hành. Cảm biến áp điện tương đối rẻ và dễ gắn. Nếu gắn nhiều cảm biến, ta cũng có thể ước đoán âm thanh đến từ đâu. Vì vậy, giám sát dựa trên âm thanh đang được chú ý như một con đường thực dụng tại hiện trường.

Nguyên lý định vị bằng nhiều cảm biến khá dễ hiểu. Cùng một âm thanh đến từng cảm biến ở các thời điểm khác nhau. Nó đến cảm biến gần trước, cảm biến xa sau. Nếu so sánh chênh lệch thời gian đến, ta có thể ước lượng vị trí phát âm. Tức là dùng chênh lệch thời gian âm đến nhiều micro để xác định hướng. Vị trí tìm được như vậy cho biết khuyết tật nằm ở đâu.

### 6.5.3 Nhận biết khuyết tật bằng cách dùng chung ánh sáng và âm thanh

Đa phương thức (Multimodal) có nghĩa là dùng chung nhiều loại dữ liệu. Nó giống việc con người dùng cả mắt và tai. Nếu chỉ nghe âm thanh thì có thứ bị bỏ sót, nếu chỉ nhìn ánh sáng thì cũng có thứ bị bỏ sót. Khi kết hợp hai bên, chúng bù vào khoảng trống của nhau. Mục này trình bày cách kết hợp âm thanh và ánh sáng.

Ánh sáng và âm thanh nhìn thấy những thứ khác nhau. Pyrometry cho thấy tốt nhiệt độ bề mặt. Phát xạ âm bắt được các biến đổi đột ngột sinh ra bên trong. Bất thường nhiệt độ trên bề mặt được camera biết trước. Vết nứt bên trong được âm thanh biết trước. Khi kết hợp hai tín hiệu, ta quan sát cùng lúc bề mặt và bên trong.

Cách kết hợp là như sau. Tín hiệu âm thanh được mạng nơ-ron chuyên cho âm thanh đọc. Ảnh ánh sáng được mạng nơ-ron chuyên cho hình ảnh đọc. Các đặc trưng do hai mạng nơ-ron rút ra được gom lại để đưa ra phán đoán cuối cùng. Có thể xem như giao cho mỗi bên việc mình làm tốt, rồi hợp lực ở phần kết luận.

Nghiên cứu thực tế đã cho thấy sức mạnh của sự kết hợp này. Một nghiên cứu thu thập đồng thời pyrometry và âm thanh. Khi kết hợp hai tín hiệu, nó phát hiện lỗ keyhole tốt hơn. Số lỗ bị bỏ sót giảm so với khi chỉ dùng một tín hiệu. Bắt được cả vết nứt theo cách kết hợp như vậy là bài toán phía trước.

Khi kết hợp, quan trọng hơn cả là khớp thời gian. Phải khớp chính xác khoảng khắc âm thanh phát ra với khoảng khắc ánh sáng thay đổi. Nếu thời gian lệch nhau, các sự kiện khác nhau có thể bị xem là cùng một sự kiện. Vì vậy, hai cảm biến được đồng bộ cực kỳ chính xác. Do mỗi cảm biến có loại nhiễu khác nhau, cũng cần xử lý trước để làm sạch tín hiệu.

Có nhiều cách kết hợp hai tín hiệu. Có cách trộn tín hiệu ngay từ đầu rồi học chung. Cũng có cách để mỗi bên học riêng, sau đó chỉ kết hợp ở phần kết luận. Các nghiên cứu gần đây thường dùng cách kết hợp các đặc trưng đã học riêng bằng một bộ phân loại nhẹ. Cách này tính toán nhẹ và thích ứng tốt với khuyết tật mới. Cách nào tốt phụ thuộc vào chi tiết và thiết bị.

Dòng chảy kết hợp nhiều giác quan sẽ còn mở rộng hơn. Ngoài ánh sáng và âm thanh, người ta còn dùng tín hiệu phản xạ của laser. Ảnh của lớp đã tạo và bản đồ nhiệt độ cũng được thu thập cùng nhau. Nhiều mắt và tai quan sát một chi tiết theo nhiều lớp. Càng có nhiều giác quan, khuyết tật bị bỏ sót càng giảm. Việc nối nhiều tín hiệu ấy thành một là vai trò của trí tuệ nhân tạo.

Cách này cũng có giới hạn rõ ràng. Phán đoán của trí tuệ nhân tạo an toàn khi dùng cho cảnh báo trong sản xuất và chọn lọc ứng viên cần kiểm tra. Quyết định đạt thật sự phải dựa trên kiểm tra không phá hủy, quan sát mặt cắt và chụp cắt lớp vi tính. Lệch thời gian giữa các cảm biến và nhiễu cũng có thể gây phán đoán sai. Trí tuệ nhân tạo là người dẫn đường giúp con người thu hẹp nơi cần nhìn.

## Tổng kết

Chương này đã xem xét công nghệ chế tạo chi tiết tên lửa bằng cách bồi đắp kim loại. Chúng ta đã thấy phương pháp nung chảy bột bằng laser để bồi đắp và phương pháp nung chảy dây để bồi đắp. Chúng ta đã điểm qua tính chất và các điểm cần cẩn trọng của ba kim loại Inconel 718, GRCo-42 và C-103. Mỗi kim loại có nhiệm vụ khác nhau, và vấn đề cần kiểm soát khi chế tạo cũng khác nhau.

Chế tạo bồi đắp đem lại tự do, đồng thời cũng đem lại nguy cơ khuyết tật. Rỗ keyhole và thiếu hợp nhất trở thành điểm khởi đầu của nứt mỏi. Ứng suất dư làm các chi tiết lớn bị cong vênh. Sự thiên lệch hướng tinh thể làm độ bền nghiêng về một phía. Những khuyết tật này liên quan trực tiếp đến an toàn của động cơ.

Các khuyết tật này đan xen với nhau. Chỉ một điều kiện laser sai cũng có thể kéo theo nhiều khuyết tật cùng lúc. Khi cố tránh keyhole, ta có thể gây thiếu hợp nhất. Khi cố giảm ứng suất, hình dạng tinh thể có thể xấu đi. Vì vậy không được chỉ nhìn một khuyết tật rồi đổi điều kiện. Phải đặt nhiều khuyết tật lên cùng bàn cân để tìm điểm cân bằng.

Trí tuệ nhân tạo giúp vấn đề này theo ba hướng. Một là nhanh chóng tìm khuyết tật trong lúc chế tạo. Mô hình đưa định luật vật lý vào ảnh pyrometry sẽ phân biệt lỗ rỗng. Mô hình biến âm thanh thành hình ảnh rồi đọc sẽ bắt vết nứt bên trong. Khi kết hợp ánh sáng và âm thanh, ta quan sát cùng lúc bề mặt và bên trong.

Tiếp theo là dự đoán nhanh ứng suất nhiệt và biến dạng thay cho phép tính chậm. Mô hình thay thế rút ngắn phân tích vốn mất nhiều ngày. Với dự đoán nhanh này, ta sàng lọc trước rất nhiều điều kiện quy trình. Nếu đưa vật lý vào, dự đoán sẽ không rời khỏi vật lý.

Cuối cùng là hỗ trợ thay đổi điều kiện quy trình tùy từng thời điểm. Nó điều chỉnh quỹ đạo quét và thời gian làm mát để giảm cong vênh. Nó điều chỉnh laser để giữ hình dạng tinh thể đồng đều. Mô hình sinh tạo vẽ trước hình dạng tinh thể từ điều kiện. Học tăng cường tự kiểm soát điều kiện trong lúc chế tạo.

Cần nói rõ một điều. Trí tuệ nhân tạo không thể thay thế mọi phép kiểm tra. Các mô hình đã thấy trong chương này nhanh và thông minh, nhưng không hoàn hảo. Nhiều kết quả vẫn ở giai đoạn nghiên cứu ban đầu, và ít thứ đã trở thành tiêu chuẩn công nghiệp. Phán đoán của trí tuệ nhân tạo an toàn khi dùng cho cảnh báo và chọn lọc ứng viên cần kiểm tra. Bảo đảm chất lượng cuối cùng phải dựa cùng lúc vào mẫu thử, kiểm tra không phá hủy và thử nghiệm đốt.

Cũng có một dòng chảy mới dùng mô hình ngôn ngữ. Trong chế tạo bồi đắp, luận văn và báo cáo chất cao như núi. Mô hình ngôn ngữ lớn (Large Language Model; LLM) đọc các văn bản này và sắp xếp tri thức. Ta có thể hỏi trong điều kiện nào thì khuyết tật nào dễ sinh ra. Một nghiên cứu năm 2025 đã thử dự đoán điều kiện gây khuyết tật bằng mô hình ngôn ngữ. Những công cụ như vậy trở thành cây cầu nối kinh nghiệm của con người với tài liệu.

Bức tranh lớn là như sau. Cảm biến trở thành mắt và tai, mô hình thay thế trở thành đầu óc, bộ điều khiển trở thành bàn tay. Khi ba phần này nối thành một, nhà máy sẽ tự quan sát chi tiết trong lúc chế tạo. Việc NASA bồi đắp một vòi phun lớn và chịu được thử nghiệm đốt kéo dài là điểm khởi đầu. Chương sau sẽ xem xét cách hoàn thiện chi tiết đã chế tạo như vậy, và cách đánh giá liệu chi tiết đó có chịu được lâu hay không.

## Tài liệu căn cứ liên quan

- 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](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>

## Chương 7. Hậu xử lý hợp kim hàng không vũ trụ in 3D và bảo đảm tuổi thọ mỗi chu kỳ cao

### Mở đầu

Hãy nghĩ đến chiếc máy in 3D dùng trong nhà. Nó nung chảy sợi nhựa rồi xếp từng lớp để tạo ra đồ vật. Linh kiện động cơ tên lửa cũng được chế tạo theo cách tương tự. Chỉ khác là vật liệu không phải nhựa, mà là bột kim loại. Tia laser làm bột kim loại nóng chảy trong chớp mắt rồi đông cứng từng lớp. Bằng cách này, có thể in nguyên khối những hình dạng phức tạp.

Phương pháp này đã làm thay đổi lớn việc chế tạo linh kiện tên lửa. Trước đây, người ta làm riêng nhiều mảnh rồi hàn chúng lại với nhau. Nay, cả linh kiện có đường làm mát bên trong cũng được chế tạo trong một lần. Số lượng linh kiện giảm và trọng lượng nhẹ hơn. Khi công đoạn ghép nhiều mảnh biến mất, thời gian chế tạo cũng ngắn lại.

Nhưng có một vấn đề. Kim loại vừa in xong không thể gắn ngay lên tên lửa. Bề mặt còn thô ráp, bên trong còn những lỗ nhỏ. Ứng suất không nhìn thấy bằng mắt cũng bị giữ lại bên trong. Nếu cho linh kiện quay nhanh trong trạng thái đó, nó có nguy cơ bị vỡ.

Vì vậy, công đoạn xử lý sau khi in rất quan trọng. Công đoạn xử lý này được gọi là hậu xử lý. Người ta cắt gọt bề mặt để làm nó nhẵn hơn. Người ta dùng nhiệt và áp suất để loại bỏ lỗ rỗng bên trong. Người ta dùng nhiệt luyện để làm tổ chức vật liệu đồng đều. Chương này lần lượt trình bày quá trình hậu xử lý, từ làm nhẵn bề mặt đến loại bỏ khuyết tật bên trong.

Hãy nghĩ đến cảnh cắt một miếng thịt dai trong bếp. Dao đi qua đậu phụ mềm rất dễ. Nhưng dao khó lách vào miếng thịt dai. Siêu hợp kim dùng cho tên lửa giống miếng thịt dai. Nó rất khó cắt gọt. Tính chất khó cắt gọt như vậy được gọi là tính khó gia công.

Còn một mối lo lớn hơn. Đó là mỏi. Mỏi là hiện tượng vật liệu bị vỡ khi một lực nhỏ lặp đi lặp lại nhiều lần. Nếu gập rồi duỗi kẹp giấy nhiều lần, nó sẽ đứt. Nó đứt dù ta không tác dụng một lực lớn trong một lần. Cánh quay của turbo bơm tên lửa nằm trong rủi ro như vậy.

Trong chương này, chúng ta xem trí tuệ nhân tạo hỗ trợ hai việc như thế nào. Việc thứ nhất là tìm điều kiện để cắt gọt tốt loại kim loại khó gia công. Việc thứ hai là biết trước linh kiện chịu được bao nhiêu lần tải lặp lại. Cả hai việc đều mất nhiều thời gian nếu chỉ dựa vào kinh nghiệm của con người. Trí tuệ nhân tạo học từ dữ liệu và nhanh chóng thu hẹp câu trả lời.

Tuy vậy, trí tuệ nhân tạo không quyết định mọi thứ. Linh kiện tên lửa liên quan đến sinh mạng con người và chi phí lớn. Dự đoán của trí tuệ nhân tạo phải được kiểm tra lại bằng thí nghiệm và kiểm định. Chương này giải thích cách trí tuệ nhân tạo và thí nghiệm cùng làm việc bằng cách kiểm chứng lẫn nhau.

Sau khi đọc chương này, người đọc sẽ thấy hai điều. Một là vì sao kim loại in 3D cần được xử lý. Hai là trí tuệ nhân tạo hỗ trợ việc xử lý và kiểm định đó như thế nào. Chương này không dùng công thức khó. Nó giải thích từng bước bằng ví dụ trong đời sống. Nếu theo dõi chậm rãi, người đọc sẽ nắm được bức tranh chung của chương này mà không thấy khó.

## 7.1 Linh kiện tên lửa chuyển sang in 3D và những bài toán mới

Trước hết, chúng ta mở rõ nội dung của mục này. Phần tim của động cơ tên lửa là một nơi cực kỳ khắc nghiệt. Một bên có nhiên liệu siêu lạnh dưới âm hơn 200 độ chảy qua. Bên kia có khí nóng hàng trăm độ đi qua. Áp suất cao và rung động lớn. Mục này xem vì sao kim loại dùng ở nơi như vậy lại được chế tạo bằng in 3D, và khi đó xuất hiện những điểm yếu nào.

Linh kiện động cơ tên lửa phải nhẹ và bền. Nếu trọng lượng nặng, khối hàng có thể mang theo sẽ giảm. Nếu độ bền yếu, linh kiện không chịu được áp suất cao. Hai yêu cầu này mâu thuẫn với nhau. Vì khi làm nhẹ, linh kiện dễ yếu đi. Chế tạo bồi đắp (Additive Manufacturing, AM) đã mở ra một con đường để giải bài toán này.

Buồng đốt làm mát tái sinh là một ví dụ tốt. Thành buồng đốt tiếp xúc với ngọn lửa nóng. Nếu cứ để nguyên, thành sẽ nóng chảy. Vì vậy, người ta tạo những đường làm mát mảnh bên trong thành. Nhiên liệu lạnh đi qua các đường đó và làm mát thành. Cấu trúc có đường dẫn ẩn trong thành như vậy rất khó chế tạo bằng phương pháp cũ. Chế tạo bồi đắp tạo được cấu trúc bên trong phức tạp như vậy trong một lần.

Công nghệ tiêu biểu để chế tạo linh kiện là nấu chảy lớp bột bằng laser (Laser Powder Bed Fusion, LPBF). Tia laser đi qua lớp bột kim loại được rải mỏng. Chỗ tia laser chạm vào nóng chảy tức thì rồi nhanh chóng đông cứng. Sau đó lại rải bột lên trên và chiếu laser. Quá trình này được lặp lại hàng nghìn lần để xếp thành linh kiện.

Kim loại dùng nhiều cho linh kiện tên lửa là siêu hợp kim nền niken (superalloy). Siêu hợp kim là kim loại đặc biệt không mất sức bền ngay cả ở nhiệt độ cao. Vật liệu tiêu biểu là Inconel 718. Các hợp kim dòng Haynes cũng được dùng. Những kim loại này chịu được lâu trong dòng khí nóng.

Vấn đề nằm ở trạng thái bên trong của kim loại được laser làm nóng chảy. Chỗ tia laser đi qua nguội rất nhanh. Nó nhanh đến mức có thể nguội hàng trăm nghìn độ trong 1 giây. Khi nguội gấp như vậy, tổ chức vật liệu đông cứng không đồng đều. Tinh thể mọc dài theo hướng xếp lớp. Tinh thể dài như vậy được gọi là hạt dạng cột (columnar grain).

Cũng xuất hiện hiện tượng tính chất thay đổi theo hướng. Hiện tượng này gọi là tính dị hướng (anisotropy). Nghĩ đến gỗ thì dễ hiểu. Gỗ dễ tách khi chẻ theo thớ. Nếu cắt ngang thớ, cần nhiều lực hơn. Kim loại bồi đắp cũng có độ bền khác nhau theo từng hướng. Tính chất này làm cho thiết kế và gia công linh kiện trở nên khó hơn.

Làm nguội nhanh còn để lại dấu vết khác. Mùa đông, nếu rót nước nóng vào cốc thủy tinh đang lạnh, cốc sẽ nứt. Đó là vì bên trong và bên ngoài giãn nở với tốc độ khác nhau. Trong kim loại bồi đắp cũng xảy ra việc tương tự. Tốc độ đông cứng khác nhau tùy vị trí. Vì vậy, lực bị giữ lại bên trong linh kiện. Lực bị giữ lại này được gọi là ứng suất dư (residual stress).

Thêm vào đó là sự lệch thành phần. Khi kim loại nóng chảy đông cứng, một số nguyên tố dồn về một phía. Trong Inconel 718, nguyên tố niobi (Nb) bị dồn lại. Ở chỗ đó hình thành các hạt cứng. Các hạt này được gọi là pha Laves (Laves phase). Pha Laves cứng hơn nhiều so với kim loại nền và dễ vỡ.

Vì trạng thái bên trong như vậy, hai bài toán xuất hiện. Thứ nhất là khó cắt gọt. Các hạt cứng làm hỏng dụng cụ cắt. Thứ hai là lỗ rỗng bên trong trở thành mầm của vết nứt mới. Trong linh kiện quay nhanh, những mầm này rất nguy hiểm. Vì vậy, các mục sau lần lượt bàn về hai bài toán này.

Có lý do khiến chế tạo bồi đắp nhanh chóng chiếm vị trí trong ngành tên lửa. Đó là vì thời gian chế tạo một linh kiện giảm mạnh. Trước đây, chỉ riêng việc làm khuôn đúc đã mất vài tháng. Chế tạo bồi đắp xếp lớp trực tiếp thành linh kiện mà không cần khuôn. Nếu sửa thiết kế, ngày hôm sau có thể in linh kiện mới. Vì có thể lặp lại thiết kế và chế tạo theo chu kỳ ngắn như vậy, tốc độ phát triển nhanh hơn.

Khả năng giảm trọng lượng cũng lớn. Khi hợp nhất nhiều mảnh thành một, mối nối biến mất. Mối nối thường làm tăng trọng lượng và dễ trở thành điểm yếu. Nếu loại bỏ mối nối, linh kiện nhẹ hơn và bền hơn. Vì vậy, các công ty tên lửa đầu tư lớn vào công nghệ này. Cơ quan Hàng không và Vũ trụ Hoa Kỳ cũng đã thử nghiệm buồng đốt bồi đắp trong thời gian dài.

Hậu xử lý nhất định phải có nhiệt luyện. Nhiệt luyện là việc nung nóng rồi làm nguội kim loại để thay đổi tổ chức của nó. Trước tiên, người ta nung đến nhiệt độ cao để hòa tan các hạt cứng vào nền. Bước này được gọi là xử lý dung dịch. Sau đó, người ta giữ ở nhiệt độ thích hợp trong thời gian dài để các hạt tăng bền mọc đều. Bước này được gọi là hóa già. Siêu hợp kim phải đi qua hai bước này mới phát huy đúng sức bền của nó.

Nói gọn lại, chế tạo bồi đắp mang đến cả lợi thế lớn và bài toán mới. Lợi thế là hình dạng phức tạp và chế tạo nhanh. Bài toán là bề mặt thô ráp và khuyết tật bên trong. Chìa khóa để giải bài toán này là hậu xử lý và kiểm định. Và trí tuệ nhân tạo hỗ trợ bên cạnh trong toàn bộ quá trình hậu xử lý và kiểm định đó.

## 7.2 Vì sao siêu hợp kim chế tạo bằng bồi đắp khó cắt gọt

Trước hết, chúng ta mở rõ nội dung của mục này. Tính khó gia công là tính chất khó cắt gọt. Siêu hợp kim chế tạo bằng bồi đắp làm cho một kim loại vốn đã khó cắt gọt lại càng khó cắt gọt hơn. Chúng ta chia thành ba lý do để xem xét. Đó là tính định hướng của tổ chức tinh thể, ứng suất bị giữ lại và hóa bền do gia công, cùng quá trình dụng cụ bị mòn.

### 7.2.1 Tổ chức tinh thể thay đổi theo hướng

Tiểu mục này trình bày ảnh hưởng của tính định hướng trong tổ chức tinh thể đối với quá trình cắt gọt. Ở phần trước, chúng ta đã nói hạt dạng cột mọc dài theo hướng xếp lớp. Ở đây, chúng ta xem kỹ vì sao tổ chức này làm cho việc cắt gọt trở nên khó hơn. Chúng ta cũng xem tác động của pha Laves cứng đối với dụng cụ.

Kim loại được tạo nên từ nhiều tinh thể nhỏ. Trong từng tinh thể, các nguyên tử xếp theo quy luật. Tinh thể đứng theo hướng nào sẽ chi phối tính chất của kim loại. Trong kim loại bồi đắp, tinh thể nghiêng về một hướng. Vì tinh thể mọc dài theo hướng xếp lớp, tính chất thay đổi tùy hướng.

Khi dụng cụ cắt đi qua tổ chức này, lực cản thay đổi. Góc gặp nhau giữa hướng tinh thể và hướng lưỡi cắt khác nhau tùy vị trí. Có chỗ được cắt khá êm. Có chỗ chống lại mạnh. Vì vậy, lực cắt tăng giảm theo chu kỳ. Khi lực không đều, gia công mất ổn định và bề mặt cũng trở nên không đồng đều.

Sự tăng giảm này tạo ra va đập lặp lại lên dụng cụ. Khi va đập tích tụ, lưỡi cắt dần bị hư hại. Rung động cũng lớn hơn. Rung động này được gọi là chatter. Khi chatter nặng, trên bề mặt để lại vân sóng. Với linh kiện chính xác, những vân như vậy là khuyết điểm lớn.

Các hạt cứng còn làm vấn đề lớn hơn. Khi Inconel 718 đông cứng, niobi dồn về ranh giới tinh thể. Ở đó xuất hiện pha Laves dạng tấm hoặc dạng khối. Các hạt này vừa cứng hơn nhiều so với kim loại nền, vừa có tính dễ vỡ.

Điều gì xảy ra khi dụng cụ cắt va vào các hạt cứng này? Nó giống như đưa dao đi qua sỏi lẫn trong đất mềm. Lưỡi dao bị mẻ nhỏ. Khi những hư hỏng li ti như vậy tiếp tục tích tụ, cuối cùng lưỡi dụng cụ bị mòn cùn.

Muốn giải vấn đề này, trước hết phải làm cho tổ chức đồng đều. Nhiệt luyện được trình bày ở phần sau đảm nhận vai trò đó. Nhiệt luyện hòa tan các hạt cứng và đưa chúng trở lại vào nền. Nó cũng làm các hạt dạng cột dài trở nên tròn hơn một chút. Khi đó, mức độ tính chất thay đổi theo hướng sẽ giảm.

Khác biệt tổ chức này liên quan trực tiếp đến tính năng của linh kiện. Nếu độ bền khác nhau theo hướng, thiết kế trở nên khó. Phải đặt linh kiện sao cho tránh hướng yếu. Khi làm tổ chức đồng đều, nỗi lo này giảm đi. Vì vậy, hậu xử lý không chỉ giúp gia công, mà còn làm tính năng tốt hơn.

Trí tuệ nhân tạo đưa thông tin tổ chức này vào dự đoán. Tinh thể hướng về phía nào, có bao nhiêu hạt cứng, đều trở thành dữ liệu đầu vào. Nó học xem lực cản cắt thay đổi ra sao theo các giá trị này. Con người khó xem từng ảnh tổ chức bằng mắt. Trí tuệ nhân tạo đọc nhanh nhiều ảnh và tìm ra quy luật.

Hướng của tinh thể được đo bằng thiết bị đặc biệt. Khi chiếu chùm electron lên bề mặt mẫu, mỗi tinh thể tạo ra một hoa văn khác nhau. Từ hoa văn này, ta biết từng tinh thể đang hướng về phía nào. Thông tin hướng này trở thành đầu vào tốt cho trí tuệ nhân tạo. Con người khó đếm hướng bằng mắt. Máy móc đọc hướng của vô số tinh thể trong một lần.

Góc xếp lớp cũng làm thay đổi tính chất. Tổ chức sẽ khác nhau tùy theo ta dựng đứng hay đặt nằm linh kiện khi xếp lớp. Dù cùng một bản vẽ, nếu hướng đặt khác nhau thì độ bền khác nhau. Vì vậy, hướng xếp lớp được quyết định ngay từ giai đoạn thiết kế. Góc này cũng ảnh hưởng đến khả năng gia công và mối. Trí tuệ nhân tạo đưa góc vào như một biến số để học ảnh hưởng đó.

Tuy vậy, tổ chức mỗi linh kiện vẫn khác nhau một chút. Dù chế tạo bằng cùng một thiết lập, vẫn có sai lệch tùy vị trí. Dao động nhỏ của laser và khác biệt kích thước bột tạo ra sai lệch. Vì vậy, trong dự đoán của mô hình vẫn còn dao động. Cách xử lý dao động này sẽ được xem kỹ ở các mục sau.

### 7.2.2 Ứng suất bị giữ lại và tính chất càng cắt càng cứng

Tiểu mục này trình bày ứng suất bị giữ lại và hóa bền do gia công. Ở phần trước, chúng ta đã nói có lực bị giữ lại trong kim loại bồi đắp. Chúng ta xem lực bị giữ lại này gây ra vấn đề gì khi cắt gọt. Chúng ta cũng xem tính chất siêu hợp kim trở nên cứng hơn trong lúc cắt gọt.

Trước hết, hãy nhắc lại ứng suất dư. Bên trong linh kiện có lực bị giữ lại theo nhiều hướng. Tập hợp lực này được biểu diễn bằng một hộp toán học gọi là tensor. Có thể xem tensor như một bảng chứa cùng lúc các giá trị theo nhiều hướng. Dù mắt không thấy, chúng vẫn đẩy kéo lẫn nhau trong linh kiện.

Lực bị giữ lại này lộ ra thành vấn đề vào khoảnh khắc cắt gọt. Khi cắt bỏ bề mặt, lực ở chỗ đó được giải phóng. Khi đó, phần còn lại mất cân bằng. Cánh mỏng hoặc linh kiện dạng vòng sẽ bị cong. Trong lúc cố đạt kích thước chính xác, hình dạng lại méo đi. Vì vậy, người ta phải định sẵn cả thứ tự sẽ cắt mặt nào trước.

Tiếp theo là hóa bền do gia công (work hardening). Hóa bền do gia công là tính chất kim loại càng biến dạng càng trở nên cứng hơn. Nếu uốn dây thép nhiều lần, chỗ đó trở nên cứng và khó uốn. Siêu hợp kim có tính chất này mạnh. Lớp mỏng trước lưỡi cắt trở nên cứng hơn trong chớp mắt.

Khi đó, dụng cụ phải cắt lại lớp vừa mới cứng lên. Cần lực lớn hơn. Nhiệt cũng sinh ra nhiều hơn. Inconel 718 là kim loại không thoát nhiệt tốt. Độ dẫn nhiệt của nó chỉ khoảng một phần ba so với thép thông thường. Vì vậy, phần nhiệt không thoát ra được tích lại giữa lưỡi cắt và chi tiết, làm nhiệt độ tăng thêm.

Pha tăng bền cũng làm dụng cụ vất vả hơn. Bên trong Inconel 718 có những hạt tăng bền rất nhỏ phân tán khắp nơi. Hạt tăng bền chính của hợp kim này là gamma double-prime có chứa niobi ( $\gamma''$ , Ni<sub>3</sub>Nb). Những hạt này giữ cho kim loại cứng chắc. Ngay cả ở nhiệt độ cao, chúng cũng không dễ bị hòa tan. Vì vậy, trong vùng cắt nóng, kim loại vẫn không mềm đi và dụng cụ tiếp tục chịu lực lớn.

Tất cả các yếu tố này chồng lên nhau và gây hư hại bề mặt. Một lớp mỏng bên dưới bề mặt đã cắt bị biến đổi. Cấu trúc bị nghiền nhỏ và màu chuyển sang trắng. Lớp này được gọi là lớp trắng (white layer). Trong lớp trắng dễ còn lại ứng suất dư kéo cùng với vết nứt vi mô. Lớp này trở thành điểm khởi đầu của vết nứt và làm giảm mạnh tuổi thọ mỏi của chi tiết.

Vì sao lớp trắng nguy hiểm. Trong lớp này có ứng suất dư kéo mà mắt thường không nhìn thấy. Ứng suất kéo là lực mở rộng vết nứt. Nếu thêm vết nứt vi mô, nó trở thành hạt giống. Khi chịu tải lặp lại, vết nứt lớn lên từ hạt giống này. Vì vậy, ở bước gia công tinh cuối cùng không được để lại lớp trắng.

Cũng có cách giảm vấn đề này. Sau khi cắt, thêm một bước hậu xử lý bằng cách gõ bề mặt bằng các hạt nhỏ. Xử lý này để lại ứng suất nén trên bề mặt. Ứng suất nén là lực ép đóng vết nứt. Nó ngược chiều với ứng suất kéo. Khi cấy ứng suất nén vào bề mặt như vậy, tuổi thọ mỏi của chi tiết tăng lên.

Trí tuệ nhân tạo cũng học đồng thời hiệu quả của nhiều loại hậu xử lý này. Điều kiện cắt, xử lý gõ bề mặt, và xử lý nhiệt đều được đưa vào làm đầu vào. Khi đó, nó dự đoán tổ hợp nào tạo ra kết quả tốt. Nó giảm mạnh công việc mà con người từng phải thử từng trường hợp. Tuy vậy, bước xác nhận cuối cùng vẫn được thực hiện bằng thử nghiệm thực tế.

Trí tuệ nhân tạo được dùng để nhận biết trước nguy cơ này. Khi đưa điều kiện cắt và giá trị ứng suất dư vào, nó dự đoán trạng thái bề mặt. Nó học lớp trắng hình thành trong điều kiện nào. Nhờ đó, có thể tránh các điều kiện nguy hiểm. Với chi tiết titan chế tạo bồi đắp, đã có mô hình học sâu dự đoán độ nhám bề mặt và độ cứng từ điều kiện quy trình. Cuốn sách này đề xuất áp dụng cùng cách tiếp cận đó cho bề mặt cắt của siêu hợp kim.

Tuy nhiên, dự đoán chỉ đáng tin trong phạm vi đã học. Với điều kiện chưa từng thấy, sai số tăng lên. Vì vậy, chi tiết thực tế bắt buộc phải được xác nhận bằng kiểm tra bề mặt. Lớp trắng rất mỏng nên mắt thường dễ bỏ sót. Trí tuệ nhân tạo lọc trước các điều kiện nguy hiểm và thu hẹp các phương án.

### 7.2.3 Cách dụng cụ bị mòn và hư hại bề mặt

Tiểu mục này bàn về nhiều cách dụng cụ bị mòn. Mòn dụng cụ (tool wear) liên quan trực tiếp đến chất lượng sản xuất. Khi dụng cụ mòn, bề mặt trở nên thô ráp. Bề mặt thô ráp trở thành điểm khởi đầu của vết nứt mỏi. Vì vậy, biết trước dụng cụ đã mòn đến mức nào là việc quan trọng trong quản lý chất lượng.

Có nhiều loại mòn dụng cụ. Thứ nhất là mòn khía. Ở đường ranh của chiều sâu cắt, một rãnh sâu bị khoét vào dụng cụ. Đây là nơi oxy trong không khí và lớp kim loại đã cứng cùng cọ xát. Oxy và lớp kim loại cứng cùng cọ xát làm rãnh ở vị trí này ngày càng sâu.

Thứ hai là bong tróc do bám dính. Niken rất dễ dính vào vật liệu dụng cụ. Phoi kim loại đã cắt bị ép dính lên lưỡi cắt. Khối này được gọi là lẹo dao. Khi khối này bong ra, nó kéo theo cả lớp phủ trên bề mặt dụng cụ. Lúc đó, lớp phủ bị xé đi và bề mặt dụng cụ hư hại nhanh chóng.

Thứ ba là mòn khuếch tán. Vùng cắt có nhiệt độ rất cao. Ở nhiệt độ cao, các nguyên tử thấm vào nhau. Nguyên tử trong dụng cụ di chuyển sang phía phoi kim loại. Khi đó, trên bề mặt dụng cụ còn lại vết lõm, và vết này được gọi là miệng hố.

Thứ tư là lớp trắng bề mặt và vết nứt vi mô. Lớp trắng đã nói ở trên cũng xuất hiện ở đây. Biến dạng nhanh và làm nguội nhanh chồng lên nhau tạo ra một lớp trắng mỏng. Trong lớp này dễ ẩn chứa vết nứt. Bề ngoài có thể trông nhẵn, nhưng bên trong có thể ẩn các khuyết tật như vết nứt.

Tất cả các dạng mòn này liên kết với nhau. Khi một dạng lớn lên, các dạng khác cũng tăng theo. Vì vậy, không thể chỉ nhìn một nguyên nhân. Phải xem nhiều tín hiệu cùng lúc. Lực cắt, rung động, âm thanh, và nhiệt độ đều là tín hiệu của mòn.

Trí tuệ nhân tạo đọc các tín hiệu này cùng nhau. Cảm biến thu thập tín hiệu theo thời gian thực trong khi cắt. Mô hình dùng các tín hiệu này để ước tính hiện tại dụng cụ đã mòn đến mức nào. Nghiên cứu gần đây đã dùng học sâu đọc ảnh dụng cụ bị mòn và dự đoán ngay mức độ mòn. Công việc mà con người từng phải dùng dụng cụ rồi xem bằng kính lúp nay được máy móc làm thay.

Cách tiếp cận tương tự cũng có tác dụng với kim loại khác. Một nghiên cứu đã dự đoán mòn dụng cụ khi cắt hợp kim niken gọi là Hastelloy C276 bằng mô hình dựa trên dữ liệu. Lý do là nhiều siêu hợp kim có cách mòn tương tự. Những kết quả này cũng có giá trị tham khảo cho siêu hợp kim dùng trong tên lửa.

Nếu dự đoán được mòn dụng cụ, gia công không người vận hành cũng trở nên khả thi. Máy tự theo dõi trạng thái dụng cụ. Nếu nguy hiểm, máy giảm tốc độ hoặc thay dụng cụ. Con người không phải luôn đứng quan sát. Trong nhà máy cắt chi tiết suốt đêm không có người như vậy, mô hình dự đoán nâng đỡ sự an toàn đó.

Bốn dạng mòn này phụ thuộc rất lớn vào nhiệt độ. Vùng cắt càng nóng, mòn càng nhanh. Vì vậy, làm mát rất quan trọng. Dầu cắt được phun để hạ nhiệt. Cũng có khi dầu cắt được bắn trực tiếp vào lưỡi cắt bằng áp suất cao. Phương pháp làm mát này cũng trở thành biến đầu vào của trí tuệ nhân tạo.

Tóm lại, dự đoán mòn dụng cụ trở thành con mắt của quản lý chất lượng. Có thể thay dụng cụ trước khi mòn trở nên nghiêm trọng. Có thể sửa điều kiện trước khi tạo ra chi tiết lỗi. Tuy nhiên, tín hiệu cảm biến có lẫn nhiễu. Rung động cũng có lẫn dao động của chính máy. Vì vậy, mô hình phải được huấn luyện tốt để chịu được nhiễu. Khi dùng nhiều cảm biến cùng lúc, ảnh hưởng của nhiễu giảm xuống.

## 7.3 Tính toán lực để cắt tốt và tối ưu hóa hậu xử lý

Ở mục trước, chúng ta đã xem vì sao việc cắt khó. Ở mục này, chúng ta xem cách vượt qua khó khăn đó. Ba nội dung sẽ được bàn lần lượt. Trước hết, xem cách tính lực tác dụng trong quá trình cắt. Tiếp theo, xem tính gia công thay đổi ra sao theo hậu xử lý. Cuối cùng, xem trí tuệ nhân tạo tìm điều kiện tối ưu như thế nào.

### 7.3.1 Lực tác dụng trong quá trình cắt và mô hình xử lý lực đó

Tiểu mục này xem xét lực tác dụng trong quá trình cắt và mô hình xử lý lực đó. Lực cắt (cutting force) là lực mà dụng cụ dùng để đẩy kim loại. Nếu biết trước lực này, có thể chọn dụng cụ và điều kiện tốt hơn. Chúng ta sẽ xem ý tưởng cơ bản để tính lực cùng với vấn đề rung động.

Lực cắt không chỉ tác dụng theo một hướng. Có lực đẩy về phía trước, lực đẩy sang bên, và lực ép xuống dưới. Vì vậy, lực cắt được biểu diễn bằng một vectơ có ba thành phần hướng. Vectơ là đại lượng có cả độ lớn và hướng. Ngược lại, ứng suất trong vật liệu được biểu diễn bằng tensor. Tensor là đại lượng chứa các giá trị theo nhiều hướng cùng lúc. Vectơ lực cắt và tensor ứng suất là hai đại lượng khác nhau. Độ lớn của lực này phụ thuộc vào chiều dày cắt và độ bền của vật liệu.

Trong khi cắt, ba việc xảy ra cùng lúc. Kim loại biến dạng, nhiệt sinh ra, và ma sát xuất hiện. Một mô hình tốt phải xử lý cả ba việc này cùng nhau. Độ bền của vật liệu khi bị ép nhanh thay đổi theo nhiệt độ. Khi nóng lên, kim loại mềm đi một chút. Quan hệ này được đưa vào mô hình bằng bảng, không phải bằng công thức.

Mô hình này dùng các hằng số khác nhau cho từng vật liệu. Hằng số được xác định bằng thí nghiệm. Dù cùng là Inconel 718, hằng số của vật liệu chế tạo bồi đắp và vật liệu rèn vẫn khác nhau. Lý do là vật liệu chế tạo bồi đắp có cấu trúc khác. Vì vậy, cần thí nghiệm riêng cho vật liệu chế tạo bồi đắp. Khi dữ liệu thí nghiệm riêng cho vật liệu chế tạo bồi đắp tích lũy đủ, mô hình mới trở nên chính xác.

Trí tuệ nhân tạo cũng hỗ trợ thí nghiệm xác định hằng số. Có phương pháp tìm ngược hằng số từ tín hiệu cắt. Đo lực và nhiệt độ, rồi chọn hằng số phù hợp với các giá trị đó. Công việc mà con người từng chĩnh bằng tay được máy thực hiện nhanh chóng. Khi đó, mô hình có thể được khớp nhanh cho vật liệu mới.

Vấn đề rung động cũng quan trọng. Ở trên đã nói về chatter. Chatter là rung động trong đó dụng cụ và chi tiết làm rung lẫn nhau. Khi đã rung một lần, ở vòng quay tiếp theo dụng cụ lại cắt đúng vết đó. Khi đó, rung động lớn hơn. Dao động nhỏ tự lớn dần lên và cản trở gia công.

Để ngăn rung động này, cần biết chiều sâu cắt an toàn. Nếu cắt quá sâu, rung động tăng lên. Nếu cắt nông, rung động lắng xuống. Ranh giới đó thay đổi theo tốc độ quay. Hình vẽ thể hiện ranh giới này được gọi là biểu đồ thùy ổn định. Dựa vào hình này, người ta chọn điều kiện không gây rung động.

Trí tuệ nhân tạo giúp phép tính này nhanh hơn. Nếu chỉ dùng mô hình vật lý, tính toán mất nhiều thời gian. Mỗi khi thay đổi điều kiện một chút, phải giải lại từ đầu. Trí tuệ nhân tạo học trước kết quả của mô hình vật lý. Sau đó, nó đưa ra ngay đáp án cho điều kiện mới. Mô hình thay thế cho phép tính nặng như vậy được gọi là surrogate. Surrogate có nghĩa là đại diện. Chỉ cần thực hiện phép tính chính xác trước vài lần. Mô hình đại diện đã học từ các kết quả đó sẽ phụ trách phần sau.

Cách này giảm mạnh thời gian tại hiện trường. Trước đây, người ta phải thay đổi điều kiện và cắt thử nhiều lần. Nay mô hình lọc trước các điều kiện nguy hiểm. Chỉ vài phương án còn lại mới được thử thực tế. Tuy nhiên, nếu mô hình vật lý sai thì trí tuệ nhân tạo cũng sai. Vì vậy, dùng cả vật lý và dữ liệu cùng nhau là cách an toàn hơn.

Phương pháp dùng cả vật lý và dữ liệu được gọi là học máy dựa trên vật lý (PIML). Phương pháp này đưa các định luật vật lý vào trí tuệ nhân tạo. Dù dữ liệu ít, định luật vật lý vẫn giữ câu trả lời đúng hướng. Nó trở thành thiết bị an toàn ngăn dự đoán vô lý. Ý tưởng này sẽ xuất hiện nhiều lần trong các chương sau của cuốn sách. Nó có ích lớn trong lĩnh vực khan hiếm dữ liệu như kỹ thuật tên lửa.

### **7.3.2 So sánh tính gia công của vật liệu rèn, trạng thái ngay sau chế tạo bồi đắp, và vật liệu đã xử lý HIP**

Tiểu mục này so sánh kim loại ở ba trạng thái. Dù cùng là Inconel 718, tính chất thay đổi theo cách chế tạo và hậu xử lý. Chúng ta sẽ đặt cạnh nhau vật liệu rèn, trạng thái ngay sau chế tạo bồi đắp, và trạng thái đã qua ép đẳng tĩnh nóng.

Trước hết, hãy xem vật liệu rèn (wrought). Vật liệu rèn là vật liệu được tạo ra bằng cách ép và nén kim loại. Nó có cùng gốc với cách thợ rèn gõ sắt trong lò rèn. Cấu trúc mịn và đồng đều. Tinh thể nhỏ và tròn. Vì vậy, dao động của lực cắt thuộc nhóm nhỏ hơn trong ba vật liệu. Mòn dụng cụ cũng ít hơn tương ứng, nên việc dự đoán dễ hơn nhiều.

Tiếp theo là trạng thái ngay sau chế tạo bồi đắp. Trạng thái này còn gọi là as-built. Đây là trạng thái được lấy ra nguyên vẹn từ máy in, chưa qua hậu xử lý. Ứng suất dư lớn và pha Laves cứng tập trung nhiều. Tinh thể dạng cột cũng thô và dài. Vì vậy, sức cản cắt tăng mạnh hơn so với vật liệu rèn. Nguy cơ dụng cụ bị vỡ cũng thuộc nhóm lớn hơn trong ba vật liệu.

Thứ ba là trạng thái đã qua ép đẳng tĩnh nóng (Hot Isostatic Pressing, HIP). HIP là hậu xử lý dùng đồng thời nhiệt độ cao và áp suất cao. Thường tiến hành ở nhiệt độ trên 1000 độ C. Áp suất cũng vượt quá 100 megapascal. Trong điều kiện này, các lỗ rỗng trong kim loại bị ép đóng lại.

Hãy nghĩ đến bột nhào làm bánh mì thì dễ hiểu. Trong bột nhào có các bọt khí nhỏ. Nếu ép và nhào lâu, các bọt này thoát ra và bột trở nên đặc hơn. HIP cũng tương tự. Áp suất cao ép các lỗ rỗng trong kim loại dính lại. Áp suất cao và nhiệt làm nguyên tử di chuyển, khiến thành lỗ rỗng áp sát vào nhau.

HIP cũng thay đổi cấu trúc. Tuy nhiên, mức độ thay đổi phụ thuộc vào nhiệt độ và thời gian giữ. Nếu nhiệt độ đủ cao, pha Laves cứng hòa tan vào nền. Nếu nhiệt độ thấp hoặc thời gian ngắn, một phần pha Laves vẫn còn lại. Tinh thể dạng cột kéo dài cũng có thể vẫn giữ nguyên tùy điều kiện. Khi cấu trúc trở nên đồng đều, khác biệt theo hướng và dao động lực cắt giảm xuống. Vì vậy, HIP được dùng cùng với xử lý dung dịch, hóa già, và xử lý nhiệt đồng nhất hóa bằng cách khớp điều kiện với nhau.

Tuy vậy, HIP cũng có giới hạn. Nó không loại bỏ được mọi lỗ rỗng. Lỗ rỗng nối với bề mặt và lỗ rỗng bị giữ bên trong hoạt động khác nhau. Bọt argon bị mắc kẹt trong quá trình chế tạo bồi đắp vẫn gây vấn đề trở lại dù đã được HIP đóng lại. Lý do là argon là khí không hòa tan trong kim loại. Khi sau đó lại chịu

hiệu suất cao, bột đã đóng sẽ phồng lên và mở ra. Hiện tượng này được gọi là nổ khí do nhiệt. Vì vậy, điều quan trọng là giảm bọt khí ngay từ đầu.

Sau HIP cũng xuất hiện vấn đề mới. Khi cấu trúc mềm hơn, phoi kim loại không dễ gãy. Phoi kéo dài và quấn quanh dụng cụ. Nếu hiện tượng quấn này nặng, nó cào xước bề mặt. Vì vậy, phải chọn hình dạng dụng cụ sao cho phoi dễ gãy.

Trí tuệ nhân tạo sắp xếp phép so sánh này thành dữ liệu. Nó gom dữ liệu cắt gọt của từng trạng thái rồi học sự khác nhau. Khi đó, nó phán đoán chi tiết mới gần với trạng thái nào. Nó cũng đề xuất chi tiết đó cần trải qua xử lý sau nào. Các nghiên cứu gần đây đang xem xét rộng rãi cách tối ưu hóa đồng thời HIP và nhiệt luyện.

HIP là một bước xử lý sau tốn chi phí và thời gian. Phải nung nóng và ép trong thời gian dài bên trong thiết bị lớn. Vì vậy, khó dùng cho mọi chi tiết. Cần phán đoán chi tiết nào thật sự cần HIP. Trí tuệ nhân tạo cũng hỗ trợ phán đoán này. Nó nhìn vào thông tin khuyết tật rồi ước lượng trước hiệu quả của HIP.

Các nghiên cứu công khai cũng là chỉ dẫn tốt. Cơ quan Hàng không và Vũ trụ Hoa Kỳ đã dùng mô phỏng để xem HIP làm giảm khuyết tật của Inconel 718 chế tạo bồi đắp như thế nào. Một nghiên cứu đã xử lý bằng HIP các chi tiết ban đầu có nhiều lỗ rỗng, rồi đo cường độ mỏi. Phần lớn lỗ rỗng đã khép lại, nhưng các bọt argon vẫn còn. Những kết quả như vậy tích lũy lại và trở thành nền tảng cho thiết kế xử lý sau.

Có lý do để lấy vật liệu rèn làm chuẩn. Vật liệu rèn là vật liệu đã được kiểm chứng trong thời gian dài. Tính chất của nó được hiểu rõ và dữ liệu cũng nhiều. Mục tiêu là đưa vật liệu chế tạo bồi đắp lên mức của vật liệu rèn. Xử lý sau là con đường tiến gần đến mục tiêu đó. Trí tuệ nhân tạo giúp tìm con đường ấy nhanh hơn.

### 7.3.3 Tìm điều kiện cắt gọt tốt bằng ít thí nghiệm

Tiểu mục này bàn về cách tìm điều kiện tốt bằng ít thí nghiệm. Có nhiều điều kiện cắt gọt. Tốc độ, lượng chạy dao, chiều sâu và cách làm mát đều là biến số. Nếu tổ hợp hết các biến số này, số thí nghiệm sẽ tăng bùng nổ. Phương pháp thiết kế thí nghiệm Taguchi và mạng nơ-ron nhân tạo được kết hợp để giải bài toán này.

Phương pháp thiết kế thí nghiệm Taguchi (Taguchi method) là cách so sánh nhiều điều kiện bằng ít thí nghiệm. Nó không thử tất cả các tổ hợp. Nó chỉ chọn các tổ hợp đại diện để thí nghiệm. Chỉ với vài tổ hợp được chọn tốt, ta đọc được xu hướng chung. Số lần thí nghiệm giảm, nên thời gian và chi phí cũng giảm theo.

Phương pháp này làm giảm mạnh số thí nghiệm. Ví dụ, có thể giảm hàng trăm tổ hợp xuống còn mười tám thí nghiệm. Trong mỗi thí nghiệm, ta đo độ nhám bề mặt và độ mòn dụng cụ. Ta cũng đo năng lượng dùng cho cắt gọt. Các giá trị thu được từ nhiều tổ hợp như vậy trở thành dữ liệu học của trí tuệ nhân tạo.

Mạng nơ-ron nhân tạo (Artificial Neural Network, ANN) học quy luật từ dữ liệu này. Mạng nơ-ron nhân tạo là cấu trúc tính toán mô phỏng kết nối thần kinh trong não người. Đầu vào gồm điều kiện cắt gọt và trạng thái nhiệt luyện. Đầu ra gồm độ nhám bề mặt và độ mòn dụng cụ. Mô hình tự tìm mối quan hệ giữa đầu vào và đầu ra rồi học thành quy luật.

Sau khi học xong, mô hình cũng dự đoán được các điều kiện chưa thí nghiệm. Nếu đưa vào tốc độ mới và lượng chạy dao mới, nó cho biết trước kết quả. Sau đó, ta gắn thêm bước tìm kiếm điều kiện tối ưu. Ta chọn điều kiện làm bề mặt mịn và làm dụng cụ mòn ít. Khi hai mục tiêu va chạm nhau, ta tìm điểm cân bằng.

Các trường hợp thực tế cũng đang tích lũy. Một nghiên cứu xử lý tiện vi mô đối với Inconel 718 chế tạo bồi đắp. Nghiên cứu xem khả năng gia công của vật liệu đã qua nhiệt luyện thay đổi ra sao. Nó so sánh

tốc độ cắt từ 30 mét đến 60 mét mỗi phút. Những dữ liệu như vậy được gom lại và trở thành vật liệu học cho mô hình.

Tổ hợp này rất hữu ích tại hiện trường. Trước hết, ta lập một bảng thí nghiệm nhỏ. Sau đó, ta thu dữ liệu bằng cảm biến và phép đo. Cuối cùng, trí tuệ nhân tạo học mối quan hệ giữa điều kiện và chất lượng. Làm như vậy, ta có thể nhanh chóng xác định điều kiện cho chi tiết mới.

Khi có nhiều mục tiêu, cần tìm điểm cân bằng. Ta muốn bề mặt mịn, dụng cụ mòn ít và thời gian ngắn. Các mục tiêu này va chạm nhau. Nếu cắt nhanh, bề mặt dễ trở nên thô. Vì vậy, nếu chỉ cố làm tốt một mục tiêu, mục tiêu khác sẽ xấu đi.

Khi đó, ta dùng phương pháp xử lý nhiều mục tiêu cùng lúc. Ta nhường một chút ở mục tiêu này để đạt mục tiêu kia. Ta gom các tổ hợp tốt thu được theo cách đó. Trong số ấy, ta chọn tổ hợp phù hợp với chi tiết. Trí tuệ nhân tạo nhanh chóng tìm ra các tổ hợp này.

Mục tiêu thực tế là làm độ nhám bề mặt nhỏ đi. Chi tiết tên lửa phải có bề mặt mịn thì mới bền mỗi. Nếu dùng điều kiện cắt gọt được chọn tốt cùng với dụng cụ tốt, ta tiến gần đến mục tiêu này. Bề mặt mịn làm giảm các điểm khởi đầu của vết nứt. Vì vậy, hoàn thiện bề mặt là nút thắt đầu tiên của bảo đảm mỗi.

Giới hạn cũng rõ ràng. Mô hình chỉ đáng tin trong phạm vi nó đã học. Với tốc độ hoặc lượng chạy dao không có trong dữ liệu học, dự đoán sẽ dao động. Nếu dữ liệu ít, sai số sẽ lớn. Vì vậy, với chi tiết quan trọng, ta xác nhận điều kiện bằng thử nghiệm cắt gọt thực tế. Ta kiểm chứng dự đoán có đúng không bằng độ nhám bề mặt và độ mòn dụng cụ.

Lựa chọn dụng cụ cũng là biến số quan trọng. Với siêu hợp kim, ta dùng dụng cụ phủ lớp phủ đặc biệt. Lớp phủ giúp chịu nhiệt và chịu mài mòn. Đôi khi cũng dùng dụng cụ làm bằng vật liệu gốm. Trí tuệ nhân tạo cũng học dụng cụ nào phù hợp với điều kiện nào. Khi đó, ta có thể cân bằng chi phí dụng cụ và chất lượng gia công.

Nhiệm vụ phía trước là mở rộng dữ liệu. Nếu gom dữ liệu thí nghiệm của nhiều tổ chức, mô hình sẽ vững hơn. Tuy nhiên, mỗi tổ chức có cách đo hơi khác nhau. Cách ghi điều kiện cắt gọt cũng không giống nhau. Vì vậy, cần quy tắc để thống nhất dữ liệu. Khi có chuẩn như vậy, độ tin cậy của dự đoán sẽ tăng lên.

## 7.4 Phá hủy mỗi chu kỳ cực cao của cánh turbopump

Đến đây, ta đã bàn về cách cắt gọt cho tốt. Từ mục này, ta bàn về chi tiết chịu được bao lâu. Ta xem vì sao cánh quay của turbopump trong tên lửa lại nguy hiểm, và khi chi tiết đó vỡ, nó để lại dấu vết nào.

### 7.4.1 Môi trường khắc nghiệt có cả nhiệt độ cực thấp và nhiệt độ cao

Tiểu mục này bàn về môi trường khắc nghiệt nơi cánh turbopump làm việc. Turbopump (turbopump) là bơm đẩy nhiên liệu và chất oxy hóa vào với áp suất cao. Cánh quay nhanh bên trong nó được gọi là bánh công tác (impeller). Ta sẽ xem vì sao tải trọng mà chi tiết này chịu lại khác thường.

Bánh công tác quay rất nhanh. Nó quay từ 30.000 đến 80.000 vòng mỗi phút. Đây là tốc độ nhanh hơn động cơ ô tô hơn mười lần. Trong khi quay, lực ly tâm kéo cánh ra phía ngoài. Nếu tốc độ quay ổn định, lực ly tâm này là tải trọng tĩnh có độ lớn gần như không đổi.

Môi trường nhiệt độ cũng khắc nghiệt. Bản thân bánh công tác tiếp xúc với chất đẩy ở nhiệt độ cực thấp. Hydro lỏng lạnh hơn âm 250 độ. Oxy lỏng cũng ở gần âm 180 độ. Khí làm việc nóng không nằm ở bánh công tác, mà ở phía turbine gắn trên cùng trục. Nhiệt này truyền dần sang phía bơm qua trục, ổ đỡ và phốt. Vì vậy, hai đầu của một trục cùng lúc đối diện với nơi rất lạnh và nơi rất nóng.

Chênh lệch nhiệt độ lớn tạo ra co và giãn. Phía lạnh co lại, phía nóng nở ra. Sự khác nhau này tạo lực bên trong chi tiết. Trên nền đó có lực ly tâm tĩnh. Phía trên nữa là tải trọng rung lặp lại chồng lên và làm chi tiết dao động.

Tải trọng lặp lại đến từ nhiều nơi. Mỗi khi cánh đi qua cạnh chi tiết cố định, lực lại tác động vào. Mất cân bằng rất nhỏ của trục quay và dao động áp suất cũng góp thêm lực. Khởi động và dừng máy cũng mỗi lần tạo một tải trọng lớn. Lực do cánh đi qua lặp lại hàng nghìn lần mỗi giây. Vì vậy, số chu kỳ lặp tích lũy rất nhanh.

Mỗi có số chu kỳ lặp lớn như vậy được gọi là mỗi chu kỳ rất cao (Very High Cycle Fatigue, VHCF). Thiết kế thông thường lấy mười triệu chu kỳ làm chuẩn. Muốn biết chi tiết thực tế chịu tải lặp đến bao nhiêu lần, phải nhân tần số kích thích với thời gian vận hành tích lũy. Trong nghiên cứu mẫu thử, người ta xử lý đến vùng 100 triệu lần và 1 tỷ lần. Ở vùng này, quy luật mỗi thay đổi, nên phải xem xét riêng.

Thử nghiệm mỗi thông thường khó xử lý vùng này. Nếu thử 1 tỷ lần ở tốc độ thường, thời gian sẽ quá lâu. Vì vậy, người ta dùng máy thử tần số cao đặc biệt. Máy này làm mẫu dao động khoảng 20.000 lần mỗi giây. Dù dùng dao động nhanh như vậy, hoàn thành một thử nghiệm vẫn mất vài ngày.

Vì thử nghiệm mất nhiều thời gian, dữ liệu rất quý. Nếu dữ liệu ít, dự đoán sẽ dao động. Vì vậy, trí tuệ nhân tạo cần cách dùng dữ liệu tiết kiệm. Nó phải đưa ra dự đoán đáng tin ngay cả với ít dữ liệu. Bài toán dùng dữ liệu tiết kiệm này sẽ được bàn kỹ hơn ở mục sau.

Lựa chọn vật liệu cũng gắn với môi trường này. Một số kim loại trở nên mềm hoặc dễ vỡ ở nơi rất lạnh. Inconel 718 có tính chất tương đối ổn định ngay cả ở nhiệt độ cực thấp. Vì vậy, nó được dùng rộng rãi cho chi tiết nhiệt độ cực thấp. Dòng Haynes phát huy sức bền tốt hơn ở nhiệt độ cao hơn. Phải chọn vật liệu theo vị trí đặt chi tiết.

Ngay trong một chi tiết, điều kiện cũng khác nhau theo từng vị trí. Chân cánh chịu lực lớn. Đầu cánh chịu rung nhiều. Vì vậy, trước hết phải tìm vị trí yếu khó chịu lực. Vị trí này được gọi là tiết diện nguy hiểm. Bảo đảm mỗi lấy tiết diện nguy hiểm này làm chuẩn.

Tìm tiết diện nguy hiểm cũng cần tính toán. Ta giải lực tác dụng lên chi tiết bằng máy tính. Tính toán này được gọi là phân tích phần tử hữu hạn. Chi tiết được chia thành các mảnh nhỏ rồi tính lực từng phần. Vị trí tập trung lực lớn là tiết diện nguy hiểm. Trí tuệ nhân tạo đôi khi thay thế nhanh phép tính này.

Rung của bánh công tác có quy luật riêng. Số lần cánh đi qua cạnh chi tiết cố định đã được xác định. Số lần này nhân với tốc độ quay sẽ thành chu kỳ rung. Ở một tốc độ quay nào đó, rung này trùng với dao động riêng của chi tiết. Khi đó, dao động tăng mạnh. Sự trùng này được gọi là cộng hưởng. Cộng hưởng làm chi tiết hư hỏng nhanh. Vì vậy, người ta thiết kế để tránh tốc độ quay nguy hiểm.

Bảo đảm mỗi của bánh công tác là cốt lõi của an toàn tên lửa. Nếu một cánh vỡ, toàn bộ bơm sẽ vỡ theo. Khi đó, động cơ dừng hoặc phát nổ. Vì vậy, chi tiết này chịu kiểm tra nghiêm ngặt hơn bất kỳ chi tiết nào khác. Phải quản lý cả chất lượng bề mặt và khuyết tật bên trong.

#### **7.4.2 Phá hủy bắt đầu từ bên trong và hoa văn mắt cá**

Tiểu mục này bàn về cách mỗi chu kỳ rất cao làm vỡ chi tiết. Ở trên đã nói rằng quy luật của vùng này khác. Bây giờ ta xem khác như thế nào một cách cụ thể. Ta cũng xem hoa văn lặn còn lại trên mặt gậy.

Mỗi thông thường bắt đầu từ bề mặt. Bề mặt là nơi dễ bị xước. Khi chịu tải lặp lại, trên bề mặt xuất hiện vết trượt. Vết nứt bắt đầu từ vết đó. Vì vậy, nếu hoàn thiện bề mặt cho mịn, tuổi thọ sẽ tăng.

Mỗi chu kỳ rất cao thì khác. Vết nứt bắt đầu không phải ở bề mặt, mà ở bên trong. Khi số lần lặp rất lớn, dù lực yếu, khuyết tật bên trong cũng không chịu nổi. Lỗ nhỏ hoặc hạt cứng bên trong trở thành hạt mầm. Bọt argon nhỏ mà HIP cũng không loại bỏ được là ví dụ nguy hiểm.

Quanh hạt mầm của vết nứt xuất hiện một vùng khác thường. Xung quanh hạt mầm, biến dạng vi mô tích lũy qua hàng trăm triệu lần. Vết nứt bắt đầu từ bên trong lớn lên ở nơi không tiếp xúc với không khí. Hai mặt của khe hẹp cọ xát nhau và nghiền cấu trúc thành các hạt rất nhỏ. Vị trí bị nghiền nhỏ như vậy được gọi là vùng hạt mịn (Fine Granular Area, FGA). Khi nhìn bằng kính hiển vi, vùng này trông tối hơn khu vực xung quanh. Vùng tối này nằm ở giữa dạng mắt cá.

Vùng tối này cần thời gian dài để hình thành. Phần lớn tuổi thọ mỗi tổng thể được dùng ở giai đoạn này. Giai đoạn trước khi vết nứt lớn đến mức nhìn thấy được là giai đoạn dài. Bên ngoài trông vẫn nguyên vẹn, nhưng bên trong hư hại đã tiến triển.

Điểm này có ý nghĩa đối với việc xác định thời điểm kiểm tra. Hư hại lộ ra bên ngoài muộn. Vì vậy, không thể yên tâm chỉ bằng cách nhìn bên ngoài. Phải nhìn vào bên trong theo chu kỳ đã định. Với tên lửa tái sử dụng, quản lý chu kỳ này càng quan trọng. Trước khi dùng lại chi tiết đã dùng một lần, phải kiểm tra bên trong.

Khi vết nứt lớn lên từ hạt mầm, nó để lại hoa văn tròn. Hoa văn này trông giống mắt cá. Vì vậy, nó được gọi là dạng mắt cá (fish-eye). Dạng mắt cá là bằng chứng cho thấy vết nứt bắt đầu từ bên trong. Nhìn mặt gãy, ta biết nguyên nhân nhờ hoa văn này.

Cuối cùng, vết nứt chạm tới bề mặt. Ngay khoảnh khắc đó, chi tiết đột ngột gãy. Vết nứt vốn lớn lên chậm bỗng bùng ra trong một lần. Kiểu phá hủy đột ngột khó nhận biết trước như vậy nguy hiểm hơn cả.

Cách phá hủy này có ý nghĩa lớn đối với kiểm tra. Nếu chỉ nhìn bề mặt, ta bỏ sót nguy cơ. Phải nhìn vào bên trong. Vì vậy, người ta dùng chung CT tia X (X-ray CT) và kiểm tra siêu âm. Các phép kiểm tra này nhìn vào bên trong mà không phá hủy chi tiết. Kiểm tra như vậy được gọi là kiểm tra không phá hủy (non-destructive testing).

Trí tuệ nhân tạo bắt tay với các phép kiểm tra này. Nó lấy kích thước và vị trí khuyết tật bên trong tìm được bằng CT làm đầu vào. Bằng các giá trị này, nó dự đoán tuổi thọ mỗi. Nó xếp hạng khuyết tật nào nguy hiểm. Nghiên cứu gần đây đã đề xuất phương pháp tìm khuyết tật nguy hiểm trong chi tiết chế tạo bồi đắp bằng cách không phá hủy. Máy móc chỉ ra trước những khuyết tật bên trong mà con người dễ bỏ sót.

Khó loại bỏ hoàn toàn khuyết tật bên trong. Vì vậy, ta thiết kế sao cho vẫn an toàn dù có khuyết tật. Cách nghĩ này được gọi là thiết kế dung sai khuyết tật. Ta xác định trước chi tiết chịu được khuyết tật lớn đến mức nào. Chi tiết có khuyết tật lớn hơn mức đó sẽ bị loại bỏ. Trí tuệ nhân tạo tạo ra dữ liệu để xác định tiêu chuẩn này.

Cách làm này cũng gắn với tiêu chuẩn kiểm tra. Thiết bị kiểm tra không thể phát hiện đến những khuyết tật cực nhỏ. Khuyết tật có thể phát hiện luôn có một giới hạn kích thước nhất định. Thiết kế phải làm sao để khuyết tật vượt quá giới hạn này vẫn an toàn. Khi đó, cả khuyết tật bị bỏ sót trong kiểm tra cũng không gây nguy hiểm. Như vậy, kiểm tra và thiết kế bù vào khoảng trống của nhau để giữ an toàn.

## 7.5 Bảo đảm độ tin cậy mỗi bằng cách dự đoán cả độ dao động

Tuổi thọ mỗi không rơi gọn vào một con số. Ngay cả các chi tiết trong cùng điều kiện cũng có tuổi thọ khác nhau. Vì vậy, dự đoán phải chứa cả độ rộng của dao động. Mục này giải thích trí tuệ nhân tạo xử lý dao động đó. Ta sẽ xem cách biến khuyết tật thành con số, mô hình chứa dao động, và cách nối kết kết quả với tiêu chuẩn an toàn.

### 7.5.1 Biến trạng thái vật liệu thành con số

Tiểu mục này bàn về cách biến trạng thái vật liệu thành con số. Trí tuệ nhân tạo học bằng các con số. Vì vậy, phải biến khuyết tật và bề mặt thành con số. Những con số này được gọi là descriptor. Ở đây, ta sẽ lần lượt xem những con số nào quan trọng đối với tuổi thọ mỏi.

Trước hết, hình dạng của khuyết tật được đưa vào con số. Khi chụp bên trong chi tiết bằng CT, ta thu được hình dạng ba chiều của lỗ rỗng. Hình dạng này được sắp xếp thành tensor. Như đã nói ở trên, tensor là một chiếc hộp chứa các giá trị theo nhiều hướng. Nó chứa thông tin lỗ rỗng kéo dài theo hướng nào, và dẹt đến mức nào.

X-ray CT có nguyên lý giống CT trong bệnh viện. Nó bắn tia X từ nhiều hướng để nhìn vào bên trong. Sau đó gom các thông tin đó lại để tạo thành hình ảnh ba chiều. Ta nhìn được lỗ rỗng bên trong mà không phải cắt chi tiết. Từng kích thước và hình dạng của lỗ rỗng được đo. Thông tin kích thước và hình dạng thu được như vậy trở thành dữ liệu nền cho dự đoán mỏi.

Kích thước khuyết tật cũng quan trọng. Một cách tốt để đo kích thước là phương pháp căn bậc hai diện tích Murakami (Murakami square root of area). Phương pháp này xem khuyết tật như một vết nứt nhỏ. Khuyết tật được chiếu ép lên mặt phẳng vuông góc với hướng tải để đo diện tích. Căn bậc hai của diện tích đó được lấy làm kích thước khuyết tật. Lý do là căn bậc hai này liên hệ tốt với cường độ của lực làm mở vết nứt. Vì vậy, một lỗ rỗng ba chiều phức tạp được chuyển thành một giá trị này để xử lý. Giá trị này càng lớn thì ảnh hưởng đến mỏi càng xấu.

Vị trí của khuyết tật cũng được xem xét cùng lúc. Dù cùng kích thước, khuyết tật gần bề mặt vẫn nguy hiểm hơn. Khuyết tật bề mặt có ít vật liệu xung quanh giữ lại hơn, nên lực tập trung mạnh hơn. Tác động của môi trường tiếp xúc với không khí cũng làm vết nứt lớn nhanh hơn một chút. Khuyết tật nằm sâu bên trong được vật liệu bao quanh bốn phía nên ít nguy hiểm hơn. Vì vậy, giá trị hiệu chỉnh được đặt khác nhau tùy theo vị trí.

Độ nhám bề mặt cũng được biến thành con số. Trên bề mặt đã gia công cắt gọt có các rãnh và đỉnh rất nhỏ. Rãnh càng sâu và sắc thì lực càng tập trung. Mức độ tập trung này được gọi là hệ số tập trung ứng suất (stress concentration factor). Nó giống như khi cắt giấy, chỉ cần rạch nhẹ một vết là giấy dễ rách hơn. Vì một vết xước nhỏ gom lực đang phân tán vào một điểm và làm vị trí đó yếu đi.

Những con số này gom lại và nối với tuổi thọ mỏi. Có một đồ thị biểu diễn quan hệ giữa độ lớn của lực và số lần lặp. Đồ thị này gọi là đường cong S-N. Thông tin về khuyết tật và độ nhám làm đường cong này dịch lên hoặc xuống. Khi khuyết tật lớn và bề mặt nhám, đường cong đi xuống. Đường cong đi xuống nghĩa là chi tiết sẽ gãy nhanh hơn ngay cả dưới cùng một lực.

Các nghiên cứu gần đây đã đưa các descriptor này vào trí tuệ nhân tạo và đạt kết quả. Một nghiên cứu đã dự đoán tuổi thọ mỏi của kim loại đắp lớp bằng thông tin khuyết tật đo bằng CT. Số nghiên cứu áp dụng phương pháp Murakami cho nhiều trường hợp đắp lớp cũng đang tăng lên. Người ta liên tục xác nhận rằng kích thước, vị trí và hình dạng khuyết tật ảnh hưởng lớn đến tuổi thọ.

Độ cứng của vật liệu cũng được đưa vào descriptor. Độ cứng cho biết vật liệu cứng đến mức nào. Trong công thức Murakami, độ cứng càng cao thì vật liệu càng được xem là chịu khuyết tật tốt hơn. Nhưng chỉ riêng độ cứng không thể giải thích hết khả năng chịu khuyết tật. Vật liệu quá cứng đôi khi lại nhạy hơn với khuyết tật. Vì vậy, phải xem cùng lúc độ cứng, kích thước khuyết tật, vị trí và tỷ số tải.

Điều kiện tải cũng được đưa vào dưới dạng con số. Lực lớn đến đâu và theo hướng nào là điều quan trọng. Lực chỉ đến từ một hướng hay trộn từ nhiều hướng cũng khác nhau. Nhiệt độ cũng được đưa vào. Vì tính chất của vật liệu khác nhau ở nơi nóng và nơi lạnh.

Nghiên cứu gần đây đã đạt kết quả tốt với các descriptor như vậy. Một nghiên cứu đã dự đoán tuổi thọ mỗi chu kỳ cao của Inconel 718 đắp lớp. Kích thước khuyết tật và độ lớn tải được dùng làm đầu vào. Vấn đề thiếu dữ liệu được giải bằng cách tăng dữ liệu bằng trí tuệ nhân tạo. Kết quả là hệ số xác định của dự đoán tăng lên đến 0.975. Sai số trung bình chỉ là 1.13 phần trăm. Hệ số xác định là giá trị cho thấy dự đoán khớp với thực tế đến mức nào. Giá trị này càng gần 1 thì dự đoán càng khớp với thực tế. Tuy nhiên, con số này đến từ dữ liệu giới hạn của một vật liệu. Vì là kết quả thu được từ dữ liệu đã được tăng thêm, nên khó áp dụng nguyên xi cho quy trình và vật liệu khác.

Tuy vậy, việc quyết định descriptor không dễ. Cần phán đoán nên đưa con số nào vào và bỏ con số nào ra. Nếu bỏ sót con số quan trọng, dự đoán sẽ lệch. Nếu đưa vào con số vô dụng, việc học sẽ bị nhiễu. Vì vậy, người hiểu rõ vật liệu và trí tuệ nhân tạo phải cùng chọn.

### 7.5.2 Mạng nơ-ron Bayes (BNN) đưa ra độ rộng của đáp án

Tiểu mục này bàn về trí tuệ nhân tạo chứa dao động. Mạng nơ-ron thông thường chỉ đưa ra một đáp án. Trong dự đoán mỗi, như vậy là chưa đủ. Vì ngay cả trong cùng điều kiện, tuổi thọ vẫn phân tán rộng. Do đó cần mô hình đưa ra cả phạm vi của đáp án.

Mô hình như vậy được gọi là mạng nơ-ron Bayes (Bayesian Neural Network, BNN). Trong mạng nơ-ron thông thường, giá trị bên trong được cố định thành một giá trị. Mạng nơ-ron Bayes xử lý giá trị bên trong như một phân bố. Nó không có một giá trị duy nhất, mà có một độ rộng giá trị. Vì vậy, đầu ra dự đoán cũng không phải một con số mà là một độ rộng.

Ở đây, từ Gaussian thường xuất hiện. Gaussian nghĩa là phân bố hình chuông. Đó là hình dạng cao ở giữa và thấp dần về hai bên. Có thể dễ hiểu nếu nghĩ đến phân bố điểm thi. Nhiều người ở gần điểm trung bình, và càng về hai đầu càng ít. Tuy nhiên, bản thân tuổi thọ mỗi thường không có dạng hình chuông. Nó phân tán theo các dạng khác như lognormal hoặc Weibull. Vì vậy, người ta lấy log của tuổi thọ rồi đặt giả định hình chuông lên giá trị đó. Giả định hình chuông này cũng là một cách thuận tiện để tính toán dễ hơn.

Trong phân bố này có hai giá trị quan trọng. Một là trung bình. Trung bình là giá trị đại diện. Giá trị còn lại là độ rộng phân tán. Độ rộng lớn nghĩa là dao động lớn. Mạng nơ-ron Bayes đưa ra hai giá trị này cùng lúc.

Hãy ví dụ bằng dự báo thời tiết để hiểu vì sao dự đoán có độ rộng. Dự báo nói xác suất mưa ngày mai là 70 phần trăm. Nó không nói dứt khoát là 100 phần trăm. Vì tương lai có dao động. Mạng nơ-ron Bayes cũng cho biết độ rộng xác suất thay vì chỉ một đáp án.

Cách học của mô hình này cũng đặc biệt. Vì giá trị bên trong là phân bố nên tính toán phức tạp. Tính toán chính xác trên thực tế là không thể. Vì vậy, người ta dùng phương pháp thay bằng một phân bố gần đúng. Quá trình đổi sang phân bố gần đúng để giảm lượng tính toán này là cốt lõi của việc học.

Khi học xong, ta rút dự đoán nhiều lần. Vì giá trị bên trong là phân bố, mỗi lần rút lại cho ra đáp án hơi khác nhau. Gộp các đáp án này lại, ta tính được trung bình và độ rộng. Nếu các đáp án giống nhau, độ rộng hẹp. Nếu các đáp án phân tán, độ rộng lớn. Kích thước của độ rộng này cho biết có thể tin dự đoán đến mức nào.

Cách làm này giống phán đoán của con người. Kỹ sư giàu kinh nghiệm cũng phân biệt giữa chắc chắn và phỏng đoán. Với chi tiết mình biết rõ, họ nói tự tin. Với chi tiết lần đầu thấy, họ nói thận trọng. Mạng nơ-ron Bayes cũng truyền đạt cả mức độ chắc chắn của mình như vậy. Vì thế, con người dễ nhìn mức độ chắc chắn đó để tin và dùng kết quả.

Cách tiếp cận này còn có một điểm tốt nữa. Khi mô hình gặp điều kiện chưa từng thấy, nó đưa ra độ rộng lớn. Nó tự gửi tín hiệu rằng mình không biết rõ. Con người nhìn tín hiệu này và thận trọng. Dự đoán có

độ rộng lớn cần được kiểm tra thêm trước khi dùng. Việc mô hình báo cho biết giới hạn của chính nó giúp ích lớn cho an toàn.

Cách tiếp cận này phù hợp với dự đoán mỗi. Mỗi là hiện tượng có dao động lớn. Độ rộng trung thực hơn một con số. Nhiều nghiên cứu gần đây đi theo hướng này. Đã có nghiên cứu dự đoán tuổi thọ mỗi của kim loại hàng không kèm theo dao động. Cũng có nghiên cứu đưa thêm định luật vật lý vào để tăng độ tin cậy.

Vì sao đưa định luật vật lý vào lại tốt? Trong mỗi có những lý thuyết đã tích lũy lâu dài. Đó là các công thức chứa quan hệ giữa lực và tuổi thọ. Các công thức này được đưa cùng vào trí tuệ nhân tạo. Khi đó, dù dữ liệu ít, đáp án vẫn không lệch quá xa. Cũng đã có nghiên cứu xử lý cùng lúc tải theo nhiều hướng. Dòng nghiên cứu này làm cho bảo đảm tuổi thọ chi tiết tên lửa vững hơn.

Phương pháp này cũng có chi phí. Vì phải rút đáp án nhiều lần nên tính toán tăng lên. Việc học cũng khó hơn mạng nơ-ron thông thường. Dù vậy, với chi tiết tên lửa, chi phí này không đáng tiếc. Vì biết được dao động là cách giữ an toàn. Nếu chỉ tin một con số rồi xảy ra tai nạn, thiệt hại sẽ lớn hơn nhiều.

### 7.5.3 Tách dao động và nối với tiêu chuẩn an toàn

Tiểu mục này bàn về cách tách dao động để xem xét. Ở trên đã nói dự đoán có độ rộng. Độ rộng này không chỉ có một loại. Hai loại dao động trộn lẫn trong đó. Khi tách hai loại này ra, ta thấy phải làm gì tiếp theo.

Thứ nhất là bất định nhận thức (epistemic uncertainty). Đây là dao động sinh ra vì mô hình không biết rõ. Nó xuất hiện lớn trong các điều kiện thiếu dữ liệu. Nếu gặp tình huống chưa từng học, mô hình thiếu tự tin. Dao động này dần giảm khi bổ sung thêm dữ liệu còn thiếu.

Thứ hai là bất định ngẫu nhiên (aleatoric uncertainty). Đây là dao động sinh ra từ chính vật liệu và tải. Ngay cả các chi tiết trong cùng điều kiện cũng có khuyết tật bên trong khác nhau. Phép đo cũng có nhiễu. Dao động này nằm trong bản thân hiện tượng, nên dù tăng dữ liệu cũng khó giảm.

Vì sao việc tách hai loại này lại quan trọng? Vì cách xử lý khác nhau. Nếu dao động nhận thức lớn, cần làm thêm thí nghiệm. Nếu dao động ngẫu nhiên lớn, cần đặt thêm biên an toàn. Nếu gộp hai loại lại, ta dễ đưa ra cách xử lý sai. Khi tách hai dao động ra để xem, việc tiếp theo cần làm sẽ rõ hơn.

Bây giờ ta nối dự đoán với tiêu chuẩn an toàn. Chi tiết tên lửa phải vượt qua tiêu chuẩn an toàn đã định. Tiêu chuẩn này chứa hai yếu tố cùng lúc. Một là bao nhiêu phần trăm chi tiết sống sót. Hai là có thể tin giá trị đó đến mức nào. Ví dụ, ta đặt giới hạn dưới cho trường hợp 90 phần trăm sống sót và tin điều đó ở mức 95 phần trăm. Giá trị được xác định cùng lúc theo tỷ lệ sống sót và mức tin cậy như vậy được gọi là giá trị cho phép của vật liệu. Giá trị chuẩn A và giá trị chuẩn B dùng trong vật liệu hàng không là những ví dụ như vậy. Chỉ lấy trung bình trừ đi phần dao động thì không bảo đảm được giá trị này. Phải đưa cả số lượng dữ liệu thử nghiệm và độ phân tán vào để tính bằng thống kê.

Giới hạn dưới này được gọi là giá trị cho phép thiết kế. Chi tiết phải chịu được lâu hơn giá trị cho phép này. Đôi khi người ta còn nhân thêm một hệ số an toàn. Hệ số an toàn là phần dư để phòng nguy cơ bất ngờ. Bằng cách đặt nhiều lớp dư như vậy, người ta ngăn tai nạn. Nếu phần dư quá lớn, chi tiết sẽ nặng hơn, nên phải tìm điểm cân bằng.

Làm như vậy sẽ khiến phán đoán nghiêng về phía an toàn. Ở điều kiện mà mô hình không tự tin, tuổi thọ được đặt thấp hơn. Ở điều kiện mà mô hình tự tin, ta có thể cho thêm một chút dư địa. Thái độ này phù hợp với chi tiết tên lửa. Dù đặt hơi bảo thủ, ngăn được tai nạn vẫn tốt hơn.

Chỉ riêng trí tuệ nhân tạo không thể hoàn tất chứng nhận. Bảo đảm mỗi cần nhiều công cụ cùng lúc. Dự đoán bằng trí tuệ nhân tạo, thử nghiệm mỗi tiêu chuẩn, kiểm tra không phá hủy và hệ số an toàn đều

cần thiết. Phương pháp thử tiêu chuẩn cũng đã được quy định. Có tiêu chuẩn quốc tế về cách đặt lực và cách phân tích dữ liệu.

Vai trò của trí tuệ nhân tạo rất rõ. Nó nhanh chóng tìm các ứng viên nguy hiểm. Nó cho biết cần thử nghiệm thêm ở đâu. Nó giúp thu được nhiều thông tin bằng ít thí nghiệm. Con người dùng thông tin đó để lập kế hoạch thử nghiệm tốt hơn, qua đó giảm mạnh thời gian và chi phí.

Phán định cuối cùng do con người và thí nghiệm đưa ra. Dự đoán của trí tuệ nhân tạo là điểm bắt đầu, không phải điểm kết thúc. Dữ liệu thử nghiệm thực tế phải chống đỡ cho dự đoán. Tài liệu chất lượng phải ghi lại quá trình đó. Chỉ khi mọi thứ này đầy đủ, chi tiết mới được đưa lên tên lửa.

Vẫn còn những nhiệm vụ phía trước. Dữ liệu mỗi vẫn còn thiếu. Dữ liệu ở vùng siêu cao chu kỳ còn ít hơn nữa. Vì thử nghiệm mất nhiều thời gian. Do đó, các phương pháp học từ ít dữ liệu vẫn tiếp tục được nghiên cứu. Cũng có cách chuyển dữ liệu của vật liệu khác sang dùng. Khi các phương pháp này phát triển, dự đoán sẽ vững hơn.

Chất lượng dữ liệu cũng quan trọng. Cách đo khuyết tật khác nhau giữa các bộ dữ liệu. Cách làm thử nghiệm cũng không giống nhau. Nếu ghi chép sơ sài, mô hình sẽ học sai. Vì vậy, cần có văn hóa ghi chép dữ liệu cẩn thận. Phải có dữ liệu tốt được ghi chép tốt thì mới có mô hình tốt.

Tiêu chuẩn và chứng nhận cũng phải phát triển cùng lúc. Các quy tắc chứng nhận hiện nay được xây dựng dựa trên vật liệu trước đây. Vật liệu chế tạo bồi đắp có tính chất khác. Vì vậy cần có quy tắc mới. Người ta cũng đang thảo luận cách đưa dự báo của trí tuệ nhân tạo vào chứng nhận. Khi cuộc thảo luận này được sắp xếp rõ ràng, linh kiện chế tạo bồi đắp sẽ được dùng rộng rãi hơn.

## Tổng kết

Trong chương này, chúng ta đã xử lý hai vấn đề lớn. Một là gia công cắt gọt tốt các kim loại khó cắt. Hai là biết linh kiện chịu được trong bao lâu. Cả hai vấn đề đều xuất hiện ở linh kiện tên lửa được làm bằng in 3D. Và trong cả hai vấn đề, trí tuệ nhân tạo đều giúp ích lớn.

Thật ra hai vấn đề này nối liền với nhau. Phải cắt gọt tốt thì bề mặt mới mịn. Bề mặt mịn thì chịu mỏi tốt. Phải loại bỏ lỗ rỗng bên trong thì mầm nứt mới giảm. Cuối cùng, gia công tốt dẫn đến tuổi thọ dài. Vì vậy phải luôn xem gia công và mỏi cùng nhau. Chương này đặt hai vấn đề gia công và mỏi vào cùng một chỗ để xử lý chung.

Ở phần đầu, chúng ta đã xem xét vì sao siêu hợp kim chế tạo bồi đắp khó cắt gọt. Tổ chức nguội nhanh có tính chất khác nhau theo hướng. Bên trong có ứng suất bị giữ lại, và các hạt cứng tụ lại. Những tính chất này làm dụng cụ mòn nhanh, đồng thời để lại lớp trắng và vết nứt vi mô trên bề mặt.

Chúng ta cũng đã thấy con đường vượt qua khó khăn này. Tính trước lực cắt và dao động để chọn điều kiện an toàn. Dùng ép đẳng tĩnh nóng để giảm lỗ rỗng bên trong và làm tổ chức đồng đều. Dùng phương pháp Taguchi và mạng nơ-ron nhân tạo để tìm điều kiện tốt từ ít thí nghiệm. Trong quá trình này, trí tuệ nhân tạo nhanh chóng thu hẹp các phương án ứng viên.

Ở phần sau, chúng ta đã xử lý vấn đề mỏi. Cánh của bơm tuabin tên lửa nằm trong vùng mỏi chu kỳ cực cao. Ở vùng này, vết nứt bắt đầu từ bên trong chứ không phải từ bề mặt. Trên mặt gãy còn lại hoa văn tròn gọi là dạng mắt cá. Vì vậy cần kiểm tra không phá hủy để nhìn vào bên trong mà không làm vỡ linh kiện.

Dự báo mỏi phải chứa cả độ dao động. Vì ngay cả các linh kiện trong cùng điều kiện cũng có tuổi thọ khác nhau. Mạng nơ-ron Bayes không đưa ra một đáp án duy nhất, mà đưa ra một khoảng đáp án. Nếu chia khoảng này thành hai loại, ta sẽ thấy việc cần làm tiếp theo. Có thể quyết định nên tăng dữ liệu hay tăng biên an toàn.

Điểm cốt lõi là sự hợp tác giữa trí tuệ nhân tạo và con người. Trí tuệ nhân tạo đọc dữ liệu nhanh và chỉ ra rủi ro. Nó giúp giảm kế hoạch thí nghiệm một cách thông minh. Nhưng chúng nhận cuối cùng do thí nghiệm và kiểm tra đảm nhận. Linh kiện tên lửa được bảo vệ bằng cả hai sức mạnh này. Sự hợp tác giữa trí tuệ nhân tạo và thí nghiệm trở thành nền tảng nâng đỡ chuyến bay an toàn.

Hãy rút gọn nội dung chương này thành một câu. Phải cắt gọt tốt, hoàn thiện tốt, và kiểm tra tốt. Rồi phải dự báo kết quả đó cùng với độ dao động. Trí tuệ nhân tạo giúp mọi bước này diễn ra nhanh hơn. Con người nhận sự trợ giúp đó để làm ra linh kiện an toàn hơn. Tương lai của linh kiện tên lửa in 3D cuối cùng phụ thuộc vào sự hợp tác này.

### Tài liệu căn cứ liên quan

- 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. (Bản tiền ấn phẩm, chưa qua bình duyệt) <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.res.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. (Bản tiền ấn phẩm, chưa qua bình duyệt) <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>

## Chương 8. Mô hình hóa dòng khí phản ứng và hóa học cháy bằng học máy dựa trên vật lý (PIML/PINNs) và học toán tử (FNO)

### Mở đầu

Khi bật bếp gas trong nhà bếp, ngọn lửa xanh hiện lên. Khí đốt là nhiên liệu gặp oxy trong không khí và sinh nhiệt. Quá trình này gọi là sự cháy. Sự cháy có ở khắp nơi trong đời sống của chúng ta. Động cơ ô tô và nồi hơi cũng vận hành nhờ sự cháy.

Động cơ tên lửa cũng tạo lực theo cùng nguyên lý. Nhưng quy mô và điều kiện khác rất nhiều. Bên trong buồng cháy tên lửa nóng hơn nhiều và có áp suất cao hơn nhiều. Nhiên liệu và oxy trộn lẫn rồi phản ứng với tốc độ rất nhanh. Trong đó, khí chảy, xoáy và phát ra âm thanh.

Rất khó tính toán tất cả những việc này cùng một lúc. Phản ứng hóa học xảy ra trong chớp mắt. Ngược lại, dòng chảy đi qua toàn bộ buồng cháy chậm hơn thế. Biến đổi nhanh và biến đổi chậm trộn lẫn trong cùng một không gian. Nếu máy tính phải theo dõi cả hai cùng lúc, khối lượng tính toán tăng rất lớn.

Cho đến nay, người ta giải bài toán bằng cách chia phương trình ra rất nhỏ. Phương pháp này gọi là động lực học chất lưu tính toán (CFD). Nó chính xác nhưng chi phí tính toán lớn. Muốn giải chi tiết một buồng cháy tên lửa lớn, siêu máy tính phải chạy trong nhiều ngày. Các nhà thiết kế đã mong có công cụ nhanh hơn từ lâu.

Trí tuệ nhân tạo giúp ích ở điểm này. Một mạng nơ-ron được huấn luyện tốt giúp giảm nhanh khối lượng tính toán. Khi có ít cảm biến, nó cũng có thể lấp đầy phần thông tin còn thiếu. Nhưng trong bài toán cháy, nếu trí tuệ nhân tạo tự ý đưa ra đáp án thì rất nguy hiểm. Trong tên lửa, một sai số nhỏ cũng có thể dẫn đến tai nạn lớn.

Vì vậy, phương pháp đưa trực tiếp định luật vật lý vào quá trình học đã xuất hiện. Phương pháp này gọi là học máy dựa trên vật lý (PIML). Mạng nơ-ron không chỉ bắt chước dữ liệu. Nó còn được tạo ra để cùng tuân thủ các định luật như bảo toàn khối lượng hay bảo toàn năng lượng. Phương pháp tiêu biểu là mạng nơ-ron dựa trên vật lý (PINN).

Chương này bàn về nhiều phương pháp của học dựa trên vật lý đó. Trước hết, chúng ta diễn giải bằng lời các phương trình chi phối sự cháy. Tiếp theo, chúng ta xem xét khó khăn tính toán do phản ứng hóa học nhanh gây ra. Sau đó, chúng ta xem một cấu trúc mới gọi là mạng Kolmogorov Arnold (KAN). Rồi chúng ta giới thiệu toán tử nơ-ron Fourier (FNO), vốn biến hàm thành hàm. Cuối cùng, chương kết thúc bằng một ví dụ khôi phục bên trong động cơ detonation quay.

Ở đây hãy thử dùng một phép so sánh. Có thể nghĩ đến hai cách vẽ bản đồ. Một cách là đi từng bước trên khắp mặt đất để đo. Cách này chính xác nhưng mất nhiều thời gian. Cách kia là chỉ đo vài nơi rồi lấp đầy phần còn lại. Cách này nhanh, nhưng nếu lấp đầy không theo quy tắc thì sẽ thành bản đồ sai lệch. Học dựa trên vật lý là đưa quy luật tự nhiên vào việc lấp đầy đó.

Sau khi đọc chương này, bức tranh lớn sẽ hiện rõ. Ta sẽ thấy trí tuệ nhân tạo đang thay đổi tính toán cháy như thế nào. Ta cũng sẽ phân biệt được điều gì đã làm được và điều gì vẫn còn khó. Công cụ mới có sức mạnh lớn, nhưng không phải vạn năng. Nó chỉ đáng tin khi đi cùng vật lý và kiểm chứng. Việc giữ cân bằng đó là ý tưởng xuyên suốt chương này.

Cuốn sách này xử lý cẩn trọng các tên gọi và con số khó. Trong các nghiên cứu gần đây, có những trường hợp chỉ có cái tên nghe có vẻ thuyết phục. Cũng có các con số hiệu năng chưa được kiểm chứng. Vì vậy, chương này chỉ chọn và giải thích các nghiên cứu đã được xác nhận. Những con số chưa kiểm chứng sẽ

không được dùng trong phần chính. Nghiên cứu ở giai đoạn đầu sẽ được nói rõ là nghiên cứu ở giai đoạn đầu.

## 8.1 Sự chuyển đổi mô hình trong kỹ thuật tính toán cháy và trí tuệ nhân tạo

### Nội dung của mục này

Mục này bắt đầu từ lý do vì sao tính toán cháy lại khó. Chúng ta xem những gì diễn ra bên trong buồng cháy tên lửa phức tạp đến mức nào. Rồi chúng ta xem giới hạn của các phương pháp tính toán truyền thống. Mục này giải thích vì sao trí tuệ nhân tạo xuất hiện để vượt qua những giới hạn đó.

Ở đây chúng ta không dùng công thức khó. Thay vào đó, chúng ta chuyển khái niệm sang các cảnh trong đời sống. Nếu hiểu mục này, mọi mục phía sau sẽ dễ hơn. Vì toàn bộ chương này cuối cùng đều là các phương pháp khác nhau để giải bài toán này.

Bên trong buồng cháy tên lửa là một môi trường cực hạn. Nhiệt độ tăng lên tới hàng nghìn độ. Áp suất đạt tới hàng trăm lần áp suất khí quyển. Nhiên liệu và chất oxy hóa phản ứng dữ dội trong thời gian ngắn. Động cơ dùng methane và oxy lỏng là một ví dụ tiêu biểu.

Phản ứng như vậy phức tạp hơn nhiều so với vẻ bên ngoài. Con đường để methane gặp oxy rồi thành nước và carbon dioxide không chỉ có một. Ở giữa xuất hiện hàng chục loại chất trung gian. Những chất này gọi là gốc tự do và có hoạt tính phản ứng lớn. Vì vậy, hàng trăm phản ứng phụ đan xen với nhau và diễn ra cùng lúc.

Hãy so sánh gốc tự do với cuộc chạy tiếp sức. Gậy chuyển tay được trao rất nhanh từ tay này sang tay khác. Không vận động viên nào giữ gậy lâu. Gốc tự do cũng sinh ra trong chớp mắt rồi chuyển sang chất tiếp theo. Sự trao đổi bận rộn này dẫn dắt quá trình cháy. Nếu bỏ lỡ tốc độ đó, ta không thể vẽ đúng hình dạng ngọn lửa.

Thêm vào đó còn có dòng rối. Dòng rối là dòng chảy trong đó khí xoáy và trộn lẫn hỗn loạn. Nó giống hình ảnh nước sông gặp đá rồi vỡ trắng xóa. Khí trong buồng cháy tên lửa cũng trộn lẫn như vậy. Dòng rối làm nhiên liệu và oxy gặp nhau nhanh. Nhưng chính chuyển động bất quy tắc đó làm tính toán khó hơn.

Muốn giải chính xác phản ứng này, phải chia không gian và thời gian ra rất nhỏ. Nơi ngọn lửa dựng thành lớp mỏng có bề dày còn mỏng hơn sợi tóc. Trong lớp đó, nhiệt độ và nồng độ thay đổi đột ngột. Nếu bỏ sót lớp mỏng này, phép tính sẽ sai lệch. Vì vậy, muốn chứa được cả lớp này, phải đặt lưới rất dày.

Hãy cảm nhận bài toán đa thang đo này bằng con số. Lớp mỏng nơi ngọn lửa đứng có kích thước vài chục micromet. Ngược lại, chiều dài toàn bộ buồng cháy là vài chục centimet. Chênh lệch kích thước lên tới hàng nghìn lần. Trong thời gian cũng xuất hiện khoảng cách tương tự. Muốn chứa toàn bộ phạm vi rộng này cùng lúc, số ô lưới tăng bùng nổ.

Các phương pháp tính toán truyền thống gồm mô phỏng số trực tiếp (DNS) và mô phỏng xoáy lớn (LES). Mô phỏng số trực tiếp giải tất cả, không bỏ sót cả những xoáy có kích thước rất nhỏ. Mô phỏng xoáy lớn chỉ giải các xoáy lớn và xấp xỉ các xoáy nhỏ. Cả hai đều chính xác nhưng khối lượng tính toán lớn. Nếu giải chi tiết một buồng cháy lớn, thời gian sẽ kéo dài.

Hãy chuyển câu nói chi phí tính toán lớn sang ngôn ngữ đời thường. Nó giống việc làm một bộ phim độ phân giải cao. Nếu vẽ tỉ mỉ từng cảnh, có thể mất nhiều tháng. Tính toán cháy cũng như vậy. Mỗi lần thay đổi điều kiện, phải vẽ lại từ đầu. Nhà thiết kế muốn thử hàng chục điều kiện. Nhưng trong thực tế, rất khó chờ nhiều ngày mỗi khi thay đổi điều kiện.

Ở đây trí tuệ nhân tạo mở ra một con đường mới. Khi mạng nơ-ron đã học tốt một lần, nó trả lời nhanh các bài toán tương tự. Nó trở thành mô hình thay thế (surrogate model) nhẹ, làm thay cho phép tính nặng. Mô hình thay thế là mô hình xấp xỉ nhanh bắt chước phép tính gốc. Nhà thiết kế có thể xem qua nhiều điều kiện trong thời gian ngắn.

Hãy so sánh mô hình thay thế với ứng dụng bản đồ. Dù không thật sự đi bộ trên đường, ứng dụng vẫn cho biết thời gian cần thiết. Nó so sánh vô số tuyến đường trong chớp mắt. Mô hình thay thế cũng vậy. Nó ước lượng kết quả mà không chạy toàn bộ phép tính thật. Nhờ đó, nhiều phương án thiết kế có thể được xem xét trong thời gian ngắn.

Tuy nhiên, ứng dụng bản đồ cũng sai trên con đường lạ. Mô hình thay thế cũng lệch khi ra ngoài phạm vi đã học. Vì vậy, ta không tin nguyên xi câu trả lời ước lượng. Trước quyết định quan trọng, phải kiểm tra lại bằng tính toán chính xác. Mô hình thay thế là la bàn giúp định hướng, không phải phán quyết cuối cùng. Thái độ này chảy xuyên suốt chương này.

Hãy giải thích thêm một chút về trạng thái siêu tới hạn. Dùng nước làm ví dụ sẽ dễ hiểu. Bình thường, nước sôi và chuyển thành hơi nước. Ranh giới giữa chất lỏng và chất khí rõ ràng. Nhưng nếu nâng áp suất và nhiệt độ lên rất cao, ranh giới đó mờ đi. Nhiên liệu trong buồng cháy áp suất cao của tên lửa ở trạng thái như vậy. Việc tính chính xác ranh giới mờ giữa chất lỏng và chất khí này rất khó.

Ví dụ thực tế cũng đã xuất hiện. DeepFlame là phần mềm đưa mạng nơ-ron vào để tính nhanh quá trình cháy siêu tới hạn. Siêu tới hạn chỉ trạng thái áp suất cao, nơi ranh giới giữa chất lỏng và chất khí trở nên mờ. Nhóm nghiên cứu đã tối ưu mã này để giải quá trình cháy rời rạc của methane và oxy ở quy mô lớn. Theo tài liệu công khai, quy mô tính toán đã được nâng tới mức hàng chục tỷ ô. Đây là phạm vi rộng hơn nhiều so với trước đây. Kết quả này là nghiên cứu giai đoạn đầu được trình bày tại một hội nghị điện toán hiệu năng cao năm 2025.

Trong ví dụ này, việc trí tuệ nhân tạo đảm nhiệm rất rõ. Tính toán phản ứng hóa học là phần mất nhiều thời gian. Mạng nơ-ron thay thế nhanh phép tính này. Nhưng các tính chất của chất lưu thật vẫn được giữ nguyên. Đây là nỗ lực nhằm có cả tốc độ và độ chính xác. Hướng đi như vậy đẩy nhanh tính toán buồng cháy tên lửa lớn.

Hãy xem vì sao tính toán quy mô lớn lại quan trọng. Động cơ tên lửa thật có hàng trăm kim phun. Kim phun là bộ phận phun nhiên liệu và oxy để trộn chúng. Ngọn lửa từ rất nhiều kim phun này ảnh hưởng lẫn nhau. Tính toán quy mô nhỏ khó chứa được tương tác này. Khi tính toán quy mô lớn mở ra, ta mới tiến gần hơn đến động cơ thật.

Có ba cách lớn để trí tuệ nhân tạo đi vào lĩnh vực cháy. Một là mô hình thay thế nhanh, làm thay phép tính mất nhiều thời gian. Cách khác là khôi phục toàn bộ trường từ các phép đo thưa. Cách còn lại là chẩn đoán và điều khiển trạng thái theo thời gian thực. Chương này tập trung vào hai cách đầu, còn cách điều khiển thứ ba sẽ được bàn ở chương sau.

Hãy lấy một ví dụ để thấy vì sao trí tuệ nhân tạo vi phạm vật lý lại nguy hiểm. Giả sử một mô hình đưa ra nhiệt độ thấp hơn thực tế. Nhà thiết kế kết luận rằng vách có thể chịu được. Nhưng nếu nhiệt độ thật cao hơn, vách sẽ nóng chảy. Sai số nhỏ dẫn đến tai nạn lớn. Vì vậy, trong tên lửa, việc tuân thủ vật lý gần bằng với an toàn.

Dĩ nhiên, giới hạn cũng rõ ràng. Trí tuệ nhân tạo thường sai khi ra ngoài phạm vi đã học. Tên lửa luôn xuất hiện điều kiện mới. Nếu không tuân thủ định luật vật lý, dự đoán trở nên nguy hiểm. Vì vậy, toàn bộ chương này bàn về cách buộc vật lý và học lại với nhau. Năm 2026 cũng đã có bài tổng quan về học dựa trên vật lý cho hệ đẩy hàng không vũ trụ. Dòng chảy này cho thấy lĩnh vực này đang phát triển nhanh.

## 8.2 Phương trình chi phối dòng phản ứng nén được đa thành phần và học dựa trên vật lý

### Nội dung của mục này

Mục này diễn giải bằng lời các phương trình cơ bản chi phối sự cháy. Nói đến phương trình nghe có vẻ khó. Nhưng ý nghĩa thì dễ hiểu. Đó là quy tắc rằng không có thứ gì tự nhiên biến mất hay tự nhiên sinh ra.

Tiếp đó, chúng ta xem cách đưa quy tắc này vào mạng nơ-ron. Chúng ta xem mạng nơ-ron dựa trên vật lý tuân thủ định luật như thế nào. Hai khái niệm này là bộ khung của chương. Mọi phương pháp phía sau đều tách ra từ hai khái niệm này.

### 8.2.1 Phương trình chi phối Navier-Stokes phản ứng

Phương trình Navier-Stokes là quy tắc mô tả chuyển động của chất lưu. Hiện tượng nước chảy và không khí thổi được giải bằng quy tắc này. Trong quá trình cháy của tên lửa, phải thêm cả phản ứng hóa học và nhiệt vào đó. Vì vậy, số phương trình cần giải cùng nhau tăng lên nhiều.

Với chất lưu thông thường không có phản ứng, chỉ cần ba cân bằng là khối lượng, động lượng và năng lượng. Nhưng chất lưu phản ứng như trong sự cháy có thêm bảo toàn loài hóa học. Vì vậy, các phương trình này chứa bốn định luật bảo toàn. Định luật thứ nhất là bảo toàn khối lượng. Lượng khí đi vào và lượng khí đi ra phải khớp nhau. Khí không tự sinh ra hay biến mất. Cân bằng khối lượng khớp đúng này trở thành nền tảng của mọi phép tính.

Tiếp theo là bảo toàn động lượng. Động lượng là đại lượng bằng khối lượng nhân với vận tốc. Khi nhận lực, động lượng thay đổi. Chênh lệch áp suất và độ nhớt tạo ra lực này. Đó là lý do khí trong tên lửa bị đẩy ra ngoài rất nhanh.

Hãy hình dung bảo toàn động lượng bằng chiếc xe đạp. Khi đạp bàn đạp, xe đạp tiến về phía trước. Lực đạp tạo ra thay đổi vận tốc. Trong tên lửa, chênh lệch áp suất trở thành lực đó. Độ nhớt là ma sát giữ các phần dòng chảy lại với nhau và làm chúng chậm đi. Hai lực là chênh lệch áp suất và độ nhớt này làm thay đổi vận tốc của khí.

Một định luật nữa là bảo toàn loài hóa học. Loài hóa học là các chất tham gia phản ứng như oxy, methane và nước. Khi phản ứng xảy ra, có chất giảm và có chất tăng. Tốc độ phản ứng quyết định lượng tăng giảm đó. Chất lưu phản ứng lập riêng một phương trình bảo toàn như vậy cho từng chất. Vì vậy, số phương trình tăng thêm bằng số loài hóa học. Cuối cùng là bảo toàn năng lượng. Nhiệt sinh ra từ phản ứng làm nóng khí và đẩy dòng chảy đi.

Có thể ví bảo toàn năng lượng với heo đất tiết kiệm. Khi khớp tiền bỏ vào và tiền đã dùng, ta biết số tiền còn lại. Trong sự cháy, phản ứng hóa học không tạo ra năng lượng mới. Năng lượng hóa học bị nhốt trong nhiên liệu chỉ chuyển thành nhiệt. Nhiệt đó làm nóng khí và sinh công rồi đi ra ngoài. Cũng có nhiệt thoát ra bằng bức xạ. Quy tắc khớp tất cả những dòng vào ra này là bảo toàn năng lượng.

Khái niệm nén được cũng quan trọng. Nén được nghĩa là mật độ khí thay đổi lớn theo áp suất. Buồng cháy tên lửa có áp suất cao và biến đổi mạnh. Vì vậy, không thể cố định mật độ. Điểm mật độ liên tục thay đổi này làm tính toán khó hơn.

Hãy xem vì sao bốn cân bằng này lại quan trọng đến vậy. Một phép tính làm vỡ cân bằng sẽ lập tức trở thành lời nói dối. Nếu khối lượng tự tăng lên, tức là lực đẩy không hề có đã được tạo ra. Nếu năng lượng rò rỉ, dự đoán nhiệt độ sẽ sai lệch. Trong thiết kế tên lửa, sai số như vậy sẽ dẫn thẳng đến tai nạn. Vì thế, trí tuệ nhân tạo cũng phải tuân thủ các cân bằng này.

Tốc độ phản ứng hóa học được xác định bằng định luật Arrhenius. Định luật Arrhenius là quy tắc nói rằng khi nhiệt độ tăng, phản ứng sẽ nhanh lên đột ngột. Nó giống như que diêm đã bén lửa thì lan ra trong chớp mắt. Tính chất đột ngột này trở thành gốc rễ của vấn đề tính toán sẽ bàn ở phần sau. Kỹ sư thiết kế tên lửa phải dự đoán chính xác phản ứng này để giữ an toàn.

Khuếch tán và truyền nhiệt cũng phải được giải cùng nhau. Khuếch tán là hiện tượng vật chất lan từ nơi có nồng độ cao sang nơi có nồng độ thấp. Nó giống như mùi hương trong bếp lan vào phòng. Truyền nhiệt là hiện tượng nhiệt di chuyển từ nơi nóng sang nơi lạnh. Trong buồng đốt, hai hiện tượng này xảy ra cùng lúc với phản ứng hóa học. Vì vậy, các phương trình đan vào nhau và được giải như một khối.

### 8.2.2 Hàm mất mát phức hợp dựa trên thông tin vật lý

Cốt lõi của mạng nơ-ron dựa trên vật lý nằm ở hàm mất mát. Hàm mất mát là thước đo xem mạng nơ-ron sai bao nhiêu. Mạng nơ-ron thông thường chỉ đo chênh lệch với dữ liệu. Mạng nơ-ron dựa trên vật lý cộng thêm sai số vật lý vào đó.

Cách hoạt động là như sau. Mạng nơ-ron nhận vị trí và thời gian làm đầu vào. Sau đó, nó đưa ra áp suất, nhiệt độ, vận tốc và nồng độ tại điểm đó. Các giá trị này phải thỏa mãn phương trình bảo toàn thì mới đúng. Nếu các giá trị này không thỏa mãn phương trình, chúng sẽ bị phạt tương ứng.

Muốn đo sai số vật lý này thì cần đạo hàm. Ở đây dùng vi phân tự động (automatic differentiation). Vi phân tự động là kỹ thuật tính đạo hàm bằng cách lần theo quá trình tính toán của mạng nơ-ron. Nó khác với sai phân hữu hạn, vốn ước lượng bằng chênh lệch của lưới. Vì vậy, nó không có sai số cắt cụt phát sinh trong sai phân hữu hạn. Tuy nhiên, sai số xấp xỉ của chính mạng nơ-ron hoặc sai số do chọn điểm mẫu vẫn còn. Thứ mà vi phân tự động loại bỏ là sai số sinh ra khi ước lượng đạo hàm.

Hãy xem vì sao cần đạo hàm. Phương trình bảo toàn xử lý tỷ lệ biến đổi. Nó xem áp suất thay đổi nhanh đến đâu, vận tốc biến đổi như thế nào. Tỷ lệ biến đổi này chính là đạo hàm. Ta dùng đạo hàm để tính sự thay đổi của các giá trị do mạng nơ-ron đưa ra. Sau đó kiểm tra xem tốc độ biến đổi thu được có thỏa mãn phương trình hay không.

Tính toán truyền thống xấp xỉ đạo hàm trên lưới. Nó ước lượng sự thay đổi bằng chênh lệch giữa các điểm lưới lân cận. Cách này có sai số lớn khi lưới thô. Vi phân tự động không dùng xấp xỉ này. Nó lấy đạo hàm trực tiếp từ biểu thức của mạng nơ-ron. Vì vậy, không cần đặt riêng một lưới dày đặc để tính đạo hàm. Tuy nhiên, vẫn phải chọn các điểm để đo sai số vật lý. Các điểm này được gọi là điểm bố trí. Nếu đặt điểm bố trí quá thưa, nghiệm sẽ dao động. Nói rằng không bị ràng buộc bởi lưới không có nghĩa là ở đâu cũng chính xác.

Hàm mất mát gồm nhiều hạng cùng lúc. Chênh lệch với dữ liệu là một hạng. Sai số của các phương trình khối lượng, động lượng, năng lượng và loài hóa học là từng hạng riêng. Điều kiện biên và trạng thái ban đầu cũng được đưa vào như các hạng. Quá trình học cân nhắc các hạng này và giảm chúng cùng nhau.

Cách này tạo ra lợi ích lớn. Dù dữ liệu ít, định luật vật lý vẫn lấp vào chỗ trống. Trong phòng thí nghiệm, người ta chỉ dùng vài cảm biến áp suất trên vách hoặc đầu dò quang học. Chỉ với các phép đo thưa như vậy thì không thể biết toàn bộ dòng chảy. Nếu đặt định luật vật lý làm quy tắc, ta có thể lấp phần ở giữa.

Tính chất này phát huy tác dụng trong bài toán giải ngược. Tính toán thông thường đưa nguyên nhân vào để nhận kết quả. Bài toán giải ngược nhìn kết quả để tìm nguyên nhân. Ví dụ là chỉ nhìn nhiệt độ vách để tìm nguồn nhiệt bên trong. Mạng nơ-ron dựa trên vật lý mạnh trong những bài toán ngược như vậy. Lý do là nó xử lý phép đo và vật lý trong cùng một hàm mất mát.

Trong thử nghiệm tên lửa, năng lực này rất có giá trị. Khó đặt nhiều cảm biến bên trong buồng đốt. Nơi đó nóng và hẹp nên cảm biến không chịu nổi. Vì vậy, người ta chỉ đo vài điểm ở bên ngoài. Khôi phục

trạng thái bên trong từ vài điểm này là nhiệm vụ then chốt. Mạng nơ-ron dựa trên vật lý bắc cầu giữa phép đo và trạng thái bên trong.

Trong hàm mất mát này có một khó khăn ẩn bên trong, đó là cân trọng số. Vấn đề là xác định tầm quan trọng của từng hạng như thế nào. Nếu đặt trọng số quá lớn cho một hạng, các hạng khác sẽ bị bỏ qua. Nếu chỉ đuổi theo sai số vật lý, dữ liệu sẽ bị bỏ lỡ. Nếu chỉ đuổi theo dữ liệu, vật lý sẽ bị vi phạm. Cách giữ cân bằng này là một chủ đề lớn của nghiên cứu.

Có thể ví việc cân trọng số này với nêm nếm khi nấu ăn. Muối quá mạnh thì mặn, đường quá mạnh thì ngọt. Nếu một thứ vượt quá mức, vị sẽ hỏng. Các hạng trong hàm mất mát cũng vậy. Nếu một hạng quá mạnh, quá trình học sẽ nghiêng về một phía. Vì vậy, các phương pháp tự điều chỉnh trọng số cũng đang được nghiên cứu.

Nghiên cứu đầu tiên phổ biến rộng rãi khái niệm này là bài báo năm 2019 của Raissi và cộng sự. Sau đó, nó lan rộng sang lĩnh vực cháy. Năm 2025, một bài tổng quan về học dựa trên vật lý cho cháy cũng đã được công bố. Bài tổng quan này chia thành các loại: phản ứng hóa học, dòng phản ứng và các vấn đề cháy khác. Nó cho thấy học dựa trên vật lý đã trở thành một dòng chảy lớn.

Cũng phải nhìn thẳng vào giới hạn. Mạng nơ-ron dựa trên vật lý tìm nghiệm tốt cho một điều kiện đã định. Nhưng khi điều kiện thay đổi, nhiều khi phải học lại. Việc chứa một vùng điều kiện rộng trong một lần vẫn còn khó. Quá trình học cũng có thể dao động ở sóng xung kích mạnh hoặc mặt ngọn lửa thay đổi gấp. Nghiên cứu nhằm bù đắp điểm yếu này vẫn đang tiếp diễn.

## 8.3 Động học phản ứng cứng và học vật lý có ràng buộc bảo toàn

### 8.3.1 Độ cứng số của phản ứng hóa học và sự sụp đổ của học tiêu chuẩn

Độ cứng xuất hiện khi có biến đổi rất nhanh và có biến đổi chậm. Tình huống này phổ biến trong hóa học cháy. Các chất trung gian như gốc tự do sinh ra rồi biến mất trong chớp mắt. Trong khi đó, nhiệt độ hoặc thành phần chính thay đổi tương đối chậm.

Hãy xem chênh lệch thời gian này lớn đến mức nào. Phản ứng nhanh diễn ra theo đơn vị một phần tỷ giây. Dòng chảy chậm chuyển động theo đơn vị một phần nghìn giây. Chênh lệch giữa hai thang đo vượt quá một triệu lần. Khoảng cách lớn này trở thành gốc rễ của độ cứng làm rung chuyển phép tính.

Nếu giải biến đổi nhanh và biến đổi chậm bằng cùng một khoảng thời gian, vấn đề sẽ phát sinh. Nếu đặt khoảng thời gian theo phản ứng nhanh, phép tính sẽ kéo dài vô tận. Nếu đặt theo dòng chảy chậm, phản ứng nhanh sẽ bị bỏ lỡ và phép tính phân kỳ. Phân tích số truyền thống cũng đã vật lộn lâu dài ở điểm này. Vì vậy, phân tích số truyền thống đã xử lý nó bằng các phương pháp tích phân thời gian chuyên dụng.

Chênh lệch thời gian này giống với việc chụp ảnh. Hãy giả sử ta muốn đưa nhịp đập cánh của chim ruồi và bước đi của rùa vào cùng một tấm ảnh. Nếu chỉnh theo chim ruồi, phải bấm cửa trập vô số lần. Nếu chỉnh theo rùa, chim ruồi sẽ bị nhòe. Rất khó đưa cả hai vào cùng một khung hình. Phản ứng nhanh và dòng chảy chậm trong cháy cũng giống như vậy.

Mạng nơ-ron dựa trên vật lý tiêu chuẩn thường sụp đổ ở đây. Mạng nơ-ron này cố giảm tất cả sai số cùng một lúc. Sai số lớn do phản ứng nhanh tạo ra chi phối quá trình học. Khi đó, nó không học đúng cân bằng của các biến chậm. Hiện tượng quá trình học lệch về một phía như vậy được gọi là bệnh lý gradient.

Hãy vẽ bệnh lý này bằng một hình ảnh. Giả sử một giọng nói ồn ào chiếm trọn cuộc họp. Lời của những người nói nhỏ bị chôn vùi. Cuộc họp chỉ nghiêng về một phía. Sai số lớn của phản ứng nhanh chính là giọng nói ồn ào đó. Vì vậy, cân bằng của các biến chậm bị đẩy xuống hàng sau trong quá trình học.

Ở đây còn ẩn thêm hai nguyên nhân. Một là tính chất mạng nơ-ron học các biến đổi gặp một cách chậm chạp. Mạng nơ-ron học dòng biến đổi thoải trước, rồi đẩy các biến đổi sắc nhọn về sau. Tính chất này được gọi là thiên lệch phổ. Biến đổi dốc đứng tại khoảnh khắc đánh lửa chính là phần sắc nhọn này. Nguyên nhân còn lại là vấn đề không giữ được quan hệ nhân quả theo thời gian. Trong cháy, khoảnh khắc trước là nguyên nhân của khoảnh khắc sau. Nhưng học tiêu chuẩn cố khớp mọi thời điểm cùng một lúc. Khi đó, trật tự nguyên nhân và kết quả bị xáo trộn. Nặng hơn, nghiệm có thể sụp xuống thành một nghiệm phẳng lì, không xảy ra phản ứng nào. Vì vậy, đã xuất hiện các phương pháp chia thời gian để học lần lượt hoặc gán trọng số khác nhau cho từng loài hóa học.

Một vấn đề khác là sự sụp đổ của định luật bảo toàn. Khi cân nhiều hạng cùng lúc, cân bằng bị vỡ. Tổng nồng độ của các loài hóa học phải luôn không đổi. Số nguyên tử trước và sau phản ứng cũng phải giống nhau. Nếu quy tắc này bị phá vỡ, sẽ xuất hiện các giá trị vô lý như nồng độ âm.

Hãy xem vì sao nồng độ âm là vấn đề lớn. Trong thực tế không có chuyển lượng của một chất là số âm. Nước không thể là âm một cốc. Nếu phép tính đưa ra giá trị như vậy, bước tiếp theo sẽ sụp đổ. Tính toán tốc độ phản ứng sẽ chạy mất kiểm soát hoặc dừng lại. Vì vậy, phải ngăn nồng độ rơi xuống âm.

Có nhiều nghiên cứu ban đầu xử lý vấn đề này. Stiff-PINN năm 2021 của Ji và cộng sự là ví dụ tiêu biểu. Nghiên cứu này đưa vào xử lý chuyên dụng để xử lý hóa học cứng. Mạng nơ-ron dựa trên vật lý đa thang đo cũng xuất hiện vào thời kỳ tương tự. Những nghiên cứu này xác nhận độ cứng là một bức tường lớn đến mức nào đối với quá trình học.

Có lý do khiến bức tường này cao hơn trong tên lửa. Cháy methane và oxygen có nhiều đường phản ứng. Số chất phải xử lý lên tới hàng chục loại. Khi số chất tăng, độ cứng cũng tăng theo. Vì vậy, cháy trong tên lửa nằm ở đầu khó của bài toán độ cứng. Phương pháp áp đặt vật lý chặt chẽ càng trở nên cấp thiết.

### 8.3.2 Kết hợp ràng buộc bảo toàn và tích phân thời gian

Chìa khóa để giải bài toán độ cứng nằm ở việc áp đặt vật lý chặt hơn. Có hai cách lớn để áp đặt vật lý. Một là ràng buộc mềm, tức là phạt khi quy tắc bị vi phạm. Cách cộng hạng sai số vật lý vào hàm mất mát thuộc loại này. Cách còn lại là ràng buộc cứng, tức là nhúng quy tắc vào cấu trúc để không thể vi phạm ngay từ đầu. Độ cứng càng nặng thì hai cách này càng thường được dùng cùng nhau.

Công cụ thứ nhất là nối bộ tích phân thời gian đã được kiểm chứng với mạng nơ-ron. Phương pháp Runge-Kutta là kỹ thuật tiêu chuẩn giải phương trình vi phân bằng cách chia thời gian thành từng bước nhỏ. Nhưng phải chọn đúng loại. Phương pháp tường minh, vốn đẩy giá trị thẳng từ phía trước, yếu trước bài toán cứng. Muốn giải ổn định thì phải chia khoảng thời gian quá nhỏ. Vì vậy, với bài toán cứng, người ta dùng các phương pháp tích phân ổn định thuộc họ ẩn. Ta đưa cấu trúc tích phân đã được kiểm chứng về ổn định này vào bên trong mạng nơ-ron. Khi đó, học sâu thừa hưởng độ ổn định của phân tích số đã tích lũy trong thời gian dài.

Công cụ thứ hai là phép chiếu cưỡng bức bảo toàn. Ta không dùng nguyên xi nồng độ do mạng nơ-ron đưa ra. Trước hết, luôn chuyển giá trị thành số dương. Tiếp theo, chỉnh lại để tổng luôn không đổi. Cuối cùng, điều chỉnh giá trị để số nguyên tử được bảo toàn. Sau quá trình này, giá trị vi phạm vật lý sẽ không xuất hiện ngay từ đầu.

Có thể ví phép chiếu với người gác cửa. Dù mạng nơ-ron đưa ra giá trị nào, nó cũng được kiểm tra trước cửa. Nếu không đúng quy tắc, nó được sửa thành giá trị đúng gần nhất. Số âm được đưa về số dương. Nếu tổng bị lệch, nó được chỉnh lại. Vì vậy, giá trị đi ra ngoài luôn tuân thủ quy tắc.

Cách này khác với cách mềm ở một điểm. Cách mềm chỉ phạt khi quy tắc bị vi phạm. Hình phạt chỉ dẫn dắt quá trình học, chứ không ngăn được vi phạm. Cách cứng chặn chính hành vi vi phạm bằng cấu trúc.

Dù quá trình học còn vụng, vật lý vẫn được giữ. Vì vậy, đây là cách phù hợp với tên lửa, nơi an toàn quan trọng hơn mọi thứ.

Một ví dụ thực tế là CRK-PINN. Nghiên cứu này giải phương trình tốc độ phản ứng cháy bằng mạng nơ-ron dựa trên vật lý. Các tác giả là Zhang Shihong, Zhang Qi và Wang Bosen. Bài báo này đã qua bình duyệt và đăng trên tạp chí Combustion and Flame. Phương pháp này dùng ràng buộc mềm, đưa định luật vật lý vào các hạng của hàm mất mát. Nó đưa định luật tác dụng khối lượng và định luật Arrhenius vào mất mát. Bảo toàn enthalpy, bảo toàn nguyên tố và bảo toàn tổng nồng độ cũng được áp đặt theo cùng cách. Khác với phép chiếu cứng ở trên, đây là hướng dùng hình phạt để dẫn dắt vật lý.

Hiệu năng cũng đã được công bố. Trong hệ phản ứng giữa hydrogen và không khí, tính toán đã nhanh hơn. Nó xử lý mười loài chất và hai mươi một phản ứng. So với tính toán tiêu chuẩn, nó nhanh hơn khoảng sáu đến mười bốn lần. Khi ghép với tính toán dòng phản ứng thực tế, nó vẫn nhanh hơn khoảng hai đến năm lần. Đây là một ví dụ đạt được tốc độ tính toán trong khi vẫn tuân thủ định luật vật lý.

Hãy xem câu nói nhanh hơn vài lần có nghĩa gì. Tính toán phản ứng hóa học là phần lớn tốn thời gian trong phân tích cháy. Một phần lớn thời gian tính toán tổng thể được dùng ở đây. Nếu làm phần này nhanh hơn mười lần, tổng thể sẽ giảm mạnh. Phép tính từng mất vài ngày có thể giảm xuống còn một ngày. Nhờ đó, chu kỳ sửa thiết kế và thử lại sẽ nhanh hơn.

Ở đây cần nói rõ để không gọi lẫn ràng buộc mềm và ràng buộc cứng. CRK-PINN là cách mềm, dẫn dắt vật lý bằng mất mát. Nó khác về tính chất với phép chiếu cứng, vốn cưỡng bức sửa giá trị. Hai cách này có thể dùng cùng nhau, cũng có thể dùng riêng. Ràng buộc mềm linh hoạt nhưng không chặn hoàn toàn vi phạm vật lý. Ràng buộc cứng chặn vi phạm nhưng làm cấu trúc phức tạp hơn. Dùng bên nào và dùng bao nhiêu là phán đoán trong thiết kế.

Bài học mà cách này đem lại rất dễ hiểu và rõ ràng. Đó là không vứt bỏ tri thức phân tích số đã được kiểm chứng. Những phương pháp như Runge-Kutta đã được mài giũa suốt nhiều thập niên. Trí tuệ nhân tạo đặt lên trên tri thức đó để tăng sức mạnh. Đây là cách tri thức vật lý cũ và học máy mới không đối đầu nhau, mà giúp đỡ nhau.

Tuy vậy, giới hạn cũng rõ ràng. Những phương pháp này đã được kiểm chứng trong hệ phản ứng mà chúng xử lý. Sự cháy methane và oxygen trong rocket phức tạp hơn. Số lượng chất và phản ứng lớn hơn nhiều. Muốn chuyển sang một hệ phản ứng mới, ta phải học lại và kiểm chứng lại. Cũng có sự đánh đổi là càng áp đặt vật lý mạnh thì tính toán càng nặng.

## 8.4 Mô hình hóa dòng phản ứng dựa trên mạng Kolmogorov-Arnold

### 8.4.1 Nền tảng đặt hàm lên đường nối và hàm kích hoạt thích nghi

Trước hết hãy hình dung cấu trúc của một mạng nơ-ron thông thường. Mạng nơ-ron gồm các nút và các đường nối giữa các nút. Ở nút có hàm kích hoạt đã định sẵn. Trên đường nối chỉ gắn một con số gọi là trọng số. Học nghĩa là điều chỉnh các con số trọng số này.

Trước hết hãy giải thích cụm từ hàm kích hoạt. Mạng nơ-ron nhận đầu vào rồi đưa nó đi qua một đường cong nào đó trước khi xuất ra. Đường cong này là hàm kích hoạt. Mạng nơ-ron thông thường cố định hình dạng của đường cong này. Nó chồng nhiều đường cong đã định sẵn lên nhau để tạo ra quan hệ phức tạp. Việc học giữ nguyên hình dạng đường cong và chỉ thay đổi cường độ chồng lên nhau.

Mạng Kolmogorov-Arnold đảo ngược cách sắp đặt này. Nó đặt hàm không phải ở nút, mà trên đường nối. Chính hàm đó được thay đổi bằng học máy. Nó không dùng hình dạng định sẵn, mà nặn đường cong một cách linh hoạt. Ý tưởng này đến từ một định lý toán học đã được chứng minh từ lâu.

Định lý đó là định lý biểu diễn Kolmogorov-Arnold. Định lý này nói rằng một hàm phức tạp có thể được biểu diễn bằng tổng của các hàm một chiều đơn giản. Cấu trúc này biểu diễn hàm trên từng đường nối bằng một đường cong linh hoạt. Đường cong này có thể thay đổi hình dạng cục bộ. Vì vậy, có thể chọn riêng những nơi biến đổi gấp như ngọn lửa rồi tinh chỉnh dày hơn ở đó.

Hãy chuyển định lý này sang hình ảnh xếp khối. Dù hình dạng phức tạp đến đâu, nó cũng là tổ hợp của vài loại khối cơ bản. Ta chia cấu trúc lớn thành những mảnh nhỏ rồi nối lại. Hàm phức tạp cũng có thể được tách thành tổng của các mảnh một chiều đơn giản. Mạng Kolmogorov-Arnold lấy sự tách nhỏ này làm cấu trúc. Có thể xem toán học cũ đã trở thành nền tảng của một mạng nơ-ron mới.

Vật liệu tạo nên đường cong này là B-spline. B-spline là phương pháp nối trơn nhiều đoạn đường cong. Nó cùng họ với công cụ mà nhà thiết kế ô tô dùng khi vẽ các đường cong mềm mại. Mỗi đoạn chỉ ảnh hưởng trong khoảng của chính nó. Vì vậy, dù sửa một chỗ, những chỗ khác không bị rung theo.

Tính cục bộ này đem lại lợi ích cho việc học. Ta có thể chia dày riêng khoảng hẹp nơi ngọn lửa đứng. Những khoảng rộng còn lại có thể để thưa hơn. Đây là cách dồn tài nguyên vào nơi cần thiết. Cách này được gọi là làm mịn lưới. Nhờ vậy, ta tiết kiệm tính toán mà vẫn đạt độ chính xác cần thiết.

Có thể cảm nhận được vì sao tính chất này tốt. Ta nhìn rõ hơn đầu vào nào ảnh hưởng đến đầu ra nào. Ta có thể biểu diễn chính xác chỉ những khoảng cần thiết. Cũng dễ quan sát và mổ xẻ hàm bằng mắt. Nó hợp với các bài toán có biến đổi cục bộ mạnh như sự cháy.

Hãy đối khác biệt giữa nút và cạnh thành một ví dụ đời thường. Mạng nơ-ron thông thường giống như chồng nhiều thấu kính có hình dạng định sẵn lên nhau. Hình dạng thấu kính cố định, chỉ điều chỉnh cường độ chồng lên nhau. Mạng Kolmogorov-Arnold gọt chính hình dạng thấu kính. Vì vậy, nó tạo ra độ khúc xạ mong muốn tự do hơn. Đối lại, việc gọt thấu kính tốn công hơn.

Dĩ nhiên cấu trúc này không phải là vạn năng. Việc học có thể khó hơn. Đường cong càng được chia nhỏ thì số giá trị phải điều chỉnh càng tăng. Tùy bài toán, mạng nơ-ron thông thường có khi tốt hơn. Nên xem cấu trúc mới này là một lựa chọn.

#### 8.4.2 Năng lực biểu diễn mặt lửa và giảm nhẹ tham số

Mặt lửa là một vùng mỏng nơi phản ứng tập trung. Ở đó, nhiệt độ vọt lên mạnh trong một khoảng cách ngắn. Khí ở nhiệt độ phòng tăng lên hàng nghìn độ chỉ trong chớp mắt. Nồng độ cũng thay đổi đột ngột. Rất khó biểu diễn bằng tính toán sự biến đổi dốc trong khoảng hẹp này.

Mạng nơ-ron thông thường yếu trước những biến đổi gấp như vậy. Với hàm kích hoạt đã định sẵn, rất khó vẽ trơn một lớp sắc nét. Nếu ép cho khớp, lỗi dạng gợn sóng sẽ xuất hiện ở vùng xung quanh. Nó cũng có tính chất học chậm thông tin tần số cao. Vì những tính chất này, kết quả tính toán mặt lửa bị nhòe.

Hãy ví những gợn sóng này với bức vẽ lá cờ. Giả sử ta chỉ dùng cọ mềm để vẽ một cột thẳng đứng. Bên cạnh mép sắc của cột sẽ xuất hiện hoa văn dạng sóng. Trong tính toán cũng vậy, bên cạnh mặt lửa sắc nét xuất hiện hoa văn giả. Hoa văn này là thứ không có trong thực tế. Những hoa văn giả như vậy làm mờ dự báo bằng các tín hiệu không tồn tại trong thực tế.

Mạng Kolmogorov-Arnold có thể bộc lộ điểm mạnh ở đoạn này. Nó vẽ tốt hơn các lớp dốc bằng đường cong linh hoạt. Nó có thể chọn riêng khoảng hẹp nơi ngọn lửa đứng rồi chia dày ở đó. Nơi không cần thì để thưa, nhờ vậy giảm lãng phí. Năng lực thích nghi bằng cách chuyển tài nguyên như vậy hợp với tính toán sự cháy.

Một trường hợp thực tế là ChemKANs. Tác giả là König, Kim và Deng. Nghiên cứu này được đăng trên tạp chí Physical Chemistry Chemical Physics năm 2025. Nó đưa tốc độ phản ứng và dòng chảy của các định

luật nhiệt động lực học vào mạng Kolmogorov-Arnold. Mục tiêu là học mô hình hóa học cháy và tăng tốc tính toán. Tuy nhiên, thứ nghiên cứu này kiểm chứng không phải là mặt lửa lan trong không gian. Nó học bài toán lò phản ứng hóa học, trong đó các chất phản ứng theo thời gian. Việc khôi phục mặt lửa không gian hoặc sóng xung kích trong buồng đốt rocket vẫn là bài toán cho tương lai.

Kết quả công bố có tính thực dụng. Mô hình này hoạt động với rất ít giá trị điều chỉnh. Nó học tốt ngay cả với dữ liệu có lẫn nhiễu mà không bị quá khớp. Nó chịu được nhiễu tối đa mười lăm phần trăm. Nó cũng cho thấy kết quả nhanh hơn khoảng hai lần so với tính toán chuẩn. Cùng nhóm nghiên cứu đó cũng công bố một nghiên cứu ban đầu tiếp theo về học tốc độ phản ứng theo áp suất.

Hãy giải thích ý nghĩa của việc mạnh trước nhiễu. Đo đạc thực tế luôn có nhiễu lẫn vào. Vì cảm biến không hoàn hảo. Mô hình yếu trước nhiễu sẽ ghi nhớ cả nhiễu đó. Điều này gọi là quá khớp. Mô hình mạnh trước nhiễu chỉ chọn và học quy luật thật.

Hãy ví quá khớp với việc học thi. Học sinh chỉ học thuộc nguyên đáp án sẽ sai khi đề thay đổi. Học sinh hiểu nguyên lý thì giải được cả bài mới. Mô hình mạnh trước nhiễu giống như học sinh đã học nguyên lý. Dữ liệu đo của rocket có nhiễu nhiễu. Vì vậy, tính chất mạnh trước nhiễu này có giá trị trong đo đạc rocket.

Điểm mạnh của mô hình này phải được nói trong phạm vi đã xác nhận. Có ba điều đã được xác nhận. Đó là số giá trị điều chỉnh ít, mạnh trước nhiễu, và nhanh hơn khoảng hai lần so với tính toán chuẩn. Những con số nêu hiệu năng lớn hơn thế vẫn cần được kiểm chứng. Vì vậy, cuốn sách này chỉ nói trong phạm vi đã được xác nhận đó.

Mạng Kolmogorov-Arnold có ý nghĩa gì với rocket. Mặt lửa cũng đứng thành lớp mỏng trong buồng đốt rocket. Phải vẽ tốt lớp mỏng này thì mới dự báo đúng nhiệt và áp suất. Ít tham số thì tính toán nhẹ hơn. Mô hình nhẹ phù hợp cho tính toán thời gian thực. Vì vậy, cấu trúc này được quan tâm như một ứng viên cho chẩn đoán thời gian thực.

Hãy nhìn vào sức nặng của câu nói rằng tham số ít. Tham số là số lượng các con số mà mô hình điều chỉnh. Nếu con số này nhiều, việc học cần nhiều dữ liệu. Tính toán cũng nặng và dùng nhiều bộ nhớ. Nếu làm cùng một việc bằng ít tham số, ta có lợi ở nhiều mặt. Nó có thể chạy cả trên máy tính nhỏ mang theo rocket.

Cũng phải nhìn cả giới hạn. Những trường hợp này đã được kiểm chứng trong các hệ phản ứng tương đối đơn giản. Nếu số loài hóa học tăng lên hàng chục, số đường nối tăng bùng nổ. Khi đó bộ nhớ và thời gian học có thể tăng mạnh. Vì lý do này, các nghiên cứu tiếp tục tinh chỉnh cấu trúc nhẹ hơn. Sự cháy rối áp suất cao của rocket phức tạp hơn nhiều. Muốn dùng cấu trúc này trong thiết kế động cơ thực tế, vẫn còn nhiều việc phải làm. Cần thêm mã nguồn công khai, thí nghiệm so sánh và kiểm chứng độc lập. Hiện nay, nên xem nó là một cấu trúc ứng viên đầy triển vọng.

## 8.5 Toán tử nơ-ron Fourier biến hàm thành hàm

### 8.5.1 Nguyên lý học toán tử và tính không phụ thuộc lưới

Toán tử là quy tắc đi từ hàm này sang hàm khác. Nghe khó, nên hãy đổi sang ví dụ. Giả sử ta đưa vào phân bố lưu lượng ở cửa vào buồng đốt. Khi đó phân bố nhiệt độ trong toàn bộ buồng đốt sẽ đi ra. Quy tắc biến bức tranh cửa vào này thành bức tranh nhiệt độ chính là toán tử.

Hãy ví ánh xạ giữa hàm và hàm với nấu ăn. Giả sử ta đưa vào toàn bộ danh sách nguyên liệu rồi nhận ra ảnh món ăn hoàn chỉnh. Nếu thay đổi chút ít nguyên liệu, món ăn cũng thay đổi theo. Học toán tử học phép biến đổi nguyên khối này. Nó học quan hệ tổng thể, không phải giá trị của từng nguyên liệu riêng lẻ. Vì vậy, dù điều kiện thay đổi, nó vẫn trả lời nhanh.

Mạng nơ-ron thông thường biến vài con số thành vài con số. Học toán tử biến một bức tranh thành một bức tranh. Vì vậy, nó xử lý toàn bộ trường dòng chảy cùng một lúc. DeepONet là một nghiên cứu đã mở rộng ý tưởng này từ sớm. Toán tử nơ-ron Fourier cũng là một phương pháp thuộc học toán tử này.

Hãy xem vì sao khác biệt này lớn. Nếu phải tính lại từ đầu cho từng điều kiện, tốc độ sẽ chậm. Khi đã học toán tử một lần, mô hình trả lời nhanh cho nhiều điều kiện. Dù điều kiện cửa vào thay đổi, không cần học lại. Trong thiết kế, người ta thử vô số điều kiện khác nhau. Vì vậy, học toán tử được chú ý như một công cụ thiết kế.

Cốt lõi của toán tử nơ-ron Fourier nằm ở biến đổi Fourier. Biến đổi Fourier là kỹ thuật đổi một hình dạng phức tạp thành tổng của nhiều sóng. Sóng tần số thấp chứa cấu trúc lớn của dòng chảy. Sóng tần số cao chứa chi tiết như hoa văn nhỏ. Mạng nơ-ron này chủ yếu học thành phần tần số thấp để vẽ nhanh dòng chảy lớn.

Cách này có lợi thế riêng. Mạng nơ-ron xử lý dòng chảy như tổng của các sóng. Nó tóm bắt nhanh dòng chảy lớn bằng vài tần số thấp. Nó thuận lợi cho việc nắm bắt quan hệ trên vùng rộng trong một lần nhìn. Tuy nhiên, phương pháp này không phải là kỹ thuật số trực tiếp giải vi phân của phương trình. Nếu chỉ giữ tần số thấp và cắt bỏ tần số cao, tính toán sẽ nhẹ hơn. Nhưng chính việc cắt đó có thể làm nhòe những phần sắc như sóng xung kích hoặc mặt lửa. Vì vậy, phải thận trọng ở nơi chi tiết quan trọng.

Hãy ví sự nhòe này với ảnh chụp. Ảnh lệch tiêu điểm có biên mờ. Đường sắc nhìn như bị loang. Tần số cao càng bị cắt nhiều thì hiện tượng loang này càng xuất hiện. Biên sắc của ngọn lửa có thể bị bệt đi. Vì vậy, những nơi như vậy cần xử lý riêng để giữ lại tần số cao.

Tính chất nổi bật ở phương pháp này là không phụ thuộc lưới. Mô hình này được định nghĩa không phụ thuộc vào độ dày của lưới tính toán. Sau khi học bằng dữ liệu độ phân giải thấp, nó có thể được áp dụng cho lưới dày hơn. Điều này gọi là siêu phân giải zero-shot. Nó giống với việc khôi phục ảnh độ phân giải thấp thành ảnh độ phân giải cao. Tuy nhiên, đây là tính chất lý thuyết và kết quả đã thấy trong một số thí nghiệm nhất định. Độ chính xác ở độ phân giải mới không tự động được bảo đảm.

Mạng nơ-ron thông thường không làm được việc này. Kích thước đầu vào và đầu ra đã được định sẵn ngay từ đầu. Nếu thay đổi lưới, phải tạo mô hình mới. Toán tử nơ-ron Fourier thì khác. Nó xử lý dòng chảy như một hàm, không phải như một lưới. Vì vậy, dù độ phân giải thay đổi, vẫn dùng cùng một mô hình.

Hãy ví biến đổi Fourier với radio. Radio nhận nhiều chương trình phát sóng bằng cách chia theo tần số. Tần số thấp và tần số cao chứa những thông tin khác nhau. Toán tử nơ-ron Fourier cũng chia dòng chảy theo tần số. Dòng chảy lớn nằm ở tần số thấp, hoa văn nhỏ nằm ở tần số cao. Mạng nơ-ron này chọn phía thấp để tóm bắt nhanh dòng chảy lớn.

Tính không phụ thuộc lưới đem lại lợi ích lớn cho tính toán rocket. Ở giai đoạn đầu phát triển, người ta xem nhanh nhiều phương án bằng lưới thô. Sau khi chọn phương án tốt, người ta kiểm tra chính xác bằng lưới dày. Lúc này không cần học lại mô hình. Vì cùng một mô hình được dùng nguyên như vậy ở hai độ phân giải, thời gian và chi phí đều giảm.

Tuy nhiên, điều này không tự động xảy ra trong mọi bài toán. Những nơi giá trị bị ngắt đột ngột như sóng xung kích rất khó. Mặt lửa mỏng và dòng rối gần vách cũng phải được xử lý thận trọng. Ở những nơi này, thông tin tần số cao rất quan trọng. Phải có xử lý riêng và kiểm chứng đi kèm thì mới đáng tin.

### 8.5.2 Toán tử nơ-ron cho dòng chảy

Sự cháy là một bài toán khó, nơi nhiệt, dòng chảy và hóa học đan vào nhau. Nghiên cứu nhằm điều chỉnh toán tử nơ-ron cho bài toán này vẫn tiếp tục. Nhiều nhóm thử các cách khác nhau. Ở đây, ta xem các trường hợp mới đã được kiểm chứng, lấy khái niệm làm trung tâm.

Một hướng là học hóa học cứng bằng toán tử. Toán tử nơ-ron Fourier mở rộng (EFNO) là một ví dụ. Tác giả là Wong, Li, Zhang, Chen và Zhou. Nghiên cứu này cố gắng giải hóa học cứng ngay cả trong các điều kiện chưa học. Nó được thử nghiệm với tự bốc cháy hydrogen không chiều và ngọn lửa tia hydrogen-ammonia ba chiều. Nó cho thấy tính chất khái quát hóa tốt ngay cả ở trạng thái chưa được học. Nghiên cứu toán tử mở rộng này được đăng trên tạp chí Combustion and Flame.

Hãy phân tích cụm từ tự bốc cháy. Khi đặt nhiên liệu và oxy trong trạng thái nóng, lửa tự bắt lên. Đó là hiện tượng vật chất tự cháy mà không cần diêm. Hóa học tại khoảnh khắc này diễn ra rất nhanh và có độ cứng lớn. Toán tử mở rộng đã cố học khoảnh khắc khó này. Đây là bài toán thường được dùng làm phép thử cho hóa học cứng.

Hãy xem vì sao toán tử mở rộng lại quan trọng. Mạng nơ-ron thông thường chỉ khớp tốt trong các điều kiện đã học. Chỉ cần nhiệt độ hoặc áp suất lệch đi một chút, kết quả đã sai. Toán tử mở rộng đã cố hạ thấp bức tường này. Trong thử nghiệm, nó đã khái quát hóa ở mức nhất định sang cả những điều kiện chưa từng học. Có thể nói nó đã cho thấy dư địa sử dụng trong các điều kiện tên lửa mới.

Một hướng khác là học trộn vụn nhiều loạn phản ứng ba chiều. Một nghiên cứu đã xử lý việc này bằng toán tử nơ-ron Fourier có thêm cấu trúc U-Net. U-Net là cấu trúc mạng nơ-ron xử lý đồng thời bức tranh lớn và chi tiết. Nó tái hiện tốt đặc tính tần số của vận tốc, nhiệt độ và mật độ trong nhiều loạn phản ứng nén được. Nó dự đoán dòng chảy nhanh hơn nhiều so với tính toán truyền thống. Kết quả như vậy cho thấy khả năng của mô hình thay thế. Nhiều nghiên cứu thuộc nhánh này vẫn đang ở giai đoạn đầu, trước bình duyệt.

Hãy làm rõ câu nói khớp tốt đặc tính tần số. Trong dòng chảy có cả xoáy lớn và xoáy nhỏ. Mỗi cỡ xoáy chứa một phần năng lượng. Tiêu chí quan trọng là có vẽ đúng phân bố năng lượng này hay không. Nếu chỉ khớp dòng chảy lớn mà bỏ lỡ xoáy nhỏ thì mới được một nửa. Một toán tử tốt tái hiện rộng phân bố này của cả dòng chảy lớn và xoáy nhỏ.

Hướng thứ ba là xử lý truyền nhiệt bức xạ bằng toán tử. Ngọn lửa nóng cũng mất nhiệt dưới dạng ánh sáng. Hiện tượng này gọi là truyền nhiệt bức xạ. Zhao và các cộng sự đã đưa ra một toán tử nơ-ron giải truyền nhiệt bức xạ trong hỏa hoạn. Đây là nghiên cứu giai đoạn đầu nhằm vào hỏa hoạn, không phải tên lửa. Vì vậy không mở rộng nó thành kiểm chứng tên lửa, mà chỉ tham khảo như một trường hợp lân cận.

Hãy cảm nhận truyền nhiệt bức xạ qua một đồng lửa trại. Khi đứng trước đồng lửa, mặt ta thấy nóng. Trước khi không khí nóng lên, ánh sáng đã mang nhiệt đến trước. Ngọn lửa trong buồng đốt tên lửa cũng tỏa nhiệt bằng ánh sáng như vậy. Bức xạ này làm nóng thành vách và làm nguội ngọn lửa. Vì thế, dự đoán chính xác phải bao gồm cả nhiệt bức xạ này.

Đưa bức xạ vào tính toán là việc khó. Ánh sáng lan ra mọi hướng và trộn lẫn với nhau. Cường độ ở mỗi hướng lại khác. Nỗ lực học dòng phức tạp này bằng toán tử đã xuất hiện trong nghiên cứu hỏa hoạn. Muốn chuyển sang tên lửa thì phải chỉnh lại điều kiện và kiểm chứng. Dù vậy, đây vẫn là một trường hợp ban đầu có ý nghĩa, cho thấy hướng đi sắp tới.

Các trường hợp này cho thấy một lợi ích chung. Khi học từ kết quả tính toán độ chính xác cao, nó trở thành mô hình thay thế nhanh. Có thể khôi phục tính toán độ phân giải thấp thành dạng độ phân giải cao. Nó trở thành công cụ hỗ trợ ước lượng toàn bộ trường trong tình huống có ít cảm biến. Có thể dùng nó để quét không gian thiết kế trong thời gian ngắn.

Cũng có điểm giao nhau giữa học toán tử và học dựa trên vật lý. Toán tử nơ-ron nhanh, nhưng không phải lúc nào cũng tuân thủ vật lý. Có thể đặt ràng buộc bảo toàn vào hàm mất mát. Đây là nỗ lực nhằm có cả tốc độ của toán tử và độ tin cậy của vật lý. Sự kết hợp này sẽ xuất hiện lại trong trường hợp detonation quay ở phần sau. Vì vậy hai công cụ này không cạnh tranh với nhau mà ghép thành một cặp.

Ở đây cần nói rõ cách đọc các trường hợp này. Cuốn sách này không đặt bảng so sánh hiệu năng của từng mô hình cụ thể. Các con số trong nghiên cứu giai đoạn đầu thay đổi lớn theo điều kiện. Thay vào đó, sách giải thích vai trò của toán tử nơ-ron bằng các trường hợp đã được kiểm chứng. Trọng tâm nằm ở cách dùng phương pháp, không phải ở con số.

Cũng có trường hợp tính toán quy mô lớn. Nhóm nghiên cứu DeepFlame đã đưa mạng nơ-ron vào để giải bài toán cháy siêu tới hạn ở quy mô lớn. Các tác giả là Guo, Mao, Liu, Tan, Jia và Chen. Theo tài liệu công khai, họ đã tính toán cháy nhiễu loạn của methane và oxy ở quy mô hàng chục tỷ ô. Đây là phạm vi vượt xa năng lực trước đây. Đó là nghiên cứu ban đầu hướng tới cháy tên lửa với hơn một trăm injector.

Muốn dùng trong thiết kế thực tế thì có nhiều mục cần kiểm tra. Phải xem khối lượng và năng lượng có được bảo toàn hay không. Phải xác nhận vị trí sóng xung kích có đúng hay không. Phải xem nồng độ các chất hóa học có rơi xuống giá trị âm hay không. Chỉ khi vượt qua các kiểm chứng như vậy, nó mới trở thành công cụ. Mô hình thay thế là cặp đôi hỗ trợ tính toán chính xác, không phải vật thay thế.

## 8.6 Khôi phục mặt ranh giới cháy sóng xung kích trong động cơ detonation quay

### 8.6.1 Bài toán phân tích động cơ detonation quay

Động cơ detonation quay (RDRE) dùng buồng đốt hình vòng. Trong đó, sóng detonation quay rất nhanh theo đường tròn. Detonation là dạng cháy nổ trong đó sóng xung kích và quá trình cháy gắn với nhau rồi cùng chuyển động. Sóng này chạy với tốc độ hàng nghìn mét mỗi giây, nhanh hơn âm thanh nhiều lần.

Cách này có lợi thế là chính quá trình cháy làm tăng áp suất. Động cơ thông thường tăng áp suất bằng máy nén hoặc bơm rồi mới đốt. Detonation đốt nhiên liệu khi sóng xung kích nén khí tức thời. Vì vậy áp suất tự tăng trong lúc cháy. Hiện tượng này gọi là cháy tăng áp. Tuy nhiên, vẫn cần áp suất phun để đẩy nhiên liệu và oxy vào. Dù áp suất tăng trong buồng đốt, vẫn có tổn thất ở phun, trộn và xả. Vì vậy lợi ích tính trên toàn bộ động cơ thay đổi theo điều kiện. Nếu dùng tốt, nó có khả năng tăng hiệu suất, nên nhiều tổ chức trên thế giới đã tham gia phát triển.

Hãy xét sự khác nhau giữa detonation và cháy thông thường. Cháy thông thường lan ra như lửa lan chậm. Detonation lan ra như một vụ nổ với sóng xung kích đi phía trước. Sóng xung kích nén và làm nóng không khí trong nháy mắt. Ngay phía sau nó, nhiên liệu cháy tức thì. Cấu trúc trong đó sóng xung kích và ngọn lửa gắn với nhau tạo ra áp suất lớn.

Vấn đề là bên trong cực kỳ phức tạp. Sóng xung kích dẫn đầu và vùng phản ứng tỏa nhiệt gắn chặt với nhau. Sóng xung kích nghiêng, mặt tiếp xúc và lớp hỗn hợp chưa cháy đan vào nhau. Các cấu trúc này tương tác ở thang thời gian một phần triệu giây. Vì vậy đây thuộc nhóm bài toán khó bậc nhất trong cơ học chất lưu.

Hãy ví tương tác này với tai nạn trên đường. Nhiều xe cùng mắc kẹt trong một giao lộ hẹp. Một thay đổi ở một điểm lan ra toàn bộ trong nháy mắt. Trong detonation, sóng xung kích và ngọn lửa cũng đan vào nhau như vậy. Một dao động nhỏ làm thay đổi toàn bộ sóng. Vì thế, dự đoán và điều khiển bên trong động cơ này là việc khó.

Đo đạc còn bức bối hơn. Những gì có thể thấy trong thử nghiệm mặt đất rất hạn chế. Gần như chỉ có tín hiệu áp suất trên thành ngoài buồng đốt. Người ta cũng chụp ánh sáng của ngọn lửa bằng camera tốc độ siêu cao. Ánh sáng này được gọi là tự phát quang. Đó là ánh sáng mà một phân tử cụ thể tự phát ra khi phản ứng. Nhưng bằng các phép đo này, ta không thể biết trực tiếp trạng thái ba chiều bên trong.

Cũng cần hiểu vì sao người ta phát triển động cơ detonation quay. Động cơ tên lửa thông thường đốt nhiên liệu ở áp suất gần như không đổi. Cháy detonation đốt nhiên liệu trong khi tự làm tăng áp suất. Có

khả năng thu được lực lớn hơn từ cùng một lượng nhiên liệu. Cũng có dư địa làm động cơ ngắn hơn và nhẹ hơn. Vì vậy cách này được xem là ứng viên cho công nghệ đẩy thế hệ tiếp theo.

Hãy ví khó khăn của động cơ này với camera. Giả sử ta chụp ảnh một con quay đang quay nhanh. Nếu màn trập chậm, con quay sẽ nhòe thành một mảng. Sóng detonation nhanh hơn con quay rất nhiều. Camera nhanh đến đâu cũng chỉ ghi được vẻ ngoài. Muốn nhìn vào bên trong, phải lắp các khoảng trống bằng tính toán.

Hãy cảm nhận tốc độ của sóng detonation bằng con số. Sóng này chạy với tốc độ hàng nghìn mét mỗi giây. Đó là tốc độ nhanh hơn viên đạn. Nó mất chưa đến một mili giây để chạy hết một vòng trong buồng đốt hình vòng. Trong khoảng thời gian ngắn đó, áp suất và nhiệt độ tăng giảm mạnh. Sự thay đổi ấy nhanh đến mức mắt người không thể theo kịp.

Vì vậy mô hình khôi phục là thứ rất cần. Cần một công cụ tái tạo toàn bộ trường dòng từ các phép đo thưa. Phải ước lượng phân bố áp suất, nhiệt độ, vận tốc và chất hóa học trong thời gian ngắn. Đây là nơi học dựa trên vật lý gặp học toán tử. Các phương pháp đã học trong chương này hợp sức tại đây.

### 8.6.2 Khôi phục và bản sao số kết hợp dữ liệu với vật lý

Công cụ đầu tiên của khôi phục là kỹ thuật siêu phân giải. Nó tái tạo các phép đo thưa hoặc thô thành một trường chi tiết. Nghiên cứu năm 2026 của Yao và Zhang đã thử làm việc này bằng mạng nơ-ron đối nghịch nhất quán chu trình. Mạng nơ-ron đối nghịch là cấu trúc trong đó bên tạo ra và bên phân biệt cạnh tranh với nhau để học. Nó giống cảnh kẻ làm giả và người giám định huấn luyện lẫn nhau. Phương pháp này học từ dữ liệu tính toán động cơ dùng hydro và không khí. Nó đã khôi phục trường dòng thô thành chi tiết hơn tám lần. Kết quả này là nghiên cứu đã đăng trên tạp chí học thuật.

Hãy nhìn bằng hình ảnh để thấy vì sao cần siêu phân giải. Khi phóng to một bức ảnh cũ, ảnh trở nên thô như bậc thang. Công nghệ siêu phân giải làm bức ảnh thô đó trở lại mượt hơn. Trong tính toán cháy cũng có kết quả lưới thô. Nếu khôi phục trường thô này thành trường chi tiết, thông tin tăng lên. Đây là cách tiết kiệm chi phí tính toán bằng tính toán thô, rồi lấy lại chi tiết.

Công cụ tiếp theo là hợp nhất đa độ trung thực. Nó nối tính toán độ chính xác thấp, rẻ tiền với tính toán độ chính xác cao, đắt tiền. Ý tưởng là có được ước lượng chính xác dù chỉ có ít dữ liệu độ chính xác cao. Một nghiên cứu tên Cheap2Rich đã áp dụng cách này cho động cơ detonation quay. Các tác giả là Bao, Jayak, Powers, Raman và Kurtz. Nghiên cứu đa độ trung thực này vẫn là nghiên cứu giai đoạn đầu, trước bình duyệt.

Hãy ví lợi thế của đa độ trung thực với việc lo liệu trong gia đình. Nếu chỉ dùng vật liệu đắt tiền, ngân sách sẽ căng. Ta dùng vật liệu rẻ để dựng nền, rồi chỉ dùng vật liệu tốt ở nơi cần thiết. Tính toán cũng vậy. Dùng tính toán độ chính xác thấp để nắm khung lớn. Chỉ tinh chỉnh các điểm quan trọng bằng tính toán độ chính xác cao. Đó là cách tiến gần tới độ chính xác cao với chi phí thấp.

Trong sự hợp nhất này còn có khái niệm đồng hóa dữ liệu. Đồng hóa dữ liệu là phương pháp liên tục khớp tính toán với đo đạc thực tế. Nó giống dự báo thời tiết nhận quan trắc rồi sửa dự báo. Trong động cơ detonation quay cũng có tín hiệu cảm biến đi vào. Tín hiệu này liên tục hiệu chỉnh tính toán. Khi đó dự đoán của mô hình gần hơn với trạng thái động cơ thực.

Một công cụ khác là mô hình sinh phản ánh quan sát. Mô hình sinh là mạng nơ-ron tự tạo ra những cảnh có vẻ hợp lý. Nếu đưa quan sát thưa làm điều kiện, nó vẽ ra toàn bộ trường. Ngay cả khi vị trí quan sát thay đổi, vẫn có dư địa sử dụng mà không phải học lại. Wang và các cộng sự đã đưa ra nỗ lực như vậy trong khôi phục buồng đốt detonation quay. Vì thế nó hữu ích trong tình huống có ít cảm biến và thiếu quan sát.

Công cụ cuối cùng là sự kết hợp giữa toán tử nơ-ron và mạng nơ-ron vật lý. Toán tử nơ-ron ước lượng nhanh toàn bộ trường. Mạng nơ-ron vật lý giữ cho nó tuân thủ các định luật bảo toàn. Khi hợp hai phương pháp lại, ta có một bộ ước lượng vừa nhanh vừa khớp vật lý. Khi tín hiệu cảm biến đi vào, nó tái tạo bên trong trong thời gian ngắn.

Kết quả khôi phục này dẫn tới bản sao số. Bản sao số là công nghệ trong đó mô hình máy tính bám theo trạng thái của thiết bị thực. Từ trường dòng đã khôi phục, có thể tính tải nhiệt trên vách và hiệu suất cháy. Có thể nhận biết trước nguy cơ injector bị bào mòn do nhiệt. Nó cũng nối với điều khiển để xử lý bất ổn cháy trong thời gian ngắn.

Hãy ví bất ổn cháy với chiếc xích đu. Khi đẩy xích đu đúng nhịp, nó dao động ngày càng lớn. Dao động áp suất trong buồng đốt cũng lớn lên nếu bắt nhịp với ngọn lửa. Khi dao động này lớn, động cơ có thể vỡ. Vì vậy phải nhận ra dao động sớm và phá nhịp đó. Khôi phục và điều khiển thời gian thực đảm nhận việc nhận ra và xử lý dao động này.

Tuy nhiên, ở đây cũng có nhiều điểm cần thận trọng. Phần lớn các nghiên cứu khôi phục này vẫn ở giai đoạn đầu. Những con số nhấn mạnh tốc độ tăng mạnh hoặc độ chính xác cao đều phụ thuộc điều kiện. Vì vậy cuốn sách này không đưa các con số đó vào phần nội dung chính. Phần này chỉ giải thích bằng khái niệm thiết kế và các trường hợp đã được kiểm chứng.

Hãy xem vì sao bản sao số hữu dụng. Không thể nhìn trực tiếp vào bên trong động cơ thực trong lúc thử nghiệm. Bản sao số dùng tín hiệu cảm biến để phán đoán trạng thái bên trong. Nó nhận ra dấu hiệu nguy hiểm nhanh hơn con người. Nó có thể gửi tín hiệu trước khi vấn đề lớn lên. Trong tên lửa tái sử dụng, cảnh báo sớm này có giá trị lớn.

Tự cập nhật cũng có rủi ro ẩn. Nếu mô hình tự sửa mình bằng chính dự đoán của nó, sai số có thể tích lũy. Nó có thể tự đẩy mình theo hướng sai. Một nghiên cứu năm 2026 đã chỉ ra vấn đề này. Nghiên cứu đó đề xuất bản sao số tự cập nhật được ổn định bằng vật lý. Nghiên cứu này không dùng tên lửa, mà cho thấy khái niệm qua một bài toán nhiệt tổng quát. Đây là cách các quy tắc vật lý nắm dây cương của quá trình cập nhật.

Hãy xem vì sao cần bộ giới hạn. Trí tuệ nhân tạo đôi khi đưa ra giá trị sai lệch. Nếu giá trị đó đi thẳng vào điều khiển, nó sẽ nguy hiểm. Bộ giới hạn ngăn giá trị vượt ra khỏi phạm vi an toàn. Đó là hàng rào giữ các giới hạn vật lý do con người đặt ra. Điều khiển tự hành nhất định phải có hàng rào như vậy.

Muốn dùng trong điều khiển động cơ thực, cần chuẩn bị nhiều hơn. Phải tích lũy đủ dữ liệu thử nghiệm tốc độ cao. Phải có kiểm chứng độc lập đi kèm. Cũng phải có bộ giới hạn bảo vệ an toàn. Dòng phát triển này cho thấy công nghệ vẫn đang trong quá trình trưởng thành. Nhưng hướng đi đã rõ. Khôi phục dùng chung dữ liệu và vật lý đang dần định hình.

## Tổng kết

Chương này đã bàn về nhiều con đường đưa trí tuệ nhân tạo vào tính toán cháy. Điểm xuất phát là một vấn đề. Cháy trong tên lửa đòi hỏi lượng tính toán lớn vì hóa học diễn ra nhanh và dòng chảy diễn ra chậm đan xen với nhau. Tính toán truyền thống chính xác nhưng chậm. Trí tuệ nhân tạo nhanh, nhưng nếu vi phạm vật lý thì nguy hiểm.

Lời giải là học máy dựa trên vật lý. Mạng nơ-ron không chỉ bắt chước dữ liệu. Nó đồng thời tuân thủ các định luật như bảo toàn khối lượng và năng lượng. Mạng nơ-ron dựa trên vật lý là phương pháp tiêu biểu. Dù dữ liệu ít, vật lý vẫn lấp vào phần còn thiếu. Sức mạnh này bộc lộ rõ trong các bài toán giải ngược.

Ở đoạn này, cần sắp xếp lại quan hệ giữa trí tuệ nhân tạo và vật lý. Trí tuệ nhân tạo mạnh trong việc tìm quy luật từ dữ liệu. Vật lý chứa các định luật đã được kiểm chứng trong thời gian dài. Chỉ một trong hai là không đủ. Nếu chỉ dùng dữ liệu, mô hình sẽ vi phạm vật lý. Nếu chỉ dùng vật lý, lượng tính toán sẽ lớn. Vì vậy, dùng hai thứ này cùng nhau là trọng tâm của toàn bộ chương này.

Trong các bài toán cứng, phải ràng buộc vật lý mạnh hơn. Chỉ dùng phạt thường không đủ. Ta đưa bộ tích phân thời gian vào mạng nơ-ron để đạt ổn định. Ta dùng phép chiếu cưỡng bức bảo toàn để ngăn các giá trị vô lý. CRK-PINN là một ví dụ về ràng buộc mềm, trong đó vật lý được đưa vào dưới dạng hàm mất mát.

Cũng cần nhắc lại các phương trình của dòng phản ứng đa thành phần. Bốn cân bằng về khối lượng, động lượng, loài hóa học và năng lượng là bộ khung. Các cân bằng này là quy tắc rằng không có gì tự nhiên sinh ra hay biến mất. Học dựa trên thông tin vật lý đưa quy tắc này vào hàm mất mát. Đạo hàm tự động đo chính xác phần dư của phương trình. Nhờ đó, nó khôi phục được toàn bộ trường ngay cả từ dữ liệu thưa.

Mạng Kolmogorov-Arnold đưa ra một cấu trúc mới. Nó đặt các hàm linh hoạt trên các đường nối. Nhờ vậy, nó có thể mô tả tốt hơn mặt ngọn lửa dốc. ChemKANs đã cho thấy khả năng này trong hóa học cháy. Nó tạo ra kết quả bền với nhiễu dù chỉ dùng ít giá trị.

Toán tử nơ-ron Fourier biến hàm thành hàm. Nó có thể học ở độ phân giải thấp rồi dùng ở độ phân giải cao. Nó trở thành mô hình thay thế nhanh và công cụ phục hồi. Nghiên cứu kết hợp toán tử mở rộng với U-Net đã đạt kết quả trong dòng phản ứng. Nghiên cứu về toán tử xử lý truyền nhiệt bức xạ cũng xuất hiện ở lĩnh vực gần kề. Tuy vậy, sóng xung kích và mặt ngọn lửa mỏng vẫn cần được xử lý thận trọng.

Động cơ detonation quay là sân khấu nơi các phương pháp này hội tụ. Mục tiêu là phục hồi bên trong từ các phép đo thưa. Người ta dùng siêu phân giải, đa độ trung thực, và kết hợp toán tử với mạng nơ-ron vật lý. Kết quả đó dẫn tới bản sao số và điều khiển thời gian thực. Hiện nay lĩnh vực này vẫn ở giai đoạn đầu và cần kiểm chứng thêm.

Có thể gom các phương pháp của chương này vào một câu. Tất cả đều là con đường dùng vật lý và dữ liệu cùng nhau. Vật lý đem lại độ tin cậy, dữ liệu đem lại tốc độ. Chỉ một phía thì khó xử lý cháy trong tên lửa. Câu trả lời nằm ở nơi hai sức mạnh gặp nhau. Cuốn sách này đã chỉ ra nơi đó bằng các trường hợp đã được kiểm chứng.

Cũng cần sắp xếp lại tính chất của các ví dụ được dùng trong chương này. Mạng nơ-ron dựa trên vật lý của Raissi và toán tử nơ-ron Fourier của Li là các nghiên cứu nền tảng được trích dẫn rộng rãi.

CRK-PINN và ChemKANs là các bài báo học thuật đã qua bình duyệt. DeepFlame, toán tử mở rộng và phục hồi đa độ trung thực là các nghiên cứu mới. Nhiều nghiên cứu trong số này vẫn là preprint ở giai đoạn đầu. Nghiên cứu giai đoạn đầu có nghĩa là còn có thể thay đổi trong tương lai.

Thái độ xuyên suốt chương này cũng đáng nhắc lại. Cuốn sách này chọn ví dụ dựa trên căn cứ. Những mô hình chỉ có tên nghe có vẻ thuyết phục được hạ xuống mức khái niệm. Các con số hiệu năng chưa được xác nhận không được đưa vào phần chính. Nghiên cứu giai đoạn đầu được nói rõ là nghiên cứu giai đoạn đầu. Mục đích là giúp người đọc phân biệt sự thật với kỳ vọng.

Các nhiệm vụ phía trước cũng rõ ràng. Cần những mô hình khớp tốt cùng lúc trong một dải điều kiện rộng. Phải xử lý ổn định sóng xung kích và mặt ngọn lửa mỏng. Cần tích lũy thêm các trường hợp được kiểm chứng bằng dữ liệu động cơ thực. Các bộ giới hạn bảo vệ an toàn cũng phải phát triển cùng lúc. Khi các nhiệm vụ này được giải quyết, công cụ sẽ tiến gần hơn tới hiện trường.

Bài học của chương này rất rõ. Trí tuệ nhân tạo làm thay đổi lớn tốc độ tính toán cháy. Tuy nhiên, nó chỉ đáng tin khi đi cùng vật lý và kiểm chứng. Căn cứ đã được xác nhận quan trọng hơn tên mô hình hào

nhoáng. Trên sự cân bằng đó, công nghệ đẩy thế hệ tiếp theo lớn lên từng bước. Khi trí tuệ nhân tạo và vật lý nắm tay nhau, bước đi ấy trở nên vững chắc.

### Tài liệu căn cứ liên quan

- 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 (Bản tiền ấn phẩm, chưa qua bình duyệt)*. 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 (Bản tiền ấn phẩm, chưa qua bình duyệt)*. 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 (Bản tiền ấn phẩm, chưa qua bình duyệt)*. 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 (Bản tiền ấn phẩm, chưa qua bình duyệt)*. 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 (Bản tiền ấn phẩm, chưa qua bình duyệt). arXiv:2604.25617. <https://arxiv.org/abs/2604.25617>
- Mousavi, M., Caldwell, C., Baltes, J., Aljaseem, M., & Lee, B. J. (2025). Physics-Informed Neural Networks in Clean Combustion: A Pathway to Sustainable Aerospace Propulsion. arXiv preprint (Bản tiền ấn phẩm, chưa qua bình duyệt). 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.

## Chương 9. Tối ưu hóa thiết kế ngược topo của kênh làm mát tái sinh, lỗ làm mát màng và inducer

### Mở đầu

Nếu cầm một chiếc nồi nóng bằng tay vào mùa hè, ta sẽ bị bỏng. Vì vậy, ta làm tay cầm nồi bằng gỗ hoặc nhựa. Mục đích là ngăn nhiệt truyền đến tay. Động cơ tên lửa cũng mang cùng một bài toán. Bên trong động cơ, nhiên liệu cháy và tạo ra khí rất nóng. Khí này nóng đến mức có thể làm chảy cả kim loại. Nhưng bức vách chứa dòng khí nóng ấy lại làm bằng kim loại. Bảo vệ bức vách đó như thế nào là một bài toán lớn trong thiết kế động cơ tên lửa.

Tên lửa giải bài toán này theo một cách khéo léo. Nhiên liệu đi vào động cơ không được đốt ngay. Trước hết, người ta cho nhiên liệu chảy qua các đường dẫn nhỏ tạo trong bức vách nóng. Khi đi qua vách, nhiên liệu lấy đi nhiệt của vách. Vách nguội xuống, còn nhiên liệu nóng lên. Sau đó nhiên liệu đã được làm nóng đi vào buồng đốt và cháy. Đây là cách vừa bảo vệ vách, vừa gia nhiệt trước cho nhiên liệu. Người ta gọi cách này là làm mát tái sinh(regenerative cooling). Vì nhiên liệu đã được làm nóng không bị bỏ đi mà được dùng lại, nên có chữ tái sinh.

Làm mát tái sinh là một phương pháp cũ. Tên lửa thời kỳ đầu đã dùng phương pháp này. Nhưng việc thiết kế kênh làm mát một cách tự do lại khó. Lý do là phương pháp gia công bằng cách cắt gọt kim loại đã trói buộc hình dạng của kênh. Gần đây, hai thứ đã tháo nút trói này. Một là in 3D kim loại. Thứ còn lại là trí tuệ nhân tạo, công cụ tìm nhanh thiết kế tốt. Khi hai thứ này gặp nhau, thiết kế làm mát đã bước sang một cục diện mới.

Chương này bàn về cấu trúc quản lý nhiệt của động cơ tên lửa. Chúng ta sẽ xem xét kênh làm mát trong vách và các lỗ làm mát màng phủ một lớp màng làm mát mỏng lên vách. Chương này cũng bàn đến inducer, tức cánh bơm phía trước đẩy nhiên liệu lên. Cả ba cấu trúc đều là bài toán cân bằng giữa nhiệt và dòng chảy. Nếu muốn làm mát vách mạnh hơn, lực cản cản trở dòng chảy sẽ tăng. Nếu muốn giảm lực cản, khả năng làm mát sẽ yếu đi. Khó có thể tìm điểm cân bằng này chỉ bằng bàn tay con người.

Trí tuệ nhân tạo bước vào đây. Trí tuệ nhân tạo được bàn trong chương này không phải là trí tuệ nhân tạo vẽ tranh hay viết văn. Nó là công cụ giúp nhanh chóng tìm đáp án tốt trong một không gian thiết kế phức tạp. Có loại trí tuệ nhân tạo thay thế phép tính dòng chảy chậm. Có loại trí tuệ nhân tạo tự chọn thí nghiệm tiếp theo cần làm. Có loại trí tuệ nhân tạo rút gọn hình dạng kênh làm mát phức tạp thành một nhóm số nhỏ. Chương này sẽ giải thích từng công cụ ấy một cách dễ hiểu.

Hãy nói rõ vì sao câu chuyện của chương này quan trọng. Trong động cơ tên lửa, nơi hay gây sự cố là quản lý nhiệt và bơm. Nếu vách chảy hoặc bơm rung, động cơ sẽ dừng. Vì vậy, khi chế tạo một động cơ mới, phần này tốn nhiều thời gian và tiền bạc. Chỉ cần thiết kế nhanh hơn một chút, an toàn hơn một chút, cũng đã giúp rất nhiều. Trí tuệ nhân tạo góp sức chính tại điểm này. Nó quét qua không gian thiết kế rộng nhanh hơn con người và chọn ra các ứng viên tốt.

Có một điều cần nói trước. Thiết kế làm mát động cơ tên lửa là lĩnh vực đòi hỏi biên độ an toàn lớn. Nếu vách chảy một lần, cả động cơ sẽ vỡ. Vì vậy, người ta không tin ngay và dùng nguyên xi đáp án do trí tuệ nhân tạo đưa ra. Trí tuệ nhân tạo được dùng để tìm nhanh các ứng viên tốt. Phán đoán cuối cùng luôn được xác nhận bằng tính toán độ chính xác cao và thử nghiệm đốt thật. Các ví dụ trong chương này cũng phải được đọc trên nguyên tắc đó.

## 9.1 Cách tên lửa bảo vệ vách buồng đốt nóng

Phần này bàn về cách động cơ tên lửa bảo vệ vách nóng. Trước hết, ta xem vì sao buồng đốt tên lửa lại nóng đến như vậy. Sau đó, ta giải thích làm mát tái sinh là gì và vì sao khó chế tạo. Cuối cùng, ta xem in 3D kim loại đang thay đổi thiết kế này như thế nào. Phần này là nền tảng cho mọi nội dung phía sau.

Trong buồng đốt chính của động cơ tên lửa, nhiên liệu và oxy gặp nhau rồi cháy. Nhiệt độ của khí được tạo ra khi đó dễ dàng vượt quá 3 nghìn độ. Điểm nóng chảy của thép khoảng 1 nghìn 5 trăm độ. Khí cháy cao hơn điểm nóng chảy của thép rất nhiều. Hơn nữa, áp suất trong buồng đốt cũng cao. Động cơ hiệu suất cao có áp suất bên trong vượt quá một trăm lần áp suất khí quyển. Khí nóng và áp suất cao đẩy và nung nóng bức vách kim loại mỏng từ bên trong.

Áp suất cao cũng là một gánh nặng lớn. Áp suất cao trong buồng đốt đẩy bức vách mỏng ra ngoài. Nó giống như khi thổi căng một quả bóng, lớp vỏ trở nên căng cứng. Thêm vào đó là chênh lệch nhiệt độ lớn, bên trong nóng và bên ngoài lạnh. Phía nóng muốn giãn ra, còn phía lạnh muốn giữ nguyên. Sự giãn co này để lại ứng suất lớn trong vách. Có thể nói áp suất và nhiệt cùng lúc hành hạ bức vách. Vì vậy, vách phải mỏng, nhưng đồng thời phải được làm mát tốt và bền chắc.

Vị trí nhiệt đổ vào vách đạt mức cao nhất là cổ họng vòi phun. Vòi phun là bộ phận hình loa phụt khí ra phía sau để tạo lực đẩy. Phần trong chiếc loa ấy thu hẹp lại và thắt eo được gọi là cổ họng vòi phun(throat). Tại đây, khí tăng tốc đến vận tốc âm thanh. Nhiệt độ của bản thân khí lại giảm đôi chút khi đi qua vòi phun. Dù vậy, tại cổ họng vòi phun, nhiệt đi vào vách đạt mức cao nhất. Lý do là dòng khí nhanh quét rất mạnh qua bức vách hẹp. Lượng nhiệt đi vào vách trên một đơn vị diện tích được gọi là thông lượng nhiệt(heat flux). Thông lượng nhiệt ở cổ họng vòi phun rất lớn. Có thể hình dung nhiệt của vài chục bếp lò công suất lớn dồn vào một diện tích hẹp chỉ bằng lòng bàn tay.

Hãy thử hình dung thông lượng nhiệt lớn đến mức nào. Nhiệt do một họng bếp gas gia đình tạo ra trải trên một diện tích bằng lòng bàn tay. Ở cổ họng vòi phun, nhiệt của vài chục họng bếp như vậy dồn vào cùng một diện tích. Lượng nhiệt ấy đổ vào chỉ một lớp vách kim loại mỏng. Chênh lệch nhiệt độ giữa mặt trong và mặt ngoài của vách lên đến vài trăm độ. Chênh lệch lớn này làm kim loại cong vênh và gây nứt. Nếu không làm mát vách, vách sẽ sụp đổ chỉ trong vài giây.

Nếu vách nhận nguyên lượng nhiệt này, vách sẽ chảy hoặc yếu đi. Vì vậy, người ta dùng làm mát tái sinh. Người ta tạo nhiều đường dẫn nhỏ trong vách. Nhiên liệu được cho chảy qua các đường dẫn này trước. Khi đi qua đường dẫn, nhiên liệu lấy đi nhiệt của vách. Vách nguội xuống và nhiên liệu nóng lên. Nhiên liệu đã nóng đi vào buồng đốt rồi cháy. Cách này bảo vệ vách và gia nhiệt trước cho nhiên liệu, nên không vứt bỏ nhiệt. Đôi khi người ta còn thêm làm mát màng. Làm mát màng là phương pháp phủ một lớp màng chất làm mát mỏng dọc theo vách. Khi dùng hai phương pháp cùng nhau, có thể bảo vệ vách an toàn hơn.

Vấn đề là chế tạo các kênh làm mát này. Trước đây, người ta cắt gọt một khối kim loại để tạo đường dẫn. Cách làm bằng cắt gọt buộc hình dạng đường dẫn phải đơn giản. Thường thì người ta khoét các rãnh chữ nhật thẳng song song nhau. Những đường dẫn như vậy khó thay đổi lớn về bề rộng hoặc hướng theo từng vị trí. Rất khó chỉ làm mát mạnh hơn riêng khu vực cổ họng vòi phun, nơi nhiệt tập trung nhiều. Nếu cố làm đường dẫn hẹp lại, nhiên liệu sẽ khó đi qua và tổn thất áp suất tăng. Khi tổn thất áp suất lớn, bơm đẩy nhiên liệu vào phải làm việc mạnh hơn.

In 3D kim loại đang tháo gỡ giới hạn này. In 3D kim loại là phương pháp trải một lớp bột kim loại mỏng rồi dùng laser làm nóng chảy để kết dính. Bộ phận được tạo ra bằng cách chồng từng lớp một. Phương pháp này có thể tạo các đường dẫn bên trong phức tạp mà cắt gọt không làm được. Cũng có thể thiết kế đường dẫn tách nhánh như cành cây rồi gom lại. Cũng có thể tạo đường dẫn uốn theo mặt cong của

vách. Ở nơi nhiệt tập trung, có thể bố trí đường dẫn dày hơn; ở nơi khác, bố trí thưa hơn. Tự do thiết kế đã mở rộng rất lớn.

Hãy lấy một ví dụ để thấy in 3D thay đổi kênh làm mát đến mức nào. Cách cũ là khoét rãnh trên vách trong rồi phủ một nắp bên ngoài để tạo đường dẫn. Phải chế tạo nhiều bộ phận rồi ghép lại. Khi có nhiều mối nối, chính các vị trí đó có thể yếu đi. In 3D in ra một khối vách có sẵn đường dẫn bên trong. Mối nối giảm, nên vách chắc hơn. Các bước chế tạo cũng giảm, tiết kiệm thời gian và vật liệu. Vì vậy, việc sửa thiết kế nhiều lần rồi thử nghiệm trở nên dễ hơn.

Vật liệu cho kênh làm mát cũng phát triển cùng lúc. Cơ quan Hàng không và Vũ trụ Hoa Kỳ(NASA) đã phát triển một hợp kim đồng tên là GRCo-42. Đồng là kim loại truyền nhiệt tốt. GRCo-42 là hợp kim pha một ít chromium và niobium vào đồng. Nó vừa truyền nhiệt tốt, vừa chịu được nhiệt độ cao. Hợp kim này có thể được in 3D thành cả một buồng đốt có khắc sẵn đường dẫn. Cơ quan Hàng không và Vũ trụ Hoa Kỳ đã tiến hành cả thử nghiệm đốt thật đối với buồng đốt làm bằng hợp kim này. Kết quả phát triển và thử nghiệm này đã được tổng hợp trong báo cáo công khai. Cơ quan Hàng không và Vũ trụ Hoa Kỳ đã phát triển vật liệu và công nghệ chế tạo này trong chương trình công nghệ đẩy thế hệ tiếp theo. Đây là một ví dụ trong đó vật liệu, chế tạo và thiết kế cùng tiến lên như một khối.

Nếu dùng thêm làm mát màng, có thể bảo vệ vách an toàn hơn. Người ta khoan nhiều lỗ nhỏ trên vách. Qua các lỗ đó, một lượng nhỏ chất lưu lạnh được đưa chảy ra. Chất lưu chảy ra tạo thành một lớp màng mỏng dọc theo vách. Lớp màng này chen vào giữa khí nóng và vách. Nếu làm mát tái sinh làm nguội vách từ bên trong, thì làm mát màng phủ một lớp bảo vệ lên bề mặt vách. Hai phương pháp hỗ trợ nhau. Ở gần cổ họng vòi phun, nơi nhiệt tập trung nhiều, người ta đôi khi dùng chồng cả hai phương pháp.

Hình dạng của kênh làm mát quyết định rất lớn đến hiệu suất quản lý nhiệt. Nếu kênh hẹp và dày đặc, diện tích tiếp xúc với vách tăng. Khi diện tích tiếp xúc lớn, kênh lấy đi nhiều nhiệt hơn. Nhưng nếu kênh quá hẹp, nhiên liệu khó đi qua. Cần nhiều lực hơn để đẩy nhiên liệu vào. Mất mát lực này được gọi là tổn thất áp suất. Khi tổn thất áp suất lớn, bơm phải quay mạnh hơn. Kênh làm mát tốt phải làm mát vách tốt, đồng thời giữ tổn thất áp suất ở mức thấp. Hai điều này kéo nhau theo hai hướng ngược nhau.

Phương pháp mới không hề dễ. Một nghiên cứu công bố năm 2025 đã phân tích một trường hợp thất bại của buồng đốt GRCo-42 in 3D. Trong một thử nghiệm đốt vào tháng 2 năm 2021, buồng đốt đã vỡ chỉ sau 9 giây. Khi xem nguyên nhân, vấn đề nằm ở vị trí quá trình in từng dừng lại một lần. Tại vị trí đó, số lỗ rỗng li ti xuất hiện nhiều hơn bình thường rất nhiều. Khi có nhiều lỗ rỗng, phần đó yếu đi. Trường hợp này cho thấy phương pháp chế tạo mới nhạy cảm với quản lý chất lượng đến mức nào. Đối lại tự do thiết kế, ta phải theo dõi quá trình chế tạo kỹ hơn.

Tự do thiết kế rộng hơn cũng đem đến nỗi lo mới. Khi gần như có thể thay đổi hình dạng đường dẫn vô hạn, việc chọn hình dạng nào tốt lại trở nên khó hơn. Cách con người tự vẽ vài hình dạng rồi tính thử không thể quét hết không gian thiết kế rộng. Chỉ riêng việc tính chính xác dòng chảy và nhiệt của một đường dẫn cũng đã khiến máy tính mất nhiều thời gian. Trí tuệ nhân tạo bước vào công việc khám phá không gian rộng này một cách nhanh và thông minh. Từ phần sau, ta sẽ lần lượt xem xét các phương pháp đó.

## 9.2 Mô hình đại diện dự đoán biến đổi đột ngột của nhiên liệu trong kênh làm mát

Phần này bàn về lý do nhiên liệu trong kênh làm mát khó dự đoán. Methane hoặc dầu hỏa dùng làm chất làm mát thay đổi mạnh tính chất ở áp suất và nhiệt độ cao. Phân tích chính xác để tính dòng chảy thất thường này diễn ra chậm. Vì vậy, người ta dùng mô hình đại diện bằng trí tuệ nhân tạo để thay thế tính toán. Ở đây, ta giải thích nguyên lý và giới hạn của mô hình đại diện đó.

Chất làm mát của tên lửa thường là methane(methane) hoặc RP-1. RP-1 là dầu hỏa được tinh chế cho tên lửa. Có thể xem nó như họ hàng của dầu hỏa ta dùng. Các nhiên liệu này chịu áp suất rất cao trong kênh làm mát. Khi áp suất vượt qua một giá trị nào đó, ranh giới giữa chất lỏng và chất khí mờ đi. Lấy nước làm ví dụ, đó là trạng thái trong đó sự phân biệt nước trong nồi sôi lên và tạo bọt biến mất. Trạng thái này được gọi là trạng thái siêu tới hạn(supercritical). Nhiên liệu ở trạng thái siêu tới hạn mang tính chất lưỡng chùng, không hẳn là chất lỏng cũng không hẳn là chất khí.

Từ siêu tới hạn có thể nghe khó. Hãy nói dễ hơn. Khi đun nước trong nồi, nước sôi và biến thành bọt. Chất lỏng và chất khí được tách biệt rõ. Nhưng nếu tăng áp suất lên rất cao, hiện tượng sôi này biến mất. Chất lỏng dần nối tiếp sang tính chất giống chất khí. Ranh giới rõ ràng nơi bọt xuất hiện không còn nữa. Trạng thái mất ranh giới như vậy là trạng thái siêu tới hạn. Nhiên liệu trong kênh làm mát chịu áp suất cao, nên chảy ở trạng thái này.

Tính chất của trạng thái siêu tới hạn thay đổi gấp theo nhiệt độ. Khi đi qua một khoảng nhiệt độ nào đó, mật độ giảm mạnh. Trong cùng khoảng đó, nhiệt dung riêng đột ngột vọt lên. Nhiệt dung riêng là lượng nhiệt cần để tăng nhiệt độ lên 1 độ. Điểm mà mật độ và nhiệt dung riêng dao động như vậy được gọi là đường giả tới hạn(pseudo-critical line). Ở gần vùng này, tính chất của nhiên liệu thay đổi đột ngột như thể đang sôi. Thực tế nó không sôi, nhưng vì giống sôi nên được gọi là sôi giả(pseudo-boiling).

Việc nhiệt dung riêng vọt lên cũng có mặt tốt cho làm mát. Khi nhiệt dung riêng lớn, nhiên liệu có thể chứa nhiều nhiệt. Cùng một lượng nhiên liệu sẽ mang đi nhiều nhiệt hơn. Nhưng nếu mật độ giảm mạnh, nhiên liệu nở ra và chảy nhanh hơn. Khi dòng chảy đã nhanh dao động gần vách, cách nó vận chuyển nhiệt bị rối loạn. Mặt tốt và mặt xấu cùng xuất hiện tại một điểm. Vì vậy, việc dự đoán làm mát trong vùng này rất khó.

Sôi giả gây vấn đề cho việc làm mát. Thành chỉ nguội khi nhiên liệu lấy nhiệt tốt khỏi thành. Nhưng gần đường tới hạn giả, khả năng lấy nhiệt đôi khi giảm đột ngột. Lớp nhiên liệu đã nóng lên gần thành giống như một tấm chăn mỏng, chặn nhiệt lại. Hiện tượng này gọi là suy giảm truyền nhiệt (heat transfer deterioration). Khi suy giảm truyền nhiệt xảy ra, nhiệt độ thành tại vị trí đó tăng vọt. Một nghiên cứu năm 2020 đã tổng hợp rằng khi gia nhiệt nhiên liệu tên lửa, sôi giả và suy giảm truyền nhiệt xuất hiện cùng nhau. Một nghiên cứu năm 2024 đã đi sâu vào lý do xảy ra suy giảm này bằng mô hình dòng chảy hai trạng thái ở quy mô vi mô. Suy giảm truyền nhiệt này cũng liên quan đến vấn đề coking sẽ nói ở phần sau. Vì khi thành đột ngột nóng lên, nhiên liệu tại vị trí đó bị phân hủy nhanh hơn.

Có một phương pháp chính xác để tính loại dòng chảy này. Đó là động lực học chất lưu tính toán (computational fluid dynamics, CFD). Dòng chảy được chia nhỏ thành các ô lưới nhỏ, rồi vận tốc, áp suất và nhiệt độ của từng ô được giải theo các định luật vật lý. Nền tảng là ba định luật bảo toàn. Một là bảo toàn khối lượng. Lượng chất lưu đi vào và lượng chất lưu đi ra phải khớp nhau. Tiếp theo là bảo toàn động lượng. Lực đẩy hoặc kéo chất lưu phải khớp với sự thay đổi vận tốc. Cuối cùng là bảo toàn năng lượng. Nhiệt đi vào và nhiệt đi ra phải khớp với sự thay đổi nhiệt độ. CFD là việc giải đồng thời ba định luật này trong mọi ô. CFD chính xác nhưng chậm. Chỉ riêng việc giải chính xác một đoạn kênh làm mát cũng có khi khiến máy tính mất gần một giờ. Trong thiết kế kênh làm mát, phải kiểm tra hàng nghìn trường hợp bằng cách thay đổi hình dạng kênh và lưu lượng. Nếu mỗi lần tính đều mất một giờ, thiết kế sẽ không thể kết thúc.

Vì vậy người ta dùng mô hình thay thế (surrogate model). Mô hình thay thế là mô hình xấp xỉ nhanh dùng để thay cho phép tính chậm. Trước tiên, nhiều trường hợp được tính trước bằng CFD để thu thập dữ liệu. Dữ liệu đó được dùng để huấn luyện mạng nơ-ron nhân tạo. Mạng nơ-ron nhận đầu vào là hình dạng kênh, lưu lượng, áp suất và nhiệt độ cửa vào. Sau đó nó đưa ra ngay nhiệt độ thành, thông lượng nhiệt và tổn thất áp suất. Các nhà nghiên cứu tại Trung tâm Hàng không Vũ trụ Đức (DLR) đã tạo ra mô hình thay thế bằng mạng nơ-ron như vậy. Thứ họ thực sự dùng là một mạng nơ-ron thông thường gồm

vài lớp ẩn xếp chồng. Phép tính từng mất một giờ bằng CFD được mô hình này hoàn tất trong chưa đến 1 giây. Ban đầu họ chỉ xử lý kênh thẳng, sau đó mở rộng để xử lý cả kênh cong.

Quy trình tạo mô hình thay thế như sau. Trước tiên xác định phạm vi của các biến thiết kế. Đó là phạm vi của chiều rộng kênh, chiều cao kênh, lưu lượng và nhiệt độ cửa vào. Trong phạm vi đó, chọn nhiều trường hợp và tính bằng CFD. Kết quả tính toán được gom lại thành các cặp đầu vào và đầu ra. Những cặp này trở thành sách giáo khoa của mạng nơ-ron. Mạng nơ-ron nhìn các cặp đó và học quy luật đi từ đầu vào đến đầu ra. Khi học xong, nó cũng trả lời nhanh với đầu vào mới. Điều này giống như con người giải nhiều ví dụ để nắm được dạng bài.

Còn có những cách khác để tăng độ chính xác của mô hình thay thế. Các cách này không phải là thứ nghiên cứu DLR ở trên đã dùng nguyên xi, mà là các kỹ thuật được dùng rộng rãi để làm cho mô hình thay thế vững hơn. Một cách là đưa thêm tri thức vật lý vào. Nếu mạng nơ-ron chỉ học từ dữ liệu, nó có thể đưa ra câu trả lời vi phạm định luật vật lý. Vì vậy người ta đưa các định luật bảo toàn khối lượng, động lượng và năng lượng vào quy tắc học. Nếu mạng nơ-ron vi phạm các định luật đó, nó bị phạt. Mạng nơ-ron được tạo theo cách này gọi là mạng nơ-ron dựa trên vật lý (physics-informed neural network, PINN). Người ta cũng có thể thêm cấu trúc gọi là kết nối phần dư (residual connection). Kết nối phần dư là đường dẫn đưa thông tin từ lớp trước đi thẳng sang lớp sau bằng cách bỏ qua một số lớp ở giữa. Cấu trúc này giúp huấn luyện ổn định các mạng nơ-ron sâu. Việc đặt những kỹ thuật này lên mô hình thay thế cho kênh làm mát là hướng mà cuốn sách này đề xuất.

Các giá trị tính chất của chất lưu siêu tới hạn cũng phải được đưa vào chính xác. Viện Tiêu chuẩn và Công nghệ Quốc gia Hoa Kỳ (NIST) công bố các giá trị như mật độ và nhiệt dung riêng của nhiều chất lưu theo nhiệt độ và áp suất. Với chất lưu tinh khiết chỉ có một thành phần như methane, có thể nối và dùng trực tiếp các giá trị chuẩn này. Làm vậy mới không bỏ lỡ sự thay đổi đột ngột gần đường tới hạn giả. Nhưng RP-1, tức dầu hỏa, lại khác. RP-1 là hỗn hợp gồm nhiều hydrocarbon, nên thành phần thay đổi đôi chút theo từng mẻ tinh chế. Vì vậy không thể dùng nguyên bảng chất lưu tinh khiết, mà phải tạo riêng một mô hình hỗn hợp thay thế để khớp tính chất. Nếu giá trị tính chất không chính xác, mô hình thay thế cũng sẽ đưa ra nhiệt độ thành sai lệch. Chất lưu siêu tới hạn có tính chất nhạy, nên dữ liệu nền này rất quan trọng.

Cũng có nghiên cứu chọn chính chất làm mát cho tốt. Một nghiên cứu năm 2026 đã thiết kế nhiên liệu thay thế mô phỏng dầu hỏa dùng cho tên lửa. Nghiên cứu này dùng thuật toán di truyền để tìm tổ hợp khớp tốt các tính chất ở trạng thái siêu tới hạn. Thuật toán di truyền là phương pháp cải thiện đáp án bằng cách chọn và pha trộn các tổ hợp tốt. Mô hình nhiên liệu dùng trong tính toán làm mát càng chính xác, dự báo càng chính xác. Trí tuệ nhân tạo cũng góp sức vào việc chọn tổ hợp nhiên liệu.

Có điểm cần thận trọng khi xử lý mô hình thay thế. Không được dựa vào tên mô hình được gán như nhãn hiệu hoặc các con số hiệu năng chưa được kiểm chứng. Vì vậy cuốn sách này giải thích bằng khái niệm chung là mô hình thay thế CFD, thay vì nêu tên một mô hình cụ thể. Nếu nguyên lý giống nhau, tên gọi là gì thì điều học được vẫn như nhau.

Mô hình thay thế này cũng đã được dùng cho động cơ thực tế. Trung tâm Hàng không Vũ trụ Đức đã xem xét hiệu năng kênh làm mát của động cơ do họ tự phát triển bằng phương pháp này. Họ cải thiện dự báo nhiệt độ thành và phản ánh vào thiết kế. Khi tính toán nhanh hơn, họ có thể so sánh nhiều thiết kế kênh trong thời gian ngắn. Nghĩa là họ đã khảo sát thiết kế trên phạm vi rộng mà không bị trói vào phép tính chính xác chậm.

Mô hình thay thế có giới hạn rõ ràng. Mô hình thay thế chỉ đáng tin trong phạm vi nó đã học. Với hình dạng kênh lạ hoặc điều kiện vận hành lạ mà nó chưa thấy khi huấn luyện, nó có thể sai lớn. Chất lưu siêu tới hạn có tính chất nhạy, nên rủi ro này càng lớn. Vì vậy không được quyết định thiết kế cuối cùng chỉ bằng mô hình thay thế. Hãy dùng mô hình thay thế để lọc nhanh các ứng viên tốt. Sau đó chỉ chọn

các ứng viên đó và kiểm tra lại bằng CFD chính xác cùng thử nghiệm đốt thực tế. Mô hình thay thế không phải công cụ đưa ra đáp án thay con người, mà là công cụ lọc nhanh các ứng viên.

### 9.3 Học chủ động để tìm lỗ làm mát màng với ít phép tính

Trước hết hãy xem vì sao cần làm mát màng. Làm mát tái sinh làm nguội thành từ phía trong. Nhưng ở nơi nhiệt tập trung quá lớn, chỉ làm mát phía trong thành có thể không đủ. Vì chính bề mặt thành tiếp xúc trực tiếp với khí nóng và bị gia nhiệt. Vì vậy người ta phủ thêm một lớp màng bảo vệ mỏng lên bề mặt. Đó là làm mát màng. Làm mát tái sinh và làm mát màng cùng bảo vệ thành từ trong và ngoài.

Làm mát màng (film cooling) là phương pháp tạo một màng mỏng bằng chất lưu lạnh ngay trước thành nóng. Người ta khoan nhiều lỗ nhỏ trên thành. Qua các lỗ đó, chất lưu lạnh được phun xiên. Chất lưu được phun ra lan mỏng theo thành và phủ lên thành như một tấm chắn. Màng này tách khí nóng khỏi thành. Thành không tiếp xúc trực tiếp với khí nóng nên nóng lên ít hơn. Nó giống cánh luồng gió chống đóng băng trên kính chắn gió ô tô thổi từ dưới lên trên.

Hiệu năng làm mát màng do vài giá trị quyết định. Một giá trị là cường độ của chất lưu được thổi vào. Giá trị này gọi là tỷ số thổi (blowing ratio). Đó là tỷ số giữa cường độ dòng chảy của chất lưu lạnh và cường độ dòng chảy của khí nóng. Góc nghiêng của lỗ, khoảng cách giữa các lỗ và hình dạng lỗ cũng quyết định hiệu năng. Chỉ cần các giá trị này thay đổi một chút, hiệu năng của màng có thể thay đổi lớn. Màng bám tốt vào thành hay nhấc lên rồi bong ra phụ thuộc vào các giá trị này.

Tỷ số thổi quá lớn cũng gây vấn đề. Khi phun chất lưu quá mạnh, màng nhấc lên khỏi thành. Tia phun không phủ được thành mà vọt lên trên. Khi đó thành lại tiếp xúc với khí nóng. Nếu tỷ số thổi quá nhỏ, màng mỏng và nhanh chóng bị cuốn đi. Vì vậy cường độ cũng có giá trị phù hợp. Phải khớp đồng thời góc, khoảng cách và cường độ thì mới tạo được màng tốt.

Một lý do khiến màng bong ra là xoáy. Khi chất lưu phun ra từ lỗ, một cặp xoáy quay sang trái và phải hình thành. Cặp xoáy này kéo khí nóng về phía thành. Khi đó màng làm mát đã mất công tạo ra bị nâng lên và bong ra. Thành lại lộ ra trước khí nóng. Nếu thiết kế lỗ tốt, có thể làm cặp xoáy này yếu đi. Vì vậy hiện nay người ta dùng lỗ dạng quạt, với đầu lỗ mở rộng như chiếc quạt. Chất lưu được phun ra lan rộng sang hai bên và phủ đều thành.

Lỗ dạng quạt trải màng làm mát rộng hơn lỗ tròn. Lỗ tròn phun chất lưu thành một tia. Tia này dễ nhấc lên khỏi thành. Lỗ dạng quạt mở rộng đầu lỗ sang hai bên và phun chất lưu rộng ra. Chất lưu lan rộng bám vào thành tốt hơn. Khi diện tích bám tăng, thành được phủ đều. Vì vậy turbine khí và động cơ hàng không thường dùng lỗ dạng quạt. Mức độ mở rộng lỗ và góc khoan lỗ quyết định hiệu năng.

Muốn tìm hình dạng lỗ tốt, phải tính dòng chảy. Muốn quan sát chi tiết đến cả xoáy, cần phân tích chính xác. Phân tích này gọi là mô phỏng xoáy lớn (large eddy simulation, LES). LES là phương pháp tính trực tiếp các xoáy lớn trong dòng chảy. Nó chính xác nhưng chi phí tính toán rất cao. Ngay cả việc giải một điều kiện cũng mất nhiều thời gian. Rất khó tính hết hàng trăm trường hợp trong khi thay đổi góc lỗ, khoảng cách và cường độ. Vì ngân sách tính toán có hạn, không thể tính bừa mọi điều kiện.

Ở đây, quá trình Gaussian (Gaussian process) giúp ích. Quá trình Gaussian là phương pháp dự báo bằng ít dữ liệu. Điểm riêng của nó là không chỉ đưa ra giá trị dự báo. Nó cũng đưa ra mức bất định của dự báo. Ở nơi có nhiều dữ liệu, nó dự báo tự tin. Ở nơi không có dữ liệu, nó đánh dấu là bất định. Điều này giống như khi vẽ bản đồ, con đường đã đi thì vẽ chi tiết, con đường chưa đi thì vẽ mờ. Phương pháp này nối các giá trị bằng một quy tắc trơn gọi là Matern.

Hãy xem cả cách quá trình Gaussian nối các giá trị. Các điều kiện gần nhau thì hiệu năng cũng có thể giống nhau. Nếu chỉ thay đổi góc lỗ một chút, hiệu năng làm mát cũng chỉ thay đổi một chút. Kernel là thứ đưa ý tưởng này vào toán học. Kernel là thước đo hai điều kiện gần nhau đến mức nào. Nếu gần, nó

xem giá trị là tương tự; nếu xa, nó xem giá trị có thể khác. Bằng thước đo này, nó nối trôn các điều kiện đã tính. Nó cũng ước lượng được giá trị của điều kiện chưa tính.

Học chủ động (active learning) thêm lựa chọn thông minh vào đây. Nó tự chọn điều kiện nào sẽ được tính bằng LES tiếp theo. Khi chọn, nó nhìn đồng thời hai điểm. Một là nơi có vẻ có hiệu năng tốt. Điểm còn lại là nơi bất định. Đào sâu nơi có vẻ có hiệu năng tốt sẽ đưa ta đến gần đáp án tốt hơn. Đào sâu nơi bất định sẽ làm phần mờ trên bản đồ trở nên rõ hơn. Quy tắc cân bằng hai điều này gọi là hàm thu nhận (acquisition function). Quy tắc thường dùng là cách đo mức cải thiện kỳ vọng so với hiện tại.

Ưu điểm của phương pháp này nằm ở việc sắp xếp thứ tự tính toán một cách thông minh. Nó không tính bừa mọi điều kiện. Nó chọn tính trước các điều kiện có nhiều thứ để học. Vì vậy ngay cả với ít phép tính LES, nó vẫn tiến gần được đến hình dạng lỗ tốt. Nó tiết kiệm phép tính đắt đỏ và nhanh chóng thu hẹp thiết kế. Cách này trước hết đã có chỗ đứng trong thiết kế lỗ làm mát của cánh turbine khí. Các nghiên cứu đã được kiểm chứng đến nay chủ yếu nhằm vào turbine khí. Người ta đã cải thiện đồng thời hiệu năng làm mát và độ đồng đều bằng tối ưu hóa Bayes xét nhiều mục tiêu cùng lúc. Một nghiên cứu báo cáo rằng trong cùng điều kiện tổn thất áp suất, hiệu năng làm mát đã cao hơn thiết kế trước đó.

Trí tuệ nhân tạo cũng được dùng cho chính phép tính dòng chảy. Các nghiên cứu vào khoảng năm 2024 đã tạo ra mạng nơ-ron dự báo hiệu suất làm mát màng. Một nghiên cứu đã dự báo hiệu suất làm mát ngay cả khi độ cong của thành và góc lỗ thay đổi. Các mô hình dự báo như vậy đưa ra giá trị gần với phép tính chính xác nhưng nhanh hơn nhiều. Chúng giúp so sánh nhanh nhiều hình dạng lỗ ở giai đoạn đầu thiết kế. Cũng có thể kết hợp hai phương pháp, theo kiểu quá trình Gaussian chọn phép tính tiếp theo, còn mạng nơ-ron ước lượng nhanh phép tính đó.

Làm mát màng trong buồng đốt tên lửa cũng mượn cùng phương pháp. Nhưng điều kiện của tên lửa khắc nghiệt hơn nhiều. Khí nóng hơn và nhanh hơn. Vì vậy không được chuyển nguyên các giá trị lỗ tối ưu thu được từ động cơ hàng không sang tên lửa. Khi kích thước lỗ, vận tốc chất lưu, độ cong thành và loại nhiên liệu thay đổi, kết quả cũng thay đổi. Có thể chuyển phương pháp thiết kế tốt, nhưng các con số tối ưu cụ thể phải được tìm lại mỗi lần. Trí tuệ nhân tạo là công cụ giúp việc tìm lại đó nhanh hơn. Hình dạng cuối cùng phải được kiểm chứng lại trong điều kiện tên lửa.

## 9.4 Thiết kế hình dạng kênh làm mát bằng trí tuệ nhân tạo

Thiết kế kênh làm mát là một bài toán trong đó chất rắn và chất lưu đan xen với nhau. Nhiệt của khí nóng trước hết đi vào thành kim loại. Nhiệt đó truyền qua kim loại về phía kênh làm mát. Trong kênh, nhiên liệu đang chảy lấy nhiệt đó đi. Dẫn nhiệt chạy trong kim loại và đối lưu do chất lưu lấy nhiệt gặp nhau tại cùng một chỗ. Phép tính giải đồng thời hai hiện tượng này được gọi là truyền nhiệt liên hợp (conjugate heat transfer, CHT). Trong kênh làm mát của tên lửa, chất rắn và chất lưu liên kết chặt chẽ với nhau, nên phép tính này rất quan trọng.

Có thể hiểu dễ hơn nếu nghĩ đến cửa sổ mùa đông. Trong phòng thì ấm, bên ngoài thì lạnh. Nhiệt độ mặt trong và mặt ngoài của kính khác nhau. Nhiệt chảy qua kính, và không khí lạnh bên ngoài lấy nhiệt đi. Dẫn nhiệt trong kính và đối lưu của không khí bên ngoài gặp nhau trên bề mặt cửa sổ. Thành tên lửa cũng như vậy. Nhiệt đi vào từ phía khí nóng, và nhiên liệu lấy nhiệt đi ở phía kênh làm mát. Hai dòng đó gặp nhau ở thành và đạt cân bằng.

Điểm cốt lõi của truyền nhiệt liên hợp là điều kiện tại biên. Hãy nghĩ đến mặt tiếp xúc giữa kim loại và nhiên liệu. Nhiệt đi từ kim loại ra mặt đó và nhiệt mà nhiên liệu nhận tại mặt đó phải bằng nhau. Vì nhiệt không tự mất đi và cũng không tự sinh ra. Phải tính nhiệt độ thành trong khi giữ đúng điều kiện này. Phép tính này cho biết chỗ nào trên thành nóng nhất, và phải làm mát chỗ đó như thế nào.

Tính toán truyền nhiệt liên hợp đòi hỏi khối lượng tính toán lớn. Vì phải giải cùng lúc dẫn nhiệt trong chất rắn và đối lưu trong chất lưu. Hơn nữa, chất lưu là dòng rối. Dòng rối là dòng chảy trong đó các xoáy lớn nhỏ trộn lẫn với nhau. Càng giải chi tiết dòng chảy này thì thời gian tính toán càng tăng. Ngay cả việc giải chính xác một kênh làm mát cũng khiến máy tính mất nhiều thời gian. Nếu có hàng nghìn phương án thiết kế, phải lặp phép tính đó hàng nghìn lần. Vì vậy cần có cách giảm tính toán.

Bây giờ đến lúc thiết kế hình dạng kênh. Ở đây dùng tối ưu hóa topo (topology optimization). Tối ưu hóa topo là phương pháp tìm xem trong một không gian đã định, chỗ nào sẽ được lấp bằng vật liệu và chỗ nào sẽ để trống. Với kênh làm mát, đó là bài toán quyết định chỗ nào giữ kim loại và chỗ nào để thành kênh rỗng cho nhiên liệu chảy qua. Con người không định sẵn hình dạng kênh. Chỉ cần đặt mục tiêu, quá trình tối ưu hóa sẽ tự vẽ ra hình dạng kênh. Mục tiêu là hạ nhiệt độ cao nhất của thành. Đồng thời cũng không để tổn thất áp suất tăng quá mức. Đã có các nghiên cứu thiết kế kênh làm mát bằng tối ưu hóa topo. Tiêu biểu là phương pháp dựa trên mật độ, vừa xem dòng chảy và nhiệt, vừa thay đổi dần cách bố trí vật liệu.

Bài toán này có quá nhiều biến. Không gian được chia thành hàng trăm nghìn ô nhỏ. Ở mỗi ô phải quyết định đó là kim loại hay khoảng trống. Số trường hợp là vô kể. Nếu tìm trực tiếp trong không gian rộng như vậy thì thời gian sẽ kéo dài không dứt. Vì vậy cần có cách thu nhỏ bài toán. Ở đây bộ tự mã hóa tích chập (convolutional autoencoder, CAE) xuất hiện.

Bộ tự mã hóa là mạng nơ-ron nén dữ liệu phức tạp thành dạng nhỏ rồi khôi phục lại. Phần trước nén hình dạng kênh thành một cụm số nhỏ. Cụm số này được gọi là vector tiềm ẩn (latent vector). Phần sau khôi phục hình dạng kênh từ vector tiềm ẩn đó. Nếu được huấn luyện tốt, chỉ với ít con số vẫn có thể chứa được hình dạng cốt lõi của kênh. Nó giống như nén một bức ảnh thành tệp nhỏ rồi mở lại. Tích chập là cách tính cùng nhìn các phần lân cận của hình ảnh. Nó phù hợp để xử lý dữ liệu trải trong không gian như hình dạng kênh.

Các trường hợp kiểm chứng áp dụng trực tiếp phương pháp dùng bộ tự mã hóa để thu nhỏ không gian thiết kế này cho kênh làm mát tên lửa vẫn còn hiếm. Cuốn sách này đề xuất hướng đó. Bản thân tối ưu hóa topo là phương pháp đã được kiểm chứng, và ý tưởng ở đây là đặt thêm bộ tự mã hóa lên trên để tăng tốc quá trình tìm kiếm.

Việc nén này làm cho thiết kế dễ hơn. Bộ tối ưu hóa không trực tiếp xử lý hàng trăm nghìn ô. Thay vào đó, nó chỉ thay đổi qua lại một vector tiềm ẩn nhỏ. Khi thay đổi vector tiềm ẩn một chút, phần sau tạo ra một hình dạng kênh có vẻ hợp lý. Có thể xem đó là việc thu nhỏ không gian thiết kế rộng thành một không gian nhỏ rồi tìm kiếm trong đó. Bộ tối ưu hóa tìm một vector tiềm ẩn tốt trong không gian nhỏ này. Kênh do vector đó tạo ra sẽ trở thành kênh làm mát tốt.

Dùng vector tiềm ẩn còn có một điểm tốt nữa. Hình dạng kênh được tạo ra nhìn chung có vẻ hợp lý. Vì bộ tự mã hóa đã học từ dữ liệu kênh thực tế. Nó tạo hình dạng mới trong khuôn dạng đã học. Vì vậy ít sinh ra các hình dạng kỳ lạ không hợp lý về mặt vật lý. Nó giảm công sức lãng phí khi tìm kiếm bừa trong một không gian rộng. Vì chỉ tìm câu trả lời trong một không gian hẹp nhưng giàu thông tin, quá trình tìm kiếm nhanh hơn. Tuy nhiên, việc hình dạng giống dữ liệu học không bảo đảm đó là kênh tốt. Phải kiểm tra riêng xem kênh có nổi thông suốt từ đầu đến cuối hay không, thành có quá mỏng hay không, sức cản dòng chảy và độ bền kết cấu có ổn hay không. Những điều kiện như vậy phải được đặt riêng thành quy tắc trong quá trình tối ưu hóa.

Cũng phải đưa vào điều kiện rằng hình dạng đó có thể chế tạo được. In 3D tạo vật thể bằng cách xếp lớp từ dưới lên trên. Phần nghiêng ngang quá nhiều mà không có giá đỡ phía dưới sẽ bị võng khi chế tạo. Vì vậy đặt ràng buộc để thành kênh không nghiêng quá một góc nhất định. Khi giữ ràng buộc này, kênh có thể tự đứng mà không cần giá đỡ. Nếu đưa cả các điều kiện này vào quá trình tối ưu hóa, thiết kế thu

được có thể đem đi in ngay. Điều đó ngăn việc một thiết kế tốt trong tính toán lại trở thành thiết kế không thể chế tạo trên thực tế.

Cũng có nghiên cứu dùng trí tuệ nhân tạo để thay cho tính toán truyền nhiệt liên hợp. Một nghiên cứu năm 2024 đã tạo một mạng nơ-ron xếp nhiều tầng bộ mã hóa và bộ giải mã. Mạng nơ-ron này nhận hình dạng linh kiện làm đầu vào và vẽ ngay phân bố nhiệt độ. Tức là thay vì chạy phép tính truyền nhiệt liên hợp chính xác mỗi lần, nó chuyển sang dự báo nhanh. Một nghiên cứu khác năm 2025 chia chất rắn và chất lưu cho hai mạng nơ-ron khác nhau xử lý. Hai mạng nơ-ron được huấn luyện để khớp giá trị với nhau tại biên. Làm như vậy bắt tốt hơn sự thay đổi nhiệt độ gấp ở mặt tiếp xúc giữa thành và nhiên liệu. Các nghiên cứu này xử lý bài toán truyền nhiệt liên hợp nói chung. Nếu dùng các mô hình dự báo nhanh như vậy trong tối ưu hóa topo của kênh làm mát tên lửa, quá trình tìm kiếm thiết kế sẽ nhanh hơn. Sự kết hợp đó là hướng cần tiếp tục tinh chỉnh trong tương lai.

Kênh do tối ưu hóa topo vẽ ra có hình dạng con người khó vẽ bằng tay. Các nhánh tách ra rồi nhập lại, uốn theo mặt cong của thành. Cũng có khi các cánh nghiêng xếp dày đặc như xương cá. Những hình dạng như vậy phân bố nhiệt đều và dẫn dòng chảy êm hơn. Dù mắt người thấy không đẹp, trong tính toán nó có thể là hình dạng tốt. Trí tuệ nhân tạo tìm ra những hình dạng bất ngờ như vậy.

Đã có trường hợp công nghiệp của cách này. Một công ty đã áp dụng công nghệ tối ưu hóa topo như vậy cho kênh làm mát tên lửa. Họ công bố rằng kết quả là nhiệt độ cao nhất của thành đã giảm mạnh. Gần đây cũng có nghiên cứu dùng bề mặt cong có quy luật cho kênh làm mát. Bề mặt cong này được gọi là bề mặt tối thiểu tuần hoàn ba chiều (triply periodic minimal surface, TPMS). TPMS là cấu trúc trong đó các bề mặt cong mềm như màng bong bóng xà phòng lặp lại có quy luật trong không gian. Cấu trúc này gấp một bề mặt rộng vào một không gian hẹp. Khi bề mặt rộng, nó trao đổi nhiệt nhiều hơn. Một nghiên cứu năm 2025 đã tạo kênh TPMS ôm theo mặt cong của thành. Nghiên cứu đó báo cáo rằng hiệu suất làm mát và độ đồng đều nhiệt độ quanh thành cùng tăng. Năm 2026 cũng xuất hiện nghiên cứu tăng truyền nhiệt bằng kênh làm mát lắp trong mạng lưới. Các kênh cong và kênh mạng lưới như vậy chỉ có thể chế tạo nhờ in 3D. Tuy nhiên, trường hợp công ty khác với bài báo học thuật và cách kiểm chứng cũng khác. Trong cuốn sách này, nó chỉ được xem như một trường hợp tham khảo cho thấy dòng chảy công nghiệp. Hiệu suất cuối cùng chỉ nên được dùng làm cơ sở thiết kế khi có dữ liệu thử nghiệm độc lập.

## 9.5 Tối ưu hóa đa mục tiêu để giảm coking, hiện tượng muội than tích tụ

RP-1, một loại nhiên liệu gốc dầu hỏa, có một điểm yếu khi được dùng làm chất làm mát. Khi đi gần thành nóng, nhiên liệu bị nhiệt tách vỡ. Hiện tượng này được gọi là nhiệt phân. Trong các mảnh bị tách ra, carbon bám cháy vào thành. Nó giống như dầu để lâu trên chảo rồi cháy đen và bám dính. Lớp vảy carbon bám lên thành như vậy được gọi là coking. Coking tích tụ dần bên trong kênh làm mát.

Việc nhiên liệu bị nhiệt tách vỡ được gọi là nhiệt phân. Đó là quá trình phân tử lớn nhận nhiệt rồi tách thành các mảnh nhỏ. Nhiệt độ càng cao, sự tách vỡ càng nhanh. Trong các mảnh bị tách ra, những phần giàu carbon kết lại với nhau. Carbon kết tụ bám cháy lên thành và tạo thành vảy. Vì vậy thành càng nóng thì coking càng dễ xảy ra. Coking phụ thuộc đồng thời vào nhiệt độ thành, loại nhiên liệu và thời gian dòng chảy đi qua.

Coking dễ xảy ra ở cổ họng vòi phun. Cổ họng vòi phun là nơi thông lượng nhiệt lớn nhất. Thành nóng đến mức đó nên sự phân hủy nhiên liệu cũng mạnh. Coking gây ra hai vấn đề. Một là lớp vảy carbon truyền nhiệt kém. Lớp vảy chặn giữa thành và nhiên liệu, khiến thành nóng hơn. Hai là lớp vảy làm hẹp kênh. Khi kênh hẹp lại, nhiên liệu khó đi qua hơn và tổn thất áp suất tăng. Khi hai vấn đề chồng lên nhau, thành trở nên nguy hiểm. Trong động cơ tái sử dụng, coking tích tụ qua nhiều lần thử nghiệm nên vấn đề càng lớn hơn.

Có quy luật trong quan hệ giữa nhiệt độ và coking. Chỉ cần nhiệt độ tăng một chút, phản ứng hóa học đã nhanh lên nhiều. Khi nhiệt độ thành tăng, phân hủy nhiên liệu nhanh lên rõ rệt. Coking cũng tích tụ nhanh theo. Vì vậy chỉ cần hạ nhiệt độ thành một chút cũng có thể giảm coking nhiều. Làm mát tốt thì thành nguội đi, và khi thành nguội thì coking giảm. Làm mát và ức chế coking liên kết với nhau như vậy.

Sự suy giảm truyền nhiệt ở mức trước gắn với coking này. Khi nhiên liệu siêu tới hạn đi qua đường giả tới hạn, truyền nhiệt gần thành đột ngột kém đi. Khi đó nhiệt độ thành tại chỗ đó tăng vọt. Khi thành nóng lên, nhiên liệu bị tách vỡ nhanh hơn và coking tích tụ nhanh hơn. Coking đã tích tụ lại chặn nhiệt nhiều hơn, khiến thành nóng hơn nữa. Nếu vòng phản hồi này tiếp diễn, thành có thể đột ngột hỏng. Có thể xem suy giảm truyền nhiệt và coking là hai hiện tượng làm nhau lớn lên.

Muốn giảm coking thì phải hạ nhiệt độ thành. Khi thành bớt nóng, nhiên liệu ít bị tách vỡ hơn. Cùng với đó, phải nhìn cả dòng chảy gần thành. Nếu nhiên liệu lướt nhanh qua thành, nó giảm khoảng thời gian để vảy bám lại. Lực mà chất lưu quét dọc theo thành ở gần thành được gọi là ứng suất cắt thành. Lực này phải đủ lớn ở một mức nhất định thì vảy mới khó bám. Thiết kế ức chế coking là bài toán điều chỉnh đồng thời nhiệt độ thành và ứng suất cắt thành.

Vấn đề là có nhiều mục tiêu và chúng va chạm với nhau. Có mục tiêu hạ nhiệt độ thành. Có mục tiêu giảm tổn thất áp suất. Có mục tiêu ức chế coking. Các mục tiêu này không cùng tốt lên với nhau. Nếu làm kênh hẹp để dòng chảy nhanh hơn, thành sẽ mát hơn nhưng tổn thất áp suất tăng. Nếu làm kênh rộng để dòng chảy chậm lại, tổn thất áp suất giảm nhưng làm mát yếu đi. Được một phía thì mất phía khác. Bài toán trong đó nhiều mục tiêu va chạm như vậy được gọi là tối ưu hóa đa mục tiêu.

Sự đánh đổi như vậy cũng quen thuộc trong đời sống. Hãy nghĩ đến lúc chọn một chiếc túi. Khó có một chiếc túi vừa nhẹ, vừa bền, vừa rẻ. Nhẹ thì dễ yếu, bền thì thường nặng hoặc đắt. Vì vậy chúng ta chọn điểm cân bằng phù hợp với cách dùng. Thiết kế kênh làm mát cũng giống như vậy. Không thể làm cho làm mát, tổn thất áp suất và ức chế coking đều đạt mức tốt nhất cùng lúc. Được một phía thì phải nhường một chút ở phía khác.

Lời giải của tối ưu hóa đa mục tiêu không chỉ có một. Thay vào đó, ta có một tập gồm nhiều lời giải tốt đánh đổi lẫn nhau. Tập này được gọi là biên Pareto (Pareto front). Các lời giải trên biên Pareto có các điểm cân bằng khác nhau. Có lời giải ưu tiên làm mát, có lời giải ưu tiên tổn thất áp suất. Nhà thiết kế nhìn tập này và chọn điểm cân bằng phù hợp với động cơ. Thay vì con người định trước một đáp án duy nhất, cách này bày ra các ứng viên tốt để chọn.

Ở đây dùng tối ưu hóa Bayes (Bayesian optimization). Đây là cách mở rộng quy trình Gaussian và học chủ động ở mức trước sang nhiều mục tiêu. Trước hết, nó vẽ bản đồ sơ bộ của các mục tiêu bằng ít phép tính. Nhìn vào bản đồ đó, nó chọn điều kiện nào sẽ được tính tiếp theo. Nó tính trước các điều kiện giúp mở rộng biên Pareto nhiều. Nhờ vậy có thể tìm ra các điểm cân bằng tốt trong khi giảm số lần phân tích cháy đất đỏ. Nếu đưa cả coking vào mục tiêu, ta có một thiết kế cân nhắc cùng lúc nhiệt độ thành, ứng suất cắt và tổn thất áp suất. Tuy nhiên, trong kênh làm mát tên lửa, vẫn khó tìm thấy trường hợp kiểm chứng đã tối ưu hóa với coking làm mục tiêu. Vật lý của nhiệt phân và lắng đọng đã được biết rõ, nên cuốn sách này đề xuất hướng đưa vật lý đó vào mục tiêu để thiết kế.

Hiện tượng coking trở thành bài toán khó hơn trong động cơ tái sử dụng. Nhưng điều đó không có nghĩa là động cơ dùng một lần rồi bỏ không gặp coking. Ngay trong một lần đốt, nếu lớp cặn làm hẹp đường dẫn hoặc làm truyền nhiệt xấu đi, thành động cơ có thể bị phá hỏng. Dù có tái sử dụng hay không, điều quan trọng là lớp cặn tích tụ nhanh đến mức nào trong thời gian nhiệm vụ. Với động cơ dùng lại nhiều lần, vấn đề tích lũy còn cộng thêm vào đó. Mỗi lần thử đốt, lớp cặn lại tích lại từng chút một. Lớp cặn đã tích tụ làm khả năng làm mát ngày càng xấu đi. Vì vậy, động cơ tái sử dụng phải được thiết kế để kìm hãm coking ngay từ đầu. Tối ưu hóa bằng trí tuệ nhân tạo giúp thiết kế nhìn trước cả sự thay đổi trong thời gian dài này.

Phương pháp này cũng có thể được nối với dữ liệu thử nghiệm tên lửa. Bãi thử Lampoldshausen của Trung tâm Hàng không Vũ trụ Đức là nơi đốt động cơ tên lửa thật để thu dữ liệu. Khi có dữ liệu thử nghiệm như vậy, mô hình nền của tối ưu hóa trở nên chính xác hơn. Dữ liệu lắp vào các hành vi coking thực tế mà tính toán thuần túy dễ bỏ sót. Đây là dòng công việc qua lại giữa thử nghiệm và tính toán để tinh chỉnh thiết kế.

Các đường dẫn có tỷ lệ hình dạng cao cũng được dùng để kim hãm coking. Tỷ lệ hình dạng là giá trị lấy chiều cao của đường dẫn chia cho chiều rộng. Đường dẫn hẹp và sâu có tỷ lệ hình dạng cao. Loại đường dẫn này có diện tích tiếp xúc với thành lớn, nên lấy nhiệt tốt. Thành mát hơn, nên sự phân hủy nhiên liệu giảm. Cũng có khi dùng đường dẫn cuốn quanh như xoắn ốc. Khi làm dòng chảy xoáy, nó quét mạnh hơn vùng gần thành. Tuy nhiên, loại đường dẫn này có thể làm tổn thất áp suất tăng, nên phải cân nhắc cùng lúc.

Phần này cũng có điểm cần thận trọng. Các con số về mức giảm coking chỉ đáng tin khi được thử nghiệm thực tế chống đỡ. Vì vậy, cuốn sách này tập trung vào nguyên lý của phương pháp thay vì nêu con số cụ thể. Trong động cơ thực tế, thành phần nhiên liệu, độ nhám bề mặt, tác dụng xúc tác của thành và thời gian vận hành đều ảnh hưởng đến coking. Có nhiều biến số, nên dự đoán khó. Giá trị dự đoán coking chưa được thử nghiệm chống đỡ phải được xử lý thận trọng. Trí tuệ nhân tạo được dùng để thu hẹp các ứng viên tốt, còn quyết định cuối cùng dựa trên thử nghiệm đốt thực tế.

## 9.6 Học máy có bổ sung vật lý để thiết kế inducer cánh trước bơm

Động cơ tên lửa phải đẩy nhiên liệu và oxy vào buồng đốt với lực lớn. Bộ phận làm việc này là turbopump. Turbopump là bơm quay với tốc độ rất cao. Nhưng cánh chính của bơm không thể làm việc đúng nếu áp suất đầu vào thấp. Vì vậy, người ta đặt một cánh phụ phía trước cánh chính. Cánh phụ này là inducer. Inducer trước hết kéo chất lỏng vào một cách êm hơn và nâng áp suất lên một chút. Nhờ đó, cánh chính phía sau quay ổn định. Nó là bộ phận giống như người gác cổng của bơm.

Hãy xem từ lý do vì sao cần turbopump. Bên trong buồng đốt có áp suất rất cao. Muốn đẩy nhiên liệu vào đó, cần áp suất cao hơn nữa. Nhiên liệu có áp suất thấp không thể đi vào buồng đốt có áp suất cao. Điều này giống như nước không tự chảy từ nơi thấp lên nơi cao. Vì vậy, bơm nâng áp suất nhiên liệu lên rất lớn. Turbopump của tên lửa làm việc này bằng chuyển động quay rất nhanh. Turbopump của động cơ lớn quay hàng chục nghìn vòng trong 1 phút.

Inducer khó xử lý hơn khi làm việc với hydro lỏng. Hydro lỏng (liquid hydrogen) là chất lỏng rất lạnh, khoảng âm 253 độ. Nó có mật độ thấp nên nhẹ. Khi inducer quay nhanh, ở phía trước cánh sẽ xuất hiện nơi áp suất tụt mạnh. Chất lỏng sôi khi áp suất hạ đủ thấp. Dù không nóng, chỉ cần áp suất thấp cũng có thể sinh bọt khí. Nguyên lý này giống như nước dễ sôi ở nhiệt độ thấp trên núi. Hiện tượng bọt khí sinh ra trong chất lỏng đang chảy như vậy gọi là cavitation.

Cavitation gây khó cho inducer. Bọt khí đã sinh ra sẽ bị ép xẹp trong chớp mắt khi chảy đến nơi áp suất cao trở lại. Khi bọt khí sụp xuống, sóng áp suất rất mạnh xuất hiện. Sóng này bào mòn dần bề mặt cánh. Nếu kéo dài, cánh sẽ bị rỗ. Cavitation cũng làm giảm hiệu suất bơm. Khi bọt khí nhiều lên, áp suất mà bơm đẩy lên có thể sụp đột ngột. Nếu vượt qua điểm sụp này, toàn bộ động cơ trở nên nguy hiểm. Vì vậy, inducer phải được thiết kế sao cho cavitation xảy ra muộn.

Có một tính chất hữu ích. Chất lỏng cực lạnh có hiệu ứng kim hãm nhiệt động lực học. Để bọt khí hình thành, nó phải lấy nhiệt ẩn hóa hơi từ chất lỏng xung quanh. Khi đó, nhiệt độ quanh bọt khí giảm nhẹ. Khi nhiệt độ giảm, áp suất sôi tại vị trí đó cũng giảm. Vì vậy, bọt khí khó lớn thêm. Có thể xem như bọt khí tự kìm sự phát triển của chính nó. Hydro lỏng có hiệu ứng kim hãm này lớn, nên cavitation bị trì hoãn một chút. Thiết kế inducer phải đưa cả hiệu ứng này vào tính toán. Các quy tắc vật lý của học máy

có bổ sung vật lý cũng phải đưa hiệu ứng này vào. Nếu không, hiệu năng cavitation sẽ bị đánh giá sai là xấu hơn thực tế.

Thiết kế inducer là việc quyết định hình dạng cánh. Góc cánh, độ cong và mức quét về phía trước quyết định hiệu năng. Có ba mục tiêu. Phải nâng áp suất tốt. Hiệu suất phải cao. Cavitation phải xảy ra muộn. Các mục tiêu này cũng va chạm với nhau. Nếu dựng cánh lên để nâng áp suất mạnh, cavitation dễ sinh ra. Nếu đặt cánh nằm xuống để tránh cavitation, áp suất tăng ít hơn. Đây là bài toán cân bằng giống các phần trước.

Năng lực nâng áp suất của inducer có một đường giới hạn. Giả sử ta cho bơm chạy trong khi hạ dần áp suất đầu vào. Đến một điểm nào đó, bơm vẫn chịu được tốt. Nhưng nếu áp suất đầu vào rơi xuống dưới một giá trị nào đó, cavitation trở nên nghiêm trọng. Ngay lúc đó, áp suất mà bơm nâng lên sụp đổ đột ngột. Điểm sụp này được gọi là điểm chặn cột áp. Cột áp là độ lớn của lực mà bơm dùng để đẩy chất lưu lên. Trong bơm thực tế, người ta phân biệt điểm bọt khí bắt đầu sinh ra và điểm cột áp sụp hoàn toàn. Có khi điểm cột áp giảm 3 phần trăm ở giữa hai điểm đó được lấy làm mốc. Giá trị này gọi là NPSH3. Cavitation ban đầu, NPSH3 và sụp đổ cột áp hoàn toàn là các điểm khác nhau. Inducer tốt trì hoãn điểm chặn này đến áp suất đầu vào thấp hơn. Nhờ vậy, nó vẫn quay ổn định ngay cả ở áp suất thấp hơn. Tính chất này cũng liên quan đến khối lượng tên lửa. Nếu inducer chịu được áp suất thấp tốt, áp suất của bình nhiên liệu có thể đặt thấp hơn. Khi áp suất bình thấp, thành bình có thể làm mỏng hơn. Khi bình nhẹ hơn, toàn bộ tên lửa nhẹ hơn. Thiết kế của một cánh nhỏ có thể làm thay đổi cả khối lượng tên lửa.

Ở đây dùng học máy có bổ sung vật lý (physics-enhanced machine learning). Học máy chỉ học dữ liệu thuần túy có thể vi phạm các định luật vật lý. Vì vậy, các định luật bảo toàn của chất lưu và vật lý cavitation được đưa cùng vào quá trình học. Trường dòng chảy được ước lượng nhanh bằng toán tử nơ-ron Fourier (Fourier neural operator, FNO). Toán tử này là mạng nơ-ron dự đoán nhanh toàn bộ bức tranh của dòng chảy trong một lần. Nếu mạng nơ-ron thông thường dự đoán giá trị tại một điểm, thì toán tử này dự đoán trọn cả trường dòng chảy. Vì nó chia dòng chảy thành các sóng có nhiều kích thước để xử lý, nó bắt được cả dòng lớn và gọn nhỏ. Vì vậy, nó phù hợp để ước lượng nhanh các giá trị lan rộng trong không gian như trường dòng chảy. Khi thay đổi hình dạng cánh, ta có thể nhanh chóng xem xét dòng chảy và rủi ro cavitation. Nó nhanh hơn nhiều so với chạy tính toán chính xác mỗi lần. Bản thân toán tử nơ-ron Fourier là phương pháp đã được kiểm chứng trong nhiều bài toán dòng chảy. Việc dùng nó cho thiết kế ngược inducer hydro lỏng là hướng mà cuốn sách này đề xuất.

Trong thiết kế inducer, trí tuệ nhân tạo nhìn vào nhiều thông tin. Nó nhìn phân bố áp suất ở phía trước cánh. Nó xem áp suất hạ mạnh ở đâu và cavitation tại đó sẽ lớn lên đến mức nào. Nó cũng xem áp suất phục hồi bao nhiêu ở phía sau cánh. Nó gom các thông tin này lại để tìm hình dạng cánh tốt. Hình dạng cánh quét về phía trước giúp trì hoãn cavitation. Trí tuệ nhân tạo tạo ra nhiều hình dạng như vậy và so sánh nhanh.

Trí tuệ nhân tạo cũng được dùng để chẩn đoán cavitation. Một nghiên cứu năm 2025 đã dùng camera tốc độ cao để quay cavitation của inducer, rồi phân tích video đó bằng mạng nơ-ron. Nghiên cứu dùng cùng lúc mạng nơ-ron nhìn mẫu không gian và mạng nơ-ron nhìn biến đổi theo thời gian. Cách này có thể nhận ra khoảnh khắc cavitation trở nên bất ổn. Trí tuệ nhân tạo bắt được những thay đổi nhanh mà mắt khó thấy. Nghiên cứu này thuộc hướng chẩn đoán nhận biết cavitation. Công dụng của nó khác với mô hình thay thế dùng để tính toán dòng chảy hoặc thiết kế ngược để tìm hình dạng cánh. Cần tách ba việc này ra để nhìn, thay vì gộp chúng lại.

Thiết kế inducer không kết thúc bằng tính toán dòng chảy. Cánh quay rất nhanh nên chịu lực lớn. Sóng sinh ra khi bọt khí cavitation bị ép xẹp cũng đập vào cánh. Vì vậy, phải xem xét cùng lúc liệu cánh có chịu được các lực này hay không. Nếu thiết kế đường dẫn làm mát là sự cân bằng giữa nhiệt và dòng

chảy, thì thiết kế inducer là sự cân bằng giữa dòng chảy, lực và cavitation. Trí tuệ nhân tạo được dùng để cân nhắc nhiều mục tiêu này cùng lúc.

Ở đây, giới hạn cũng rõ ràng. Thử nghiệm hydro cực lạnh khó và tốn kém. Không có nhiều thiết bị thử nghiệm có thể xử lý chất lỏng lạnh hơn âm 250 độ. Vì vậy, dữ liệu kiểm chứng cho thiết kế inducer không dồi dào. Các con số giảm cavitation chỉ đáng tin khi được thử nghiệm bơm thực tế chống đỡ. Vì vậy, cuốn sách này không dùng các con số chưa được xác nhận. Thay vào đó, sách tập trung vào việc mô hình có bổ sung vật lý nhìn vào thông tin nào để giúp thiết kế. Dù trí tuệ nhân tạo thu hẹp được các ứng viên cánh tốt, phán đoán cuối cùng vẫn phải dựa trên thử nghiệm bơm thực tế.

## Tổng kết

Chương này đã xem xét cách thiết kế cấu trúc quản lý nhiệt của động cơ tên lửa bằng trí tuệ nhân tạo. Có ba cấu trúc được bàn đến. Đường dẫn làm mát trong thành, lỗ làm mát màng phủ một lớp màng làm mát lên thành, và inducer cánh trước bơm đẩy nhiên liệu lên. Cả ba cấu trúc đều mang cùng một kiểu bài toán. Đó là bài toán cân bằng giữa nhiệt và dòng chảy. Khi đạt mạnh một phía, ta mất phía kia. Trí tuệ nhân tạo đã đi vào việc nhanh chóng tìm điểm cân bằng này.

Phía nền có hai thay đổi lớn. Một là in 3D kim loại. Nó giúp tạo ra các đường dẫn bên trong phức tạp mà gia công cắt gọt không làm được. Tự do thiết kế đã mở rộng rất nhiều. Thay đổi kia là vật liệu mới như hợp kim đồng GRCop-42, vốn truyền nhiệt tốt. Khi thiết kế tự do gặp vật liệu tốt, cấu trúc làm mát mới trở thành khả thi. Tuy nhiên, phương thức chế tạo mới nhạy cảm với quản lý chất lượng, nên phải theo dõi kỹ quá trình chế tạo.

Trí tuệ nhân tạo được dùng theo nhiều cách. Mô hình thay thế ước lượng nhanh hành vi của nhiên liệu siêu tới hạn thay cho tính toán chất lưu chậm. Học chủ động Gaussian process tiết kiệm các phép tính đắt đỏ trong khi tìm lỗ làm mát màng tốt. Autoencoder tích chập thu nhỏ hình dạng đường dẫn phức tạp thành một không gian số nhỏ, làm cho thiết kế topo dễ hơn. Tối ưu hóa Bayes đa mục tiêu tìm điểm cân bằng giữa các mục tiêu va chạm nhau để tìm hãm coking. Học máy có bổ sung vật lý và toán tử nơ-ron Fourier giúp thiết kế kiểm soát cavitation của inducer.

Các phương pháp trong chương này nối với nhau. Mô hình thay thế ước lượng nhanh dòng chảy. Học chủ động chọn phép tính tiếp theo dựa trên ước lượng đó. Autoencoder thu nhỏ không gian thiết kế để việc tìm kiếm dễ hơn. Tối ưu hóa Bayes tìm điểm cân bằng của nhiều mục tiêu. Học có bổ sung vật lý buộc mô hình giữ các định luật vật lý. Không phải chỉ dùng một công cụ, mà là nối nhiều công cụ lại để dùng. Việc chọn và kết hợp cho đúng bài toán là phần việc của người thiết kế.

Có một nguyên tắc xuyên suốt. Trí tuệ nhân tạo không phải là công cụ quyết định đáp án thay con người. Nó là công cụ nhanh chóng thu hẹp các ứng viên tốt trong không gian thiết kế rộng. Mô hình thay thế chỉ đáng tin trong phạm vi nó đã học. Các con số hiệu năng trong ví dụ của doanh nghiệp khác với kiểm chứng học thuật. Không dùng những con số chưa xác nhận được nguồn. Làm mát động cơ tên lửa là lĩnh vực mà chỉ cần thành nóng chảy một lần, động cơ sẽ vỡ. Vì vậy, phán đoán cuối cùng luôn dựa trên tính toán độ chính xác cao và thử nghiệm đốt thực tế.

Các phương pháp trong chương này có điểm chung. Tất cả đều là kỹ thuật nhằm tiết kiệm tính toán đắt đỏ. Tính toán chính xác thì đúng, nhưng chậm. Nếu chạy phép tính chậm đó mỗi lần, thiết kế sẽ không thể kết thúc. Vì vậy, trí tuệ nhân tạo nhận phần ước lượng nhanh. Mô hình thay thế làm thay tính toán, học chủ động chọn thứ tự tính toán, còn autoencoder thu nhỏ bài toán. Thời gian tiết kiệm được dùng để xem xét nhiều ứng viên hơn. Càng xem xét nhiều, xác suất tiến gần đến đáp án tốt càng tăng.

Hướng đi tiếp theo cũng nằm trên nguyên tắc này. Học máy đưa tri thức vật lý vào sâu hơn đang phát triển. Dòng chảy nối thiết kế với chẩn đoán cũng đang mở rộng. Trí tuệ nhân tạo đưa ra các phương án

thiết kế, thử nghiệm kiểm chứng các phương án đó, và dữ liệu thử nghiệm ấy lại nuôi lớn trí tuệ nhân tạo. Vòng lặp này đang từng bước đẩy thiết kế quản lý nhiệt của động cơ tên lửa tiến lên. Nếu rút gọn chương này thành một câu thì như sau. Trên nền tự do thiết kế đã mở rộng, trí tuệ nhân tạo giúp giải nhanh bài toán cân bằng khó giữa nhiệt và dòng chảy. Và đáp án ấy luôn được xác nhận bằng thử nghiệm. Công cụ đã nhanh hơn, nhưng nguyên tắc kiểm chứng vẫn giữ nguyên. Chương sau sẽ bàn về việc khi vận hành thực tế động cơ được tạo ra như vậy, trí tuệ nhân tạo chẩn đoán và điều khiển động cơ như thế nào.

## Tài liệu căn cứ liên quan

- 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., & Oswald, 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., & Oswald, 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., & Oswald, 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>

# Chương 10. Chẩn đoán lỗi thời gian thực, bản sao số và điều khiển thích ứng thời gian thực cho lực đẩy dựa trên học tăng cường trong động cơ tên lửa nhiên liệu lỏng (LRE)

## Mở đầu

Khi lái ô tô lâu năm, người lái nhận biết tình trạng xe qua âm thanh. Nếu động cơ phát ra tiếng lạ, họ đưa xe đến xưởng sửa chữa. Nếu đèn cảnh báo trên bảng đồng hồ sáng lên, họ giảm tốc độ. Con người phán đoán bằng cách nhìn nhiều tín hiệu cùng lúc như vậy. Động cơ tên lửa cũng tương tự. Chỉ có điều động cơ tên lửa nhanh hơn, nóng hơn và nguy hiểm hơn ô tô rất nhiều.

Động cơ tên lửa nhiên liệu lỏng tạo ra lực đẩy lớn bằng cách đốt nhiên liệu lỏng và chất oxy hóa lỏng. Động cơ này xử lý nhiệt và áp suất khổng lồ trong thời gian ngắn. Chỉ một bộ phận lệch đi một chút cũng có thể làm toàn bộ hệ thống sụp đổ. Vì vậy, việc liên tục theo dõi tình trạng trong khi động cơ đang hoạt động rất quan trọng. Việc quan sát tình trạng như vậy trong suốt thời gian động cơ chạy được gọi là giám sát trạng thái thời gian thực.

Chương này bàn về cách trí tuệ nhân tạo quan sát và điều chỉnh động cơ tên lửa. Trí tuệ nhân tạo viết bằng tiếng Anh là Artificial Intelligence, viết tắt là AI. Nội dung chính gồm bốn câu chuyện. Một là chẩn đoán để nhanh chóng tìm ra bất thường trong tín hiệu cảm biến. Tiếp theo là bản sao số (Digital Twin), tức là vẽ lại phần bên trong động cơ không nhìn thấy được bằng máy tính. Một nội dung nữa là cách đo sự bất định ngay cả khi động cơ bị lão hóa hoặc có sai số chế tạo. Cuối cùng là điều khiển thông minh, điều chỉnh nhiều van cùng lúc.

Khi các công nghệ này gắn lại thành một hệ thống, động cơ tự chăm sóc sức khỏe của mình. Dòng công nghệ này là điều bắt buộc trong thời đại tên lửa tái sử dụng. Khi dùng một động cơ nhiều lần, trạng thái của nó thay đổi từng chút một sau mỗi lần dùng. Con người khó có thể tự tay hiệu chỉnh lại mọi thứ mỗi lần. Trí tuệ nhân tạo là công cụ đọc nhanh những thay đổi đó và phản ứng với chúng. Từ đây, ta sẽ lần lượt tháo gỡ nguyên lý vận hành của các công nghệ này.

Trước khi đọc chương này, cần giữ một thái độ rõ ràng. Trí tuệ nhân tạo không phải phép thuật. Nó là công cụ nhìn dữ liệu và tìm quy luật. Nếu dữ liệu xấu, câu trả lời cũng xấu. Vì vậy, chương này nêu cả công dụng lẫn giới hạn của từng công nghệ. Tôi không dùng các con số hiệu năng chưa được kiểm chứng. Thay vào đó, tôi tập trung giải thích nguyên lý cho dễ hiểu. Mục tiêu là giúp người đọc hiểu bộ khung của công nghệ.

## 10.1 Những thách thức của giám sát trạng thái thời gian thực trong LRE

### Nội dung của mục này

Mục này giải thích vì sao việc quan sát động cơ tên lửa theo thời gian thực lại khó đến vậy. Trước hết, ta hình dung môi trường khắc nghiệt của động cơ tên lửa nhiên liệu lỏng. Sau đó, ta xem các giới hạn của phương pháp giám sát đã được dùng cho đến nay. Cuối cùng, ta nói vì sao cần giám sát bằng trí tuệ nhân tạo.

Trước hết, cần giải thích một thuật ngữ. LRE, thuật ngữ xuất hiện nhiều trong chương này, là viết tắt của động cơ tên lửa nhiên liệu lỏng (Liquid Rocket Engine). Từ đây trở đi, tôi sẽ viết là động cơ tên lửa nhiên liệu lỏng hoặc động cơ. Loại động cơ này được dùng rộng rãi khi đưa vật thể nặng lên không gian.

Động cơ tên lửa nhiên liệu lỏng là một cỗ máy phức tạp, trong đó nhiều bộ phận cùng chuyển động. Buồng đốt là căn phòng nơi nhiên liệu và chất oxy hóa được đốt cháy. Kim phun là bộ phận phun nhỏ nhiên liệu và chất oxy hóa để trộn chúng với nhau. Bơm tua-bin là bơm đẩy chất đẩy vào hệ thống ở áp suất cao. Van đóng mở dòng chảy của chất đẩy. Kênh làm mát lấy nhiệt ra để thành động cơ không bị nóng chảy. Các bộ phận này phải chuyển động như một cơ thể thì động cơ mới chạy đúng.

Vấn đề nằm ở chỗ môi trường này rất khắc nghiệt. Áp suất bên trong buồng đốt lên tới hàng chục atm. Nhiệt độ ngọn lửa vượt quá 3 nghìn độ. Bơm tua-bin quay hàng chục nghìn vòng mỗi phút. Những giá trị này khác hẳn nôi áp suất trong bếp hay động cơ ô tô. Một dao động nhỏ cũng lan thành lực lớn. Vì vậy, động cơ luôn vận hành trong nguy cơ hư hỏng.

Môi trường khắc nghiệt này cũng tàn nhẫn với cảm biến. Cảm biến là đôi mắt đọc trạng thái. Nhưng nếu cảm biến ở gần nơi nóng, bản thân nó sẽ hỏng. Nếu rung động mạnh, dây dẫn cũng có thể đứt. Nghĩa là phải nhìn động cơ bằng đôi mắt đã mờ. Vì vậy, khó có thể tin vào một giá trị cảm biến duy nhất. Phải nhìn nhiều cảm biến cùng lúc và lọc tín hiệu. Ngay cả khi một cảm biến bất thường, ta kiểm tra lại bằng cảm biến khác.

Lượng tín hiệu cũng không hề nhỏ. Một cảm biến đọc dày đặc tạo ra hàng chục nghìn giá trị trong 1 giây. Trên động cơ có từ hàng chục đến hàng trăm cảm biến như vậy. Khi đó, hàng triệu giá trị đổ về trong 1 giây. Con người không thể nhìn bằng mắt hết lượng giá trị này. Hơn nữa, phán đoán phải kết thúc trong mili giây. Mili giây là 1/1000 giây. Trong thời gian chớp mắt một lần, phải phán đoán hàng trăm lần. Tốc độ và khối lượng này không thể xử lý nếu thiếu sự trợ giúp của máy móc.

Ta lấy ví dụ về việc một bất thường của động cơ lan thành tai nạn nhanh đến mức nào. Áp suất bên trong buồng đốt bắt đầu dao động lớn theo nhịp đều. Dao động này có thể tăng lên gấp nhiều lần chỉ trong vài mili giây. Dao động đã lớn sẽ gõ vào kim phun và thành buồng. Khi việc gõ này lặp lại, kim loại nứt ra. Khí nóng rò qua khe nứt. Quá trình này xảy ra trong nháy mắt. Vì vậy, phải bắt được dấu hiệu đầu tiên của bất thường càng sớm càng tốt.

Một điểm khó nữa là thử nghiệm trên mặt đất và bay có điều kiện khác nhau. Trong thử nghiệm trên mặt đất, động cơ được cố định chắc chắn. Có thể gắn nhiều cảm biến. Nếu có vấn đề, có thể tắt ngay. Khi bay thì khác. Vì trọng lượng, không thể gắn nhiều cảm biến. Con người cũng không thể đứng cạnh để quan sát. Vì vậy, động cơ đang bay phải tự phán đoán chỉ bằng số cảm biến ít ỏi, và trí tuệ nhân tạo hỗ trợ phán đoán đó.

Chẩn đoán động cơ tên lửa có mục tiêu giống chẩn đoán các máy móc khác. Nhưng điều kiện khắc nghiệt hơn rất nhiều. Thời gian ngắn, dữ liệu ít và rủi ro lớn. Vì vậy, khó có thể bê nguyên phương pháp của lĩnh vực khác sang. Phải chỉnh lại phương pháp cho phù hợp với tên lửa. Các công nghệ mà chương này giới thiệu đều là kết quả của quá trình chỉnh sửa như vậy, và công việc đó vẫn tiếp tục đến nay.

Thử nghiệm trên mặt đất cũng nguy hiểm như vậy. Thử nghiệm cố định động cơ xuống đất rồi đốt lửa được gọi là thử nghiệm hot-fire. Trong thử nghiệm này, động cơ tạo ra lực giống thực tế. Trong không gian, tình hình còn khó hơn. Để tên lửa hạ cánh trở lại, động cơ được tắt rồi bật lại trên không. Khi đó, chất đẩy trong bồn chứa bị sóng sánh. Dao động sóng sánh này được gọi là sloshing. Thời điểm mở van cũng lệch nhau từng chút một. Những dao động gấp như vậy xuất hiện rồi biến mất trong chớp mắt, nên rất dễ bị bỏ lỡ.

Nếu tưởng tượng cảnh tên lửa hạ cánh trở lại, ta sẽ dễ hiểu hơn. Khi phóng, động cơ tạo ra lực lớn nhất. Khi hạ cánh, nó phải giảm lực thật nhiều để hạ xuống nhẹ nhàng. Trong quá trình này, lực của động cơ tăng giảm trong một dải rộng. Chất lỏng trong bồn chất đẩy cũng sóng sánh theo. Van đóng mở nhanh. Nhiệt độ và áp suất thay đổi gấp. Bất thường dễ xuất hiện trong khoảnh khắc ngắn và khắc nghiệt này. Vì vậy, phải theo dõi khoảnh khắc đó mà không bỏ lỡ.

Việc tái khởi động động cơ trên không cũng khó. Ngay cả trên quỹ đạo, trọng lực Trái Đất vẫn lớn. Chỉ có điều tên lửa và chất đẩy cùng rơi tự do, nên gần như không cảm thấy trọng lượng. Trạng thái này gọi là vi trọng lực. Trong vi trọng lực, chất lỏng không nằm yên ở một vị trí trong bồn. Vì vậy, phải dùng gia tốc nhỏ hoặc sức căng bề mặt để gom chất đẩy về một phía. Nếu bật động cơ khi chưa có quá trình này, dòng chảy không đều. Áp suất và lưu lượng nhảy vọt trong khoảnh khắc. Chẩn đoán dễ bị dao động trong trạng thái quá độ như vậy. Có thể nhìn nhầm trạng thái bình thường thành bất thường. Vì vậy, hệ chẩn đoán phải học cách chịu được trạng thái quá độ này.

Cho đến nay, giám sát động cơ thường dùng phương pháp giá trị giới hạn. Nếu một giá trị vượt qua đường đã định trước, động cơ sẽ bị tắt ngay. Đường này được gọi là redline. Ví dụ, nếu áp suất vượt giá trị đã định, động cơ dừng lại. Phương pháp này dễ hiểu và nhanh, nên từ lâu đã được dùng như một phương pháp đáng tin cậy.

Phương pháp redline giống cầu dao điện trong nhà. Nếu dùng quá nhiều điện, cầu dao sập xuống. Khi vượt giới hạn, nó cắt không điều kiện. Cách này tốt để ngăn tai nạn lớn. Phán đoán gọn nên nhanh và chắc. Động cơ tên lửa cũng đặt cơ chế cắt an toàn như vậy làm nền tảng. Nếu giá trị vượt đường nguy hiểm, động cơ lập tức tắt.

Tuy nhiên, phương pháp redline dễ bỏ lỡ tín hiệu bất thường yếu. Tín hiệu cảm biến luôn lẫn nhiễu. Nếu nhiễu lớn, cảnh báo sẽ vang lên dù mọi thứ vẫn bình thường. Đây gọi là báo động giả. Nếu báo động giả xảy ra thường xuyên, con người sẽ không còn tin cảnh báo. Ngược lại, dao động nhỏ của kim phun hoặc mài mòn nhỏ ở phốt lại ẩn bên trong giá trị giới hạn. Những tín hiệu yếu như vậy khó thấy cho đến khi chúng lớn lên, và khi đã lớn thì thường đã quá muộn để can thiệp.

Chẩn đoán bằng trí tuệ nhân tạo cố lấp khoảng trống này. Con người chỉ có thể nhìn một hoặc hai đồng hồ đo. Trí tuệ nhân tạo nhìn cùng lúc hàng trăm tín hiệu cảm biến. Nó xem xét cả hình dạng và quan hệ giữa nhiều tín hiệu. Vì vậy, nó có thể đọc dấu hiệu bất thường trước khi giá trị vượt đường giới hạn. Một bài tổng quan đăng trên tạp chí hàng không vũ trụ năm 2026 đã hệ thống hóa xu hướng này. Bài báo giải thích rằng các phương pháp dựa trên dữ liệu trong chẩn đoán động cơ tên lửa nhiên liệu lỏng đã tăng mạnh. Đồng thời, bài báo cũng chỉ ra khó khăn là dữ liệu hỏng hóc thực tế còn ít.

Có thể gói lại rằng giám sát động cơ tên lửa có ba bức tường. Tín hiệu nhanh và nhiễu. Bất thường yếu bộc lộ muộn. Cũng phải giữ cân bằng giữa báo động giả và bỏ sót. Trí tuệ nhân tạo là nỗ lực vượt qua cả ba bức tường đó cùng lúc. Từ mục sau, ta sẽ lần lượt xem cụ thể từng phương pháp.

Giám sát bằng trí tuệ nhân tạo không loại bỏ phương pháp redline. Hai phương pháp được dùng cùng nhau. Redline là tuyến phòng thủ cuối cùng để ngăn tai nạn lớn. Trí tuệ nhân tạo đọc trước các dấu hiệu yếu ở phía trước tuyến đó. Có thể xem đây là lưới an toàn xếp thành nhiều lớp. Lớp lưới bên trong bắt những gì lớp lưới bên ngoài bỏ sót. An toàn tên lửa được bảo vệ bằng cách chồng nhiều phương pháp lên nhau như vậy, chứ không dựa vào chỉ một phương pháp.

## 10.2 Nhìn nhiều cảm biến cùng lúc để tìm bất thường sớm

### Nội dung của mục này

Mục này bàn về cấu trúc trí tuệ nhân tạo phát hiện nhanh bất thường trong tín hiệu cảm biến của động cơ. Trước hết, ta giải thích thuật ngữ telemetry và tín hiệu tần số siêu cao. Sau đó, ta giới thiệu mạng nơ-ron nhìn đồng thời dòng thời gian và quan hệ giữa các kênh. Mạng nơ-ron như vậy được gọi là mạng chú ý tích chập thời gian (TCAN). Cấu trúc này không phải mô hình đã được kiểm chứng riêng cho tên lửa, mà là phương pháp dùng trong nhiều lĩnh vực. Cuối cùng, ta cũng xem các giới hạn mà kiểu chẩn đoán này vẫn chưa vượt qua.

Trước hết, cần giải thích thuật ngữ. Telemetry là việc gửi và nhận tín hiệu về trạng thái của một thiết bị ở xa. Trên động cơ có các cảm biến đo áp suất, rung động, nhiệt độ và lưu lượng. Các giá trị cảm biến này đi vào máy tính theo thời gian thực. Tần số siêu cao có nghĩa là đọc các tín hiệu này với khoảng cách rất dày. Một số cảm biến đọc hàng chục nghìn đến hàng trăm nghìn lần trong 1 giây. Phải đọc dày như vậy mới bắt được dao động nhanh.

Cảm biến đo áp suất dùng phần tử áp điện. Phần tử áp điện là vật liệu tạo ra một lượng điện nhỏ khi bị nén. Khi áp suất thay đổi nhanh, tín hiệu điện cũng thay đổi nhanh. Vì vậy, cảm biến này bắt tốt dao động áp suất nhanh. Gia tốc kế đo mức độ rung của bộ phận. Khi dùng hai loại cảm biến này cùng nhau, ta nhìn đồng thời áp suất và rung động. Nghĩa là ta đọc trạng thái động cơ từ nhiều góc độ.

Giám sát đọc chậm sẽ bỏ lỡ bất thường nhanh. Hãy tưởng tượng chụp cánh chim ruồi bằng một chiếc máy ảnh cũ. Nếu màn trập chậm, cánh sẽ nhòe thành một mảng. Chỉ màn trập nhanh mới ghi được khoảnh khắc của cánh. Cảm biến cũng vậy. Nếu khoảng đọc thưa, dao động nhanh biến mất. Vì vậy, muốn thấy dao động tần số cao nguy hiểm, phải đọc thật dày.

Hãy xem vì sao cần giám sát nhanh như vậy. Bất thường của động cơ tên lửa lớn lên trong chớp mắt. Có hiện tượng áp suất gần kim phun dao động đều. Hiện tượng này được gọi là chugging. Đó là hiện tượng động cơ lắc lư ở tần số thấp. Nếu nặng, van và đường ống rung lên rồi hỏng. Dao động như vậy lớn lên rất nhanh. Con người không thể nhìn bằng mắt rồi phản ứng kịp. Vì vậy, máy móc phải thay con người phán đoán theo đơn vị mili giây.

Bất thường của một bộ phận đôi khi xuất hiện trước dưới dạng tín hiệu của bộ phận khác. Ví dụ, trục của bơm tua-bin rung trước. Một lát sau, rung động đó dẫn tới thay đổi áp suất buồng đốt. Con người khó nhìn ra ngay độ trễ thời gian và mối liên hệ này. Vì mạng nơ-ron nhìn nhiều cảm biến được gom lại một chỗ, nó có thể học độ trễ thời gian và quan hệ giữa các kênh như vậy.

Cấu trúc được giới thiệu ở đây kết hợp hai phương pháp. Một là tích chập thời gian. Tích chập thời gian là phương pháp rút ra hình dạng từ dòng thời gian của tín hiệu. Nó nhìn dòng chảy dài bằng cách mở rộng độ rộng từ cửa sổ ngắn đến cửa sổ dài. Phương pháp còn lại là attention. Attention là cách quyết định nên tập trung hơn vào đâu trong nhiều tín hiệu. Nó đặt trọng số lớn hơn cho kênh đang quan trọng. Khi kết hợp hai phương pháp này, ta xem xét đồng thời dao động ngắn và quan hệ giữa các kênh. Cấu trúc kết hợp này được gọi là mạng chú ý tích chập thời gian (TCAN).

Ta hãy giải thích kỹ hơn cách tích chập theo thời gian nhìn thấy dòng diễn biến dài. Tích chập là phương pháp quét tín hiệu bằng một cửa sổ nhỏ. Nếu cửa sổ hẹp, nó chỉ nhìn thấy một khoảnh khắc ngắn. Bất thường trong tên lửa đôi khi xuất hiện qua một khoảng thời gian dài. Vì vậy phải nhìn trên một khoảng thời gian rộng. Phương pháp dùng lúc này là mở rộng khoảng cách của cửa sổ. Cách này gọi là tích chập giãn. Nó bỏ qua từng giá trị một rồi quét một khoảng thời gian rộng. Giống như khi bước lên cầu thang cách hai bậc một, với cùng số bước ta lên được cao hơn. Làm như vậy có thể nhìn thấy dòng diễn biến dài với ít tính toán hơn.

Cũng có thể lấy ví dụ về cách attention phân chia sự tập trung. Hãy tưởng tượng cảnh ta chọn nghe giọng của bạn mình trong một lớp học ồn ào. Tai tập trung vào âm thanh quan trọng trong nhiều âm thanh. Attention cũng làm việc tương tự. Nó đặt trọng số vào kênh đang làm lộ rõ bất thường hiện tại. Khi tình huống thay đổi, trọng số cũng thay đổi. Vì vậy trong nhiều cảm biến, nó chọn tín hiệu quan trọng theo từng thời điểm. Tích chập theo thời gian nhìn vào dòng thời gian, còn attention chọn kênh.

Cách kết hợp này được dùng rộng rãi trong các nghiên cứu gần đây. Một nghiên cứu năm 2025 đề xuất mạng nơ-ron tích chập theo thời gian có thêm attention. Nghiên cứu này giải thích rằng nó bắt được cả phụ thuộc ngắn hạn và phụ thuộc dài hạn. Một nghiên cứu khác cùng năm kết hợp tích chập theo thời gian với graph attention. Nghiên cứu này xem nhiều cảm biến như một đồ thị và học cả quan hệ giữa

chúng. Những nghiên cứu này không dành riêng cho tên lửa. Nhưng chúng phù hợp với bài toán cùng lúc tìm bất thường trong nhiều cảm biến. Chẩn đoán động cơ tên lửa cũng là cùng loại bài toán, vì phải cùng lúc tìm bất thường trong nhiều cảm biến. Điều đã được kiểm chứng là cấu trúc này tìm bất thường tốt trong dữ liệu chuỗi thời gian nói chung. Hiện vẫn chưa có chứng cứ thực nghiệm rằng nó áp dụng nguyên vẹn cho cảm biến tên lửa và các kiểu hỏng hóc của tên lửa. Cuốn sách này đề xuất chuyển cấu trúc này sang hướng chẩn đoán tên lửa.

Các nghiên cứu khác trong lĩnh vực tên lửa cũng cho thấy hướng tương tự. Chẩn đoán bằng deep learning đọc đồng thời nhiều tín hiệu cảm biến đang tăng lên. Một nghiên cứu năm 2022 dùng mạng nhớ dài ngắn hạn có thể diễn giải. Mạng nơ-ron này phân loại bất thường của động cơ bằng dữ liệu từ nhiều cảm biến. Một nghiên cứu năm 2024 bàn về phát hiện lỗi dựa trên học đặc tính. Tất cả dòng nghiên cứu này đều là nỗ lực kết hợp nhiều tín hiệu để phát hiện bất thường sớm.

Việc phép chẩn đoán này được tính toán ở đâu cũng quan trọng. Nếu gửi tín hiệu đến một máy tính ở xa để phán đoán, sẽ mất thời gian. Muốn đưa ra câu trả lời trong mili giây, phải tính toán ngay bên cạnh động cơ. Cách tính trực tiếp bên cạnh thiết bị như vậy gọi là điện toán biên. Người ta đưa mạng nơ-ron vào một chip chuyên dụng nhỏ và cho nó chạy. Ngay khi nhận tín hiệu, nó phán đoán tại chỗ. Có như vậy mới nhận ra nguy hiểm trong mili giây. Vì vậy mạng nơ-ron chẩn đoán phải nhỏ và nhanh. Mạng nơ-ron lớn dù chính xác nhưng nếu chậm thì giá trị sử dụng giảm.

Mô hình chẩn đoán thường đi qua ba bước. Trước hết, nó chỉnh tín hiệu cảm biến cho đồng đều. Việc này gọi là chuẩn hóa. Lý do là mỗi cảm biến có kích thước và đơn vị khác nhau. Tiếp theo, nó cắt ra một cửa sổ thời gian ngắn và đưa vào mô hình. Cuối cùng, nó phân loại trạng thái thành nhiều kiểu. Nó phân biệt là bình thường, bất thường cảm biến, bất thường bơm, hay dao động cháy. Mô hình chẩn đoán lặp lại ba bước này liên tục với tốc độ rất cao.

Có nhiều loại bất thường mà phép chẩn đoán này muốn bắt. Có chugging, hiện tượng lắc lư ở tần số thấp gần kim phun. Có dao động tần số cao, phát ra như tiếng rít với rung động nhanh hơn. Có vấn đề phốt của turbopump bị mòn và rò rỉ. Cũng có vấn đề ổ trục bị mòn làm rung động tăng lên. Mỗi bất thường để lại một dạng tín hiệu khác nhau. Mô hình chẩn đoán học sự khác nhau giữa các dạng này. Nó nhìn xem dạng nào xuất hiện rồi phân biệt loại bất thường.

Ở đây có một điểm cần cẩn trọng. Một số bài viết bàn về cấu trúc này nêu các con số như độ chính xác hoặc độ trễ. Nhưng các con số đó không được xác nhận bằng bài báo công khai. Nếu viết những con số chưa được xác nhận như sự thật, người đọc sẽ bị dẫn sai. Vì vậy cuốn sách này không dùng những con số đó. Thay vào đó, cuốn sách chỉ giải thích cấu trúc này như nguyên lý chẩn đoán lỗi sớm. Trọng tâm là cùng nhìn nhiều cảm biến để bắt bất thường từ sớm.

Trong động cơ thực tế, khi xử lý kết quả chẩn đoán, người ta chia mức độ nguy hiểm của hỏng hóc. Một tín hiệu chí mạng, như áp suất vượt xa ngưỡng nguy hiểm, chỉ cần xuất hiện một mình cũng lập tức tắt động cơ. Nếu bắt buộc xác nhận nhiều lớp đối với tín hiệu như vậy, việc dừng sẽ chậm lại và tai nạn sẽ lớn hơn. Ngược lại, với dấu hiệu yếu, chỉ hành động khi nhiều tín hiệu cùng chỉ về một bất thường. Lý do là để không phá hỏng một thử nghiệm đang bình thường chỉ vì một phán đoán sai của một mô hình. Vì vậy đầu ra chẩn đoán được dùng cùng logic an toàn. Phải có giám sát của con người và thiết bị an toàn dựa trên quy tắc đi kèm.

Muốn tin mô hình chẩn đoán, ta phải biết vì sao nó phán đoán như vậy. Mạng nơ-ron giống như một chiếc hộp khó nhìn vào bên trong. Nó đưa ra đáp án nhưng không nói lý do. Ở nơi nguy hiểm như tên lửa, lý do rất quan trọng. Vì vậy các nghiên cứu cho thấy cảm biến nào dẫn dắt phán đoán đang tăng lên. Cách này gọi là chẩn đoán có thể diễn giải. Con người nhìn lý do đó rồi quyết định có tin chẩn đoán hay không. Phải biết được lý do thì mới có thể dùng trí tuệ nhân tạo trong phán đoán an toàn.

Một khó khăn khác là dữ liệu ít. Dữ liệu hồng học thật của động cơ tên lửa rất hiếm. Lý do là hồng học đồng nghĩa với một tai nạn đắt đỏ. Vì vậy dữ liệu bình thường thì nhiều, còn dữ liệu hồng học thì gần như không có. Khi dữ liệu ít, trí tuệ nhân tạo khó học tốt. Để giải quyết việc này, các nhà nghiên cứu dùng nhiều phương pháp. Một cách là tạo dữ liệu giả mô phỏng tình huống hồng học bằng mô hình vật lý. Cách khác là cố ý đưa một số hồng học đã định trước vào mô hình thử nghiệm. Quy trình cố ý đưa hồng học vào như vậy gọi là tiêm lỗi. Tạo dữ liệu giả và tiêm lỗi có nghĩa khác nhau. Thiếu dữ liệu vẫn là bài toán lớn của lĩnh vực này.

Cũng có phương pháp chuyển điều đã học từ một động cơ sang động cơ khác. Cách này gọi là học chuyển giao. Nó giống như khi biết đi xe đạp thì học đi xe máy nhanh hơn. Ta chuyển kiến thức đã học sang tình huống mới. Nhờ vậy, dù dữ liệu của động cơ mới ít, mô hình vẫn học nhanh. Một bài tổng quan trên tạp chí hàng không vũ trụ năm 2026 cũng chỉ ra khó khăn này. Đó là dữ liệu hồng học thực tế ít và loại hồng học thì nhiều. Bài báo này sắp xếp nghiên cứu chẩn đoán xoay quanh ba vấn đề thực tế này.

Mục tiêu của chẩn đoán không chỉ là ngăn tai nạn. Nó còn được dùng để bảo dưỡng đúng mức. Nếu không có bất thường, không cần thay linh kiện. Nếu thấy dấu hiệu bất thường, người ta sửa trước. Cách bảo dưỡng bằng cách nhìn vào trạng thái như vậy gọi là bảo dưỡng dựa trên tình trạng. Động cơ tái sử dụng tiết kiệm chi phí nhờ cách này. Chẩn đoán là công cụ bảo vệ cả an toàn và chi phí.

### 10.3 Bản sao số vẽ ra bên trong buồng đốt không nhìn thấy

Mục này bàn về dao động nguy hiểm bên trong buồng đốt. Dao động này gọi là bất ổn cháy. Trước hết, ta giải thích âm thanh và ngọn lửa làm tăng lực của nhau như thế nào. Sau đó, ta xem cách khôi phục dòng áp suất không nhìn thấy bằng máy tính. Cách này dùng mạng nơ-ron tuân thủ phương trình vật lý. Cuối cùng, ta chỉ ra giá trị sử dụng và giới hạn của cách này.

Trước hết hãy giải thích thuật ngữ. Bất ổn cháy là hiện tượng áp suất dao động mạnh bên trong buồng đốt. Khi ngọn lửa và âm thanh phối hợp với nhau, dao động này lớn lên. Dao động áp suất nhỏ làm rung sự giải phóng nhiệt của ngọn lửa. Ngọn lửa bị rung lại làm dao động áp suất lớn hơn. Nếu phản hồi này lặp lại, dao động lớn lên như quả cầu tuyết. Nếu nghiêm trọng, thành buồng đốt và kim phun bị phá hỏng. Trong lịch sử tên lửa, hiện tượng này là một vấn đề nan giải lâu đời.

Dao động này đã là vấn đề lớn từ giai đoạn đầu phát triển tên lửa. Ngay cả các động cơ cỡ lớn đưa con người lên Mặt Trăng cũng từng khổ sở vì vấn đề này. Áp suất bên trong buồng đốt cộng hưởng mạnh và phá hỏng động cơ. Các nhà nghiên cứu đã nhiều lần thay đổi hình dạng kim phun để khống chế nó. Điều đó cho thấy dao động cháy khó xử lý đến mức nào. Đến nay, mỗi khi chế tạo động cơ mới, người ta vẫn kiểm tra vấn đề này.

Dao động có nhiều dạng. Hãy nghĩ buồng đốt là một hình trụ. Âm thanh có thể cộng hưởng theo chiều dài của hình trụ. Nó cũng có thể quay theo chu vi rồi cộng hưởng. Nó cũng có thể cộng hưởng ngang qua đường kính. Mỗi dạng có tần số khác nhau. Nếu biết dạng nào đang lớn lên, việc ứng phó sẽ dễ hơn. Vì vậy biết hình dạng và cường độ của dao động là việc quan trọng.

Có một nguyên lý lâu đời giải thích phản hồi này. Nếu nhiệt sinh ra nhiều hơn vào đúng lúc áp suất tăng, dao động sẽ lớn lên. Nguyên lý này gọi là tiêu chuẩn Rayleigh. Nghĩ đến cảnh đẩy xích đu thì dễ hiểu. Nếu đẩy khi xích đu đi về phía trước, nó sẽ lên ngày càng cao. Nếu đẩy ở thời điểm ngược lại, xích đu sẽ dừng. Ngọn lửa và âm thanh cũng theo cơ chế đó mà làm nhau lớn lên hoặc lắng xuống. Thời điểm trao đổi lực quyết định ổn định hay bất ổn.

Phản hồi này cũng giống quan hệ giữa âm thoa và micro. Trong hội trường, nếu đưa micro lại gần loa, sẽ có tiếng rít sắc. Đó là vì âm thanh từ loa đi vào micro rồi lại được khuếch đại. Phản hồi này làm âm thanh lớn lên như quả cầu tuyết. Trong buồng đốt, có thể xem ngọn lửa là loa và âm thanh là micro. Nếu

pha của hai bên khớp nhau, âm thanh sẽ lớn lên đến mức khó kiểm soát. Vì vậy cần thiết kế cắt đứt phản hồi này.

Vấn đề là khó nhìn trực tiếp vào bên trong buồng đốt. Bên trong quá nóng và áp suất quá cao. Không thể đặt cảm biến ở bất cứ đâu. Người ta chỉ gắn được cảm biến áp suất ở vài điểm trên thành. Nhưng dao động nguy hiểm xảy ra trong toàn bộ buồng đốt. Chỉ với vài điểm trên thành thì khó biết toàn cảnh. Vì vậy cần một phương pháp vẽ ra toàn bộ từ ít quan sát.

Ở đây bản sao số xuất hiện. Bản sao số là mô hình máy tính bám sát thiết bị thực. Trong khi động cơ thật chạy, mô hình cũng chạy cùng. Nó nhận giá trị cảm biến rồi liên tục chỉnh mô hình cho khớp với thực tế. Làm như vậy có thể ước đoán cả trạng thái ở nơi không có cảm biến. Tên gọi này xuất phát từ việc mô hình chạy song song với vật thật như một cặp song sinh.

Bản sao số có điểm giống hệ thống định vị xe hơi. Hệ thống định vị nhận tình hình đường thực tế rồi vẽ lại đường đi. Bản sao số cũng nhận tín hiệu của động cơ thật rồi vẽ lại trạng thái. Điểm khác là tốc độ và độ chính xác. Bản sao của tên lửa phải khớp trạng thái sau mỗi mili giây. Có như vậy mới theo được dao động nhanh. Muốn đạt tốc độ này, phải làm cho mô hình nhẹ.

Suy ra bên trong bằng vài cảm biến trên thành là một bài toán giải ngược. Thông thường ta biết nguyên nhân rồi tìm kết quả. Ở đây ta chỉ biết áp suất trên thành, tức kết quả. Từ kết quả đó, ta lần ngược lại trường áp suất bên trong, tức nguyên nhân. Bài toán tìm nguyên nhân từ kết quả như vậy gọi là bài toán ngược. Bài toán ngược khó vì có thể có nhiều đáp án. Phương trình vật lý chỉ giữ lại đáp án phù hợp với vật lý trong số nhiều đáp án đó.

Phương pháp lấp đầy phần bên trong của bản sao số này là mạng nơ-ron dựa trên vật lý. Mạng nơ-ron này gọi là mạng nơ-ron dựa trên vật lý (Physics-Informed Neural Network, PINN). Từ đây sẽ gọi là mạng nơ-ron vật lý. Mạng nơ-ron thông thường chỉ học từ dữ liệu. Mạng nơ-ron vật lý học sao cho vừa theo dữ liệu vừa tuân thủ phương trình vật lý. Khi học, nếu nó vi phạm phương trình thì bị phạt. Vì vậy dù dữ liệu ít, nó vẫn đưa ra đáp án phù hợp với vật lý.

Ta xem kỹ hơn cách mạng nơ-ron vật lý học. Mạng nơ-ron thông thường chỉ giảm sai khác với đáp án đúng. Mạng nơ-ron vật lý giảm đồng thời hai thứ. Một là sai khác với giá trị cảm biến. Hai là mức độ vi phạm phương trình vật lý. Nó học để hạ cả hai khoản phạt cùng lúc. Khi đó, ngay cả ở nơi không có cảm biến, đáp án vẫn tuân thủ vật lý. Có thể nói quy tắc vật lý lấp vào chỗ trống của dữ liệu. Nó giống như đường phác trong tranh tô màu dẫn dắt màu sắc.

Phương pháp này có thể trở thành công cụ giám sát giá rẻ. Muốn tính toán chính xác bên trong buồng đốt, máy tính lớn phải chạy lâu. Trong lúc thử nghiệm thì không có thời gian như vậy. Mạng nơ-ron vật lý, một khi đã học xong, sẽ đưa ra đáp án nhanh. Vì vậy ngay cả trong khi thử nghiệm, nó vẫn có thể vẽ trạng thái bên trong. Bằng bức tranh nhanh này, ta quan sát dấu hiệu của dao động nguy hiểm. Có thể xem như nhìn vào bên trong mà không cần dừng thử nghiệm.

Phương trình mô tả dạng âm thanh trong buồng đốt là phương trình Helmholtz. Phương trình này cho biết âm thanh cộng hưởng theo dạng nào trong một không gian kín. Nó giống cơ chế một số nốt nhất định vang tốt trong nhạc cụ hơi. Mạng nơ-ron vật lý học trong khi tuân thủ phương trình này và các điều kiện biên. Nó nhận đầu vào là giá trị từ một số cảm biến trên thành. Sau đó, nó khôi phục dòng áp suất trong toàn bộ buồng đốt. Nó cho thấy như một bức tranh rằng dao động mạnh lên ở vị trí nào.

Âm thanh trong buồng đốt thay đổi tùy theo cách nó phản ứng ở biên. Mặt kim phun và cổ họng vòi phun phản xạ một phần âm thanh và hấp thụ một phần. Mức độ phản xạ và hấp thụ này gọi là điều kiện biên âm học. Mạng nơ-ron vật lý cũng học trong khi tuân thủ các điều kiện biên này. Có như vậy nó mới vẽ âm thanh của buồng đốt thật, chứ không phải âm thanh của một căn phòng trống. Tiêu chuẩn Rayleigh đã nói ở trên và các điều kiện biên này quyết định bất ổn cháy trong tên lửa.

Hướng này đang được nghiên cứu sôi nổi gần đây. Mạng nơ-ron vật lý có thể mạnh trong việc khôi phục trường áp suất âm thanh từ ít quan sát. Lý do là áp suất chỉ được đo ở vài điểm, còn vận tốc thì hầu như không đo được. Một nghiên cứu năm 2024 học tương tác nhiệt âm trong buồng đốt bằng mạng nơ-ron vật lý. Nhiệt âm là hiện tượng nhiệt và âm thanh đan xen với nhau. Nghiên cứu này tìm các giá trị như suy giảm áp suất và liên kết giải phóng nhiệt từ dữ liệu. Một nghiên cứu năm 2026 ước tính tính chất hấp thụ âm của vách buồng đốt bằng mạng nơ-ron vật lý giải bài toán theo từng miền. Các nghiên cứu này xử lý việc học nhiệt âm của buồng đốt nói chung và ước tính tính chất biên. Hiện vẫn chưa có minh chứng khôi phục toàn bộ trường áp suất theo thời gian thực bằng ít cảm biến trong buồng đốt động cơ tên lửa nhiên liệu lỏng thực tế. Cuốn sách này đề xuất việc khôi phục theo thời gian thực đó như một hướng cần tiến tới.

Mạng nơ-ron vật lý cũng được dùng trong nghiên cứu cháy sạch của lực đẩy hàng không vũ trụ. Một bài báo năm 2026 xem dòng nghiên cứu này là con đường hướng tới lực đẩy bền vững. Nó nối phương trình vật lý với dữ liệu để mô tả quá trình cháy chính xác hơn. Nghiên cứu như vậy không chỉ dừng ở tên lửa mà còn mở rộng sang động cơ hàng không. Ý tưởng về mạng nơ-ron tuân theo vật lý đang lan sang nhiều lĩnh vực.

Bản sao số đã được dùng trước trong bảo trì động cơ hàng không. Một động cơ hàng không cỡ lớn đã dùng bản sao số để phát hiện trước nhiều bất thường. Người ta cho biết nó đã dự đoán mài mòn cánh turbine và vấn đề cháy. Trường hợp này cho thấy mô hình giám sát thực sự có ích. Động cơ tên lửa khắc nghiệt hơn động cơ hàng không và có ít dữ liệu hơn. Vì vậy cần kiểm chứng nhiều hơn để chuyển hướng này sang tên lửa.

Mạng nơ-ron vật lý cũng có điểm yếu. Khi đưa phương trình vào quá trình học, việc huấn luyện trở nên khó hơn. Phải cân bằng đúng trọng số giữa phạt vật lý và phạt dữ liệu. Nếu trọng số lệch, đáp án sẽ nghiêng về một phía. Càng xử lý quá trình cháy phức tạp thì việc điều chỉnh này càng khó. Vì vậy mạng nơ-ron vật lý không phải là công cụ vạn năng. Phải chọn phương trình và cấu trúc phù hợp với bài toán. Kiến thức của con người vẫn cần cho việc lựa chọn này.

Nhìn rộng hơn, học máy dựa trên vật lý đang phát triển nhanh trong toàn bộ nghiên cứu cháy. Một bài tổng quan xuất bản năm 2025 đã hệ thống hóa dòng nghiên cứu này. Bài này giải thích theo từng nhóm như hóa học cháy, dòng phản ứng và nhiều vấn đề khác. Nó cho rằng khi đưa kiến thức vật lý vào quá trình học, đáp án vẫn phù hợp với vật lý dù dữ liệu ít. Một bài tổng quan khác năm 2026 tập trung vào cháy rối trong lực đẩy hàng không vũ trụ. Bài này giải thích rằng nó nối độ nghiêm ngặt của vật lý với sức mạnh của dữ liệu.

Ưu điểm của phương pháp này nằm ở chỗ dùng cả dữ liệu và vật lý. Khi cảm biến ít, rất khó lấp đầy các khoảng trống chỉ bằng dữ liệu. Phương trình vật lý thu hẹp các khoảng trống đó. Vì vậy chỉ với vài cảm biến rẻ tiền, ta vẫn có thể vẽ ra toàn bộ trạng thái. Việc có thể chạy nhanh để giám sát cũng rất hữu ích.

Giới hạn cũng rõ ràng. Phản ứng ngọn lửa và nhiễu loạn rất phức tạp. Phương trình Helmholtz là mô hình rút gọn nhìn âm thanh theo cách giản lược. Nó không chứa được mọi chi tiết của quá trình cháy thực tế. Vì vậy bản sao số này được dùng để giám sát nhanh. Phán đoán tinh vi phải được bổ sung bằng phân tích độ trung thực cao và thử nghiệm thực tế. Các con số chi tiết như tần số cụ thể hay biên độ áp suất chưa được xác nhận bằng bài báo công khai. Vì vậy cuốn sách này không dùng các con số đó. Thay vào đó, sách chỉ giải thích nguyên lý phương pháp này vận hành như thế nào và hữu ích ở đâu.

Để bản sao số chạy tốt, nó phải liên tục khớp với thực tế. Nếu mô hình lệch dần khỏi thực tế, bức tranh sẽ sai. Vì vậy mô hình được hiệu chỉnh liên tục bằng giá trị cảm biến. Nếu hiệu chỉnh chậm, bản sao sẽ bỏ lỡ thực tế. Nếu hiệu chỉnh quá thường xuyên, nó sẽ chạy theo cả nhiễu và dao động. Việc giữ cân bằng giữa hai phía này là điểm then chốt của công nghệ. Một bản sao tốt không bỏ lỡ thực tế nhưng cũng không bị nhiễu chi phối.

Cách xử lý dao động cháy không chỉ là giám sát. Người ta cũng đưa cấu trúc hấp thụ âm vào vách. Cấu trúc này được gọi là thiết bị hấp thụ âm. Người ta cũng thay đổi bố trí vòi phun để cắt vòng phản hồi. Giám sát bằng trí tuệ nhân tạo là công cụ hỗ trợ thiết kế như vậy. Nó báo trước dao động sẽ lớn lên trong điều kiện nào. Thiết kế và giám sát phải đi cùng nhau thì mới kiểm soát được dao động.

## 10.4 Cách đo sự lão hóa và khác biệt cá thể của động cơ

Mục này xử lý vấn đề động cơ lão hóa hoặc khác nhau theo từng cá thể. Trước hết, nó giải thích vì sao động cơ tái sử dụng thay đổi từng chút một. Sau đó, nó xem xét phương pháp đo độ bất định của thay đổi đó. Phương pháp này là mạng nơ-ron đưa ra cả đáp án lẫn độ rộng của đáp án. Cuối cùng, mục này chỉ ra vì sao độ bất định đó quan trọng đối với an toàn. Có một điểm cần nói trước. Các nghiên cứu được nêu ở đây là dự đoán tuổi thọ của máy móc nói chung. Chúng chưa được kiểm chứng bằng dữ liệu động cơ tên lửa, và việc áp dụng cho tên lửa là khả năng trong tương lai.

Động cơ tái sử dụng dùng cùng một động cơ nhiều lần. Có thể hiểu là chuyển từ cách dùng cốc giấy một lần rồi bỏ sang cách dùng cốc sứ nhiều lần. Cách này làm giảm mạnh chi phí vận tải vũ trụ. Nhưng mỗi lần dùng, linh kiện lại mòn đi một chút. Bề mặt vòi phun bị mài mòn. Cánh turbine bị xói mòn. Kênh làm mát thay đổi do môi nhiệt. Hiện tượng hiệu năng thay đổi theo thời gian như vậy được gọi là lão hóa theo năm tháng.

Một biến số khác là sai lệch chế tạo. Dù làm theo cùng một bản vẽ, kích thước thực tế vẫn khác nhau từng chút. Mỗi linh kiện có sai số rất nhỏ. Vì sai số này, phản ứng của từng động cơ khác nhau. Khi lão hóa và sai lệch chế tạo chồng lên nhau, động cơ có cá tính riêng. Dù đưa cùng một lệnh van, kết quả của từng động cơ vẫn hơi khác nhau, nên bộ điều khiển phải biết sự khác biệt này.

Thay đổi này ảnh hưởng trực tiếp đến hiệu năng động cơ. Mài mòn và xói mòn làm lực đẩy giảm dần. Khi lực đẩy giảm, quỹ đạo của tên lửa thay đổi. Mức độ cháy hiệu quả đến đâu cũng thay đổi. Một giá trị thể hiện mức hiệu quả này là vận tốc xả đặc trưng. Khi giá trị này giảm, cùng một lượng chất đẩy tạo ra ít lực hơn. Vì vậy việc đo trước lão hóa và sai lệch rất quan trọng.

Cần giải thích thêm một chút về thuật ngữ vận tốc xả đặc trưng. Giá trị này là chỉ số cho biết buồng đốt cháy hiệu quả đến mức nào. Nó được xác định bằng áp suất buồng đốt, diện tích cổ họng vòi phun và lưu lượng chất đẩy. Nó khác với vận tốc khí xả mà vòi phun thực sự phun ra hoặc hiệu suất vòi phun. Nó chỉ là thước đo mức độ quá trình cháy diễn ra tốt đến đâu. Nếu giá trị này cao, quá trình cháy hiệu quả. Nếu thấp, chất đẩy được đốt kém hiệu quả hơn. Khi vòi phun bị mài mòn, sự trộn kém đi và giá trị này giảm. Vì vậy có thể ước đoán sức khỏe của động cơ qua thay đổi của giá trị này.

Có thể so sánh thay đổi này với đôi giày thể thao đã đi lâu. Giày mới có đế chắc. Đi lâu, đế mòn và trơn. Dù chạy bằng cùng một lực, người chạy vẫn đi được ít hơn. Động cơ cũng giảm hiệu năng từng chút khi dùng nhiều. Vấn đề là mức giảm đó khác nhau ở từng động cơ. Có động cơ mòn nhanh, có động cơ mòn chậm. Vì vậy phải xem riêng trạng thái của từng động cơ.

Hiểu được lão hóa là chìa khóa của tái sử dụng an toàn. Nếu không biết trạng thái, ta mắc hai sai lầm. Ta có thể bỏ một động cơ vẫn còn dùng được quá sớm. Đó là lãng phí. Ngược lại, ta có thể ép dùng một động cơ đã mệt. Điều đó dẫn đến tai nạn. Khi biết chính xác trạng thái, ta dùng động cơ đủ mức cần dùng nhưng không vượt qua ranh giới nguy hiểm, nhờ đó tránh được cả hai sai lầm.

Ở đây mạng nơ-ron Bayes biến phân được dùng. Tiếng Anh là Variational Bayesian Neural Network. Từ đây sẽ gọi là mạng nơ-ron Bayes. Mạng nơ-ron thông thường chỉ đưa ra một đáp án. Ví dụ, nó trả lời dứt khoát rằng lực đẩy giảm 3 phần trăm. Mạng nơ-ron Bayes đưa ra cả phạm vi của đáp án. Nó cho biết cả đáp án và mức độ chắc chắn, chẳng hạn lực đẩy có nhiều khả năng giảm trong khoảng từ 2 đến 4 phần trăm.

Có thể hình dung dễ hơn cách mạng nơ-ron Bayes đưa ra độ rộng. Giả sử ta hỏi nhiều người thay vì chỉ hỏi một người. Nếu nhiều câu trả lời giống nhau, mức chắc chắn cao. Nếu câu trả lời mỗi người một khác, mức chắc chắn thấp. Mạng nơ-ron Bayes chứa nhiều ý kiến như vậy bên trong một mô hình. Vì vậy nó đưa ra cả giá trị trung tâm của đáp án lẫn độ rộng phân tán. Độ rộng hẹp nghĩa là mô hình tự tin, còn độ rộng lớn nghĩa là mô hình vẫn chưa biết rõ.

Có hai loại bất định. Một loại đến từ dao động của chính dữ liệu. Đó là dao động không thể loại bỏ, giống như nhiễu cảm biến. Loại kia đến từ việc mô hình học chưa đủ. Nó lớn lên trong tình huống mô hình chưa từng trải qua. Khi lần đầu gặp một kiểu lão hóa mới, bất định này tăng lên. Nếu tách hai loại ra, việc ứng phó trở nên dễ hơn. Ta có thể phán đoán nên thu thêm dữ liệu hay thử nghiệm thêm.

Mạng nơ-ron Bayes khó học hơn mạng nơ-ron thông thường. Lý do là nó phải học phân bố của đáp án chứ không phải một đáp án. Nếu muốn tìm chính xác phân bố này, lượng tính toán sẽ lớn. Vì vậy các nhà nghiên cứu dùng phương pháp xấp xỉ. Họ học bằng cách đổi phân bố sang một dạng dễ xử lý hơn. Phương pháp này được gọi là suy luận biến phân. Chữ biến phân trong tên gọi đến từ phép xấp xỉ này. Vì dùng xấp xỉ, kết quả nhất thiết phải được kiểm tra bằng dữ liệu thực tế.

Dự đoán lão hóa không phải là vấn đề riêng của tên lửa. Động cơ máy bay, turbine gió và máy móc nhà máy cũng có cùng mối lo. Lý do là tất cả đều là máy móc được dùng lâu dài. Vì vậy nghiên cứu trong lĩnh vực này trao đổi qua lại với nhau. Phương pháp đã thành công ở máy móc khác được chuyển sang tên lửa. Điều kiện khắc nghiệt của tên lửa cũng có khi sinh ra phương pháp mới. Khi các lĩnh vực trao đổi phương pháp như vậy, công nghệ phát triển nhanh hơn.

Mức chắc chắn này được gọi là độ bất định. Biết độ bất định rất quan trọng đối với an toàn. Khi mô hình trả lời tự tin và khi mô hình không biết rõ, ta phải xử lý khác nhau. Nếu mô hình nói rằng nó không biết, con người sẽ thận trọng hơn. Ta đặt biên an toàn lớn hơn. Ta kiểm tra bằng thử nghiệm bổ sung. Nguyên lý này giống như dự báo thời tiết cho biết xác suất mưa. Nếu xác suất là nửa nọ nửa kia, ta mang ô và quan sát tình hình.

Hướng này đang phát triển nhanh trong nhiều lĩnh vực. Nghiên cứu dự đoán tuổi thọ còn lại của linh kiện là ví dụ tiêu biểu. Tuổi thọ còn lại được gọi là tuổi thọ hữu ích còn lại. Một nghiên cứu năm 2024 dự đoán tuổi thọ này bằng mạng nơ-ron Bayes. Nó đưa ra đáp án dưới dạng phân bố xác suất chứ không phải một giá trị. Một nghiên cứu năm 2025 xử lý nhiều bất định cùng lúc bằng mạng nơ-ron hồi quy Bayes. Các nghiên cứu này không dành riêng cho tên lửa. Nhưng bài toán xử lý trạng thái của máy móc đang lão hóa là như nhau. Cũng có thể dùng cùng ý tưởng đó cho dự đoán lão hóa của động cơ tên lửa.

Dữ liệu động cơ hàng không là đối tượng đã được thử nghiệm rộng rãi cho phương pháp này. Các nhà nghiên cứu cạnh tranh phương pháp bằng dữ liệu lão hóa động cơ công khai. Họ so sánh dự đoán của nhiều phương pháp trên cùng một dữ liệu. Làm như vậy có thể phân định phương pháp nào tốt hơn. Dữ liệu tên lửa còn ít, nhưng các dữ liệu công khai như vậy giúp xây nền. Phương pháp đã được xây nền sau đó được chuyển sang dữ liệu tên lửa để kiểm tra.

Bộ điều khiển nhận thông tin bất định này để phán đoán. Nếu độ bất định lớn, nó vận hành thận trọng. Nó đặt lực đẩy thấp hơn một chút hoặc chừa biên rộng hơn. Nếu độ bất định nhỏ, nó khai thác hiệu năng nhiều hơn. Làm như vậy có thể vận hành theo trạng thái của động cơ. Cách này an toàn hơn cách dùng một quy tắc cố định cho mọi động cơ.

Thông tin bất định cũng được dùng trong quyết định bảo trì. Nếu mô hình dự đoán lão hóa một cách tự tin, việc lập kế hoạch trở nên dễ hơn. Nếu mô hình nói rằng nó không biết rõ, ta mở linh kiện ra xem trực tiếp. Làm như vậy giảm bảo trì mù. Vì bảo trì dựa trên trạng thái đã biết, ta tiết kiệm linh kiện. Đồng thời không bỏ sót linh kiện nguy hiểm. Biết độ bất định bảo vệ cả tiết kiệm lẫn an toàn.

Cũng có giới hạn. Trí tuệ nhân tạo đưa ra xác suất không có nghĩa là xác suất đó tự nhiên đúng. Nếu mô hình học sai, nó sẽ đưa ra sự tự tin sai. Vì vậy phải hiệu chỉnh và kiểm chứng bằng dữ liệu thử nghiệm thực tế. Đó là việc kiểm tra xem độ bất định do mô hình đưa ra có khớp với sai số thực tế hay không. Xác suất chưa qua kiểm tra này rất khó tin. Ở nơi nguy hiểm như tên lửa, kiểm chứng này phải nghiêm ngặt hơn. Các con số như khoảng tin cậy chi tiết chưa được xác nhận bằng bài báo công khai. Vì vậy cuốn sách này không dùng các con số đó. Thay vào đó, sách chỉ giải thích khung ý tưởng của việc đo độ bất định.

## 10.5 Điều khiển bằng học tăng cường để điều chỉnh nhiều van cùng lúc (LUMEN, DLR P8.3)

Mục này xử lý điều khiển thông minh điều chỉnh nhiều van của động cơ cùng lúc. Trước hết, nó giải thích học tăng cường là gì. Sau đó, nó xem xét vì sao điều khiển động cơ tên lửa lại khó. Tiếp theo, nó giới thiệu một trường hợp thí nghiệm thực tế của Đức. Cuối cùng, mục này chỉ ra các nhiệm vụ còn lại trước khi dùng công nghệ này trong bay.

Học tăng cường là phương pháp học bằng thử và sai. Nó giống với cảnh một em bé học đi. Có thể nghĩ rằng nếu ngã thì nhận phạt, nếu đi tốt thì nhận thưởng. Hành động tốt được thưởng, hành động xấu bị phạt. Sau nhiều lần thử, hệ thống tìm cách nhận được nhiều thưởng. Phương pháp thêm mạng nơ-ron vào đây được gọi là học tăng cường sâu (Deep Reinforcement Learning, DRL). Từ đây sẽ gọi là học tăng cường sâu.

Trong học tăng cường, hai tâm thế va vào nhau. Một là tâm thế muốn dùng phương pháp tốt mà mình đã biết. Tâm thế kia là muốn thử một phương pháp mới. Nếu lúc nào cũng chỉ dùng điều đã biết, ta không tìm được con đường tốt hơn. Nếu lúc nào cũng chỉ làm cái mới, ta gặp nguy hiểm và dễ lạc hướng. Vì vậy cần có sự cân bằng giữa hai bên. Có thể xem sự cân bằng này như cuộc kéo co giữa tâm thế tìm cái mới và tâm thế dùng điều đã biết. Ở giai đoạn đầu học, hệ thống thử nhiều cách; càng về sau, nó càng dùng điều đã biết.

Trong tên lửa, không thể thực hiện quá trình thử và sai này bằng động cơ thật. Vì không thể học bằng cách làm hỏng động cơ. Vì vậy trước hết ta học bằng mô hình trong máy tính. Mô hình này gọi là trình mô phỏng. Trong trình mô phỏng, dù thất bại hàng nghìn lần cũng không sao. Ta thử thoải mái để tìm cách điều khiển tốt. Điều này giống như chơi nhiều ván trò chơi để học mẹo.

Hãy xem vì sao điều khiển động cơ tên lửa lại khó. Các giá trị bên trong động cơ liên kết chặt với nhau. Khi mở van chất oxy hóa, áp suất tăng. Khi áp suất tăng, tốc độ quay của turbopump thay đổi. Tỷ lệ giữa nhiên liệu và chất oxy hóa cũng thay đổi theo. Tỷ lệ này gọi là tỷ lệ hỗn hợp. Khi động vào một giá trị, các giá trị khác kéo nhau thay đổi. Vì vậy không thể điều chỉnh riêng một van. Cần một cách điều khiển xử lý nhiều giá trị cùng lúc.

Trong tên lửa hạ cánh, việc điều khiển này còn khó hơn. Để tên lửa quay xuống trở lại, phải giảm lực đẩy rất mạnh. Có khi phải hạ xuống dưới một phần ba lực đẩy tối đa. Việc điều chỉnh lực đẩy trong phạm vi rộng như vậy gọi là deep throttling. Nghĩa là siết sâu. Khi đó vẫn phải giữ tỷ lệ hỗn hợp ở mức phù hợp. Nếu tỷ lệ hỗn hợp lệch đi, quá trình cháy xấu đi và linh kiện bị hư hại. Phải điều chỉnh nhiều giá trị cùng lúc, nhanh và an toàn.

Trung tâm Hàng không Vũ trụ Đức nghiên cứu vấn đề này bằng động cơ thật. Trung tâm Hàng không Vũ trụ Đức được viết tắt là DLR. DLR đã chế tạo một động cơ thử nghiệm tên là LUMEN. LUMEN là động cơ dùng bơm, sử dụng oxy lỏng và methane. Lực đẩy thuộc cấp 25 kilonewton. Đây là kích thước tương ứng với động cơ tên lửa cỡ nhỏ. Động cơ này vận hành lần đầu tại bộ thử P8.3 vào tháng 3 năm 2024. P8.3 là cơ sở thử nghiệm tại DLR Lampoldshausen. LUMEN là nơi thử công nghệ động cơ tái sử dụng. Tại đây người ta nghiên cứu đồng thời giám sát trạng thái và điều khiển thông minh.

LUMEN dùng một chu trình cụ thể. Chu trình này gọi là expander bleed. Đây là cách quay turbine bằng một phần nhiên liệu đã được làm nóng trong kênh làm mát. Cấu trúc của cách này tương đối gọn và an toàn. Vì vậy nó phù hợp để thử nhiều công nghệ. LUMEN là bộ thử mở mà doanh nghiệp và viện nghiên cứu cùng sử dụng. Người ta có thể đưa phương pháp điều khiển mới vào động cơ thật để thử.

Bộ thử P8.3 bắt đầu vận hành từ năm 2021. Bộ thử này được xây bằng cách mở rộng cơ sở trước đó. Trước đây chỉ có thể thử buồng đốt. P8.3 thử được toàn bộ động cơ, cả turbopump. Vì vậy có thể kiểm tra phương pháp điều khiển trong điều kiện thật. LUMEN và P8.3 trở thành nền tảng cho nghiên cứu điều khiển bằng trí tuệ nhân tạo.

DLR gắn điều khiển bằng trí tuệ nhân tạo vào động cơ này. Tài liệu chính thức của DLR giải thích rằng họ mô hình hóa truyền nhiệt trong kênh làm mát bằng mạng nơ-ron. Tài liệu cũng nói rằng điều khiển bằng trí tuệ nhân tạo dùng để khớp lực đẩy và theo dõi bất ổn cháy. Luận án tiến sĩ năm 2025 của Kai Dresia trình bày chi tiết nghiên cứu này. Luận án này điều khiển động cơ tên lửa bằng học tăng cường sâu. Sau đó so sánh kết quả với điều khiển dự báo mô hình truyền thống. Luận án báo cáo rằng khi điều khiển nhiều giá trị cùng lúc, sai số trung bình đạt khoảng 1,3 phần trăm. Giá trị này là kết quả do luận án của Dresia đưa ra. Đây là giá trị từ mô phỏng và nghiên cứu trên bộ thử, không phải giá trị đã được chứng minh trong bay cho toàn bộ vùng vận hành của một động cơ hoàn chỉnh.

Trong nghiên cứu trước đó, DLR đã đưa điều khiển bằng mạng nơ-ron vào thử nghiệm thật. Năm 2023, họ kiểm chứng điều khiển bằng mạng nơ-ron trong thử nghiệm turbopump của LUMEN. Họ cũng tiến hành nghiên cứu mô hình hóa truyền nhiệt trong kênh làm mát bằng mạng nơ-ron. Năm 2024, họ công bố kết quả thử nghiệm linh kiện của LUMEN và tiến độ nghiên cứu. Những nghiên cứu này tích lũy dần rồi dẫn tới điều khiển động cơ thật. Đây không phải là một lần thành công riêng lẻ, mà là sự tích lũy qua nhiều năm.

Hãy nhìn kỹ hơn một chút vào khung mà bộ điều khiển dùng để xử lý trạng thái và hành động. Bộ điều khiển đọc trạng thái ở từng khoảnh khắc rồi đưa ra hành động. Hành động đó tạo ra trạng thái tiếp theo. Từ trạng thái tiếp theo, nó lại đưa ra hành động. Quá trình này nối tiếp không ngừng. Bài toán trong đó trạng thái và hành động nối đuôi nhau như vậy gọi là bài toán quyết định tuần tự. Học tăng cường phù hợp để xử lý quyết định tuần tự này. Nó học không phải một phán đoán riêng lẻ, mà học dòng phán đoán nối tiếp.

Hãy xem nguyên lý bộ điều khiển học tăng cường phán đoán như thế nào. Trước hết bộ điều khiển đọc trạng thái của động cơ. Đó là các giá trị như áp suất buồng đốt, vị trí van, tốc độ quay của bơm, nhiệt độ. Những giá trị này được gọi là trạng thái. Bộ điều khiển nhìn vào trạng thái rồi quyết định sẽ di chuyển van bao nhiêu. Chuyển động này gọi là hành động. Nếu kết quả của hành động tốt, nó nhận thưởng; nếu xấu, nó nhận phạt. Nếu gần với lực đẩy mục tiêu, nó được thưởng. Nếu tỷ lệ hỗn hợp lệch hoặc rung động tăng, nó bị phạt. Quy tắc chứa thưởng phạt này gọi là hàm phần thưởng. Bộ điều khiển tinh chỉnh thao tác van theo hướng nhận được nhiều thưởng.

Thiết kế tốt hàm phần thưởng quyết định thành bại của điều khiển. Quy tắc thưởng phạt phải chứa đúng mục tiêu. Nếu chỉ thưởng cho lực đẩy, bộ điều khiển có thể bỏ qua rung động. Vì vậy lực đẩy, tỷ lệ hỗn hợp và rung động đều được đưa vào thưởng phạt. Việc di chuyển van quá gấp cũng bị phạt. Vì thao tác gấp gây quá tải cho linh kiện. Phải cân chỉnh tốt trọng số của nhiều mục tiêu. Việc cân chỉnh trọng số này là công việc cốt lõi trong thiết kế điều khiển.

Trong công nghiệp, vận hành tự động động cơ tái sử dụng cũng là chủ đề nóng. Một động cơ mới của một doanh nghiệp Mỹ nhắm tới tái sử dụng lặp lại. Động cơ này được biết là đã giảm số linh kiện và đổi sang hướng dễ chế tạo hơn. Có thông tin nói rằng họ bỏ tấm chắn nhiệt để giảm khối lượng và độ phức tạp. Những thay đổi này thuộc xu hướng dùng động cơ thường xuyên và bảo trì nhanh. Tuy nhiên đây là ví

dự phát triển của doanh nghiệp. Đây không phải là giá trị hiệu năng đã được kiểm chứng học thuật. Cuốn sách này chỉ nêu ví dụ này như một tham chiếu về xu hướng.

Điểm mạnh của cách này nằm ở chỗ nó xử lý nguyên cả bài toán liên kết chặt. Con người khó viết quy tắc thủ công cho từng van. Vì các giá trị liên kết với nhau nên quy tắc trở nên phức tạp. Học tăng cường tự học các quan hệ liên kết từ dữ liệu. Nó nhìn nhiều giá trị cùng nhau rồi tìm điểm cân bằng. Tuy nhiên một bộ điều khiển đã học một lần không tự thích nghi với tình trạng lão hóa mà nó chưa từng gặp. Có nhiều con đường để ứng phó với thay đổi của động cơ. Có cách đọc trạng thái trước rồi phản ánh vào điều khiển. Có điều khiển bền vững, được tạo ra để không dao động dù có thay đổi. Cũng có cách cho bộ điều khiển học lại bằng dữ liệu mới. Thích nghi trực tuyến trong khi vận hành tạo ra một rủi ro mới, đó là thay đổi điều khiển chưa được kiểm chứng. Vì vậy chọn con đường nào phải dựa trên đánh giá an toàn.

Có thể so sánh điều khiển học tăng cường với phương pháp hiện có. Một phương pháp đã dùng lâu là điều khiển dự báo mô hình. Phương pháp này dùng mô hình toán học của động cơ để nhìn trước rồi quyết định van. Nếu có mô hình chính xác, nó hoạt động tốt. Nhưng nếu động cơ cũ đi hoặc mô hình lệch đi, hiệu năng giảm. Học tăng cường học trực tiếp từ dữ liệu nên linh hoạt hơn trước những thay đổi như vậy. Hai phương pháp có ưu và nhược điểm riêng. Vì vậy các nhà nghiên cứu so sánh và kết hợp chúng.

Hãy xem vì sao việc giữ tỷ lệ hỗn hợp phù hợp lại quan trọng. Tỷ lệ giữa nhiên liệu và chất oxy hóa chính là tỷ lệ hỗn hợp. Nhiệt độ ngọn lửa thường cao nhất gần tỷ lệ mà nhiên liệu và chất oxy hóa vừa khớp. Nếu rời khỏi tỷ lệ này, dù dư nhiên liệu hay dư chất oxy hóa, nhiệt độ đều giảm. Tên lửa thường cố ý vận hành với lượng nhiên liệu hơi nhiều hơn. Mục đích là hạ nhiệt độ ngọn lửa chạm vào vách để bảo vệ vách. Nhưng nếu đưa quá nhiều nhiên liệu vào, lực giảm. Với nhiên liệu nhóm kerosene, lúc này muối bám lại. Động cơ dùng hydro hầu như không có muối. Như vậy tỷ lệ hỗn hợp phù hợp khác nhau tùy tổ hợp chất đẩy. Vì vậy trong khi thay đổi lực đẩy, vẫn phải giữ tỷ lệ hỗn hợp. Đây là lý do cần điều khiển nhiều giá trị cùng lúc.

Ở đây cần xử lý số liệu một cách thận trọng. Các giá trị hiệu năng chi tiết như thời gian deep throttling nhanh, phạm vi sai số tỷ lệ hỗn hợp, tỷ lệ giảm rung động chưa được xác nhận bằng các bài báo công khai. Nếu dùng giá trị hiệu năng chưa được xác nhận, ta sẽ thổi phồng sự thật. Vì vậy cuốn sách này không dùng các con số đó. Giá trị đã được xác nhận là sai số trung bình khoảng 1,3 phần trăm trong điều khiển nhiều biến. Chỉ giá trị này được dùng làm căn cứ.

Cần nói rõ ý nghĩa của sai số trung bình 1,3 phần trăm này. Đây là kết quả khi khớp nhiều biến cùng lúc. Tức là không phải giữ một mục tiêu, mà giữ nhiều mục tiêu cùng nhau. Với quy tắc do con người viết thủ công, khó đạt được sự cân bằng này. Trí tuệ nhân tạo đã học các quan hệ liên kết và tìm ra điểm cân bằng. Tuy nhiên giá trị này là kết quả trong điều kiện giới hạn của một nghiên cứu cụ thể. Đây không phải là giá trị có thể chuyển nguyên sang mọi động cơ. Nếu điều kiện thay đổi, phải kiểm tra lại.

Vẫn còn một quãng đường trước khi dùng trực tiếp công nghệ này cho động cơ bay. Bộ điều khiển học tăng cường trước hết học trong mô phỏng. Trong máy tính, sai lầm không gây nguy hiểm. Sau đó người ta nối phần cứng vào để thử nghiệm. Thử nghiệm này gọi là hardware-in-the-loop. Người ta chạy linh kiện thật cùng với mạng nơ-ron. Tiếp theo kiểm tra an toàn bằng thử nghiệm đốt trên mặt đất. Phải vượt qua tất cả các bước này rồi mới bàn đến bay. Vì chứng nhận tên lửa rất nghiêm ngặt.

Ở đây có một khó khăn lớn là khác biệt giữa trình mô phỏng và thực tế. Mô hình máy tính không giống hết động cơ thật. Một cách điều khiển học tốt trong mô hình có thể lệch đi trong thực tế. Khác biệt này được gọi là khoảng cách giữa trình mô phỏng và hiện thực. Muốn giảm khoảng cách này, phải làm mô hình chính xác hơn. Phải liên tục chỉnh mô hình bằng dữ liệu thử nghiệm thật. Những bộ thử như LUMEN được dùng cho việc chỉnh sửa này. Phải có dữ liệu thật thì mới có thể tin trí tuệ nhân tạo.

Một nguyên tắc nữa là đặt nhiều lớp thiết bị an toàn. Bộ điều khiển học tăng cường có thể đưa ra lệnh lạ. Để chuẩn bị cho tình huống đó, cần đặt kèm thiết bị an toàn dựa trên quy tắc. Nếu lệnh của bộ điều khiển vượt qua ranh giới nguy hiểm, thiết bị an toàn sẽ chặn lại. Trí tuệ nhân tạo nâng hiệu năng, còn quy tắc giữ mức sàn an toàn. Phải có hai lớp này thì mới có thể gắn trí tuệ nhân tạo vào động cơ thật.

Van cũng có giới hạn vật lý. Van không thể mở hoặc đóng nhanh hơn tốc độ đã định. Tốc độ tối đa này gọi là giới hạn tốc độ truyền động của van. Nếu trí tuệ nhân tạo đưa ra lệnh gấp vượt quá giới hạn này, đường ống sẽ chịu va đập. Hiện tượng búa nước, khi chất lỏng đột ngột dừng lại và đập vào ống, là một ví dụ. Vì vậy lệnh của bộ điều khiển được giới hạn trong tốc độ mà van có thể thực hiện. Ngoài ra, nếu trí tuệ nhân tạo dao động, hệ thống lập tức chuyển sang bộ điều khiển cổ điển bằng một thiết bị chuyển mạch an toàn. Bộ điều khiển cổ điển bảo vệ động cơ an toàn bằng các quy tắc đã được kiểm chứng lâu dài.

Giá trị của điều khiển học tăng cường thể hiện rõ trong tái sử dụng. Khi dùng một động cơ nhiều lần, trạng thái mỗi lần lại khác. Quy tắc điều khiển cố định không theo kịp thay đổi này. Bộ điều khiển học tăng cường đọc trạng thái rồi điều chỉnh van cho phù hợp. Động cơ cũ được xử lý theo trạng thái cũ, động cơ mới được xử lý theo trạng thái mới. Sự linh hoạt này làm giảm chi phí của tên lửa tái sử dụng. Vì vậy nhiều tổ chức trên thế giới đang dồn sức cho nghiên cứu này.

Tuy nhiên điều khiển bằng trí tuệ nhân tạo hiện vẫn ở giai đoạn thử nghiệm mặt đất. Nó chưa được dùng rộng rãi trong động cơ bay thật. Chứng nhận bay đòi hỏi thử nghiệm và hồ sơ trong thời gian dài. Phải chứng minh rằng cách điều khiển mới an toàn trong mọi tình huống. Việc chứng minh này cần nhiều dữ liệu và thời gian. Vì vậy điều khiển bằng trí tuệ nhân tạo hiện đang lớn lên trong phòng thí nghiệm và trên bệ thử. Khi kết quả tích lũy đủ, một ngày nào đó nó sẽ dẫn tới bay.

## 10.6 Kết luận và triển vọng tương lai

Mục này xem bốn công nghệ ở trên được kết nối thành một thể như thế nào. Sau đó trình bày vì sao sự kết hợp này cần thiết trong thời đại tên lửa tái sử dụng. Cuối cùng, mục này nêu các nguyên tắc an toàn cần có cho quản lý động cơ bằng trí tuệ nhân tạo.

Vận hành thông minh động cơ tên lửa nhiên liệu lỏng nối liền bốn công nghệ. Chẩn đoán bằng cảm biến nhìn nhiều tín hiệu cùng lúc, đọc sớm các dấu hiệu yếu và tìm bất thường sớm. Bản sao số khôi phục dòng áp suất bên trong không nhìn thấy bằng mạng nơ-ron tuân thủ phương trình vật lý. Mô hình Bayes đo sự bất định của lão hóa và sai lệch, rồi đưa ra câu trả lời cùng với độ rộng của câu trả lời. Bộ điều khiển học tăng cường điều chỉnh nhiều van cùng lúc và tìm điểm cân bằng của các giá trị liên kết với nhau.

Bốn công nghệ này trao đổi dữ liệu với nhau. Khi chẩn đoán phát hiện bất thường, bộ điều khiển phản ứng. Khi bản sao số cho biết trạng thái bên trong, chẩn đoán trở nên chính xác hơn. Khi mô hình Bayes cho biết độ bất định, bộ điều khiển hành động thận trọng hơn. Như vậy bốn công nghệ vận hành trong một vòng khép kín. Đây là cấu trúc trong đó động cơ tự theo dõi trạng thái của mình và tự điều chỉnh. Cấu trúc này được gọi là quản lý sức khỏe tự chủ.

Bốn công nghệ này bù đắp điểm yếu cho nhau. Chỉ chẩn đoán thì khó biết chính xác nguyên nhân của bất thường. Khi bản sao số cho thấy bên trong, nguyên nhân trở nên rõ ràng. Chỉ bộ điều khiển thì không biết động cơ đã cũ đi ra sao. Khi mô hình Bayes cho biết độ bất định, việc điều khiển trở nên an toàn hơn. Khoảng trống của một công nghệ được công nghệ khác lấp đầy. Vì vậy, ta không tách chúng ra, mà gộp chúng thành một hệ. Chỉ khi được gộp lại, quản lý tự hành mới hoàn chỉnh.

Dòng phát triển này chưa phải là sản phẩm hoàn thiện. Phần lớn vẫn ở giai đoạn nghiên cứu và thử nghiệm. Nhiều kết quả còn là nghiên cứu ban đầu chưa qua bình duyệt. Vì vậy, còn quá sớm để khẳng

định chắc chắn về thành quả. Cuốn sách này nói rõ điểm đó. Hiện nay là giai đoạn hiểu khung xương và hướng đi của công nghệ. Cần thêm thời gian trước khi nó được dùng rộng rãi trong động cơ bay thực tế.

Dòng phát triển này cần thiết trong thời đại tên lửa tái sử dụng. Khi dùng một động cơ nhiều lần, trạng thái của nó mỗi lần đều khác. Chỉ các quy tắc cố định thì khó xử lý hết mọi thay đổi. Các giá trị thay đổi quá nhanh và liên kết quá chặt để con người chỉnh bằng tay mỗi lần. Trí tuệ nhân tạo là công cụ đọc nhanh các thay đổi này và đáp ứng. Muốn giảm chi phí vận tải vũ trụ, phải có quản lý tự hành như vậy.

Tuy nhiên, điều khiển bằng trí tuệ nhân tạo bắt buộc phải có thiết bị giới hạn. Một là khả năng giải thích. Con người phải biết vì sao mô hình phán đoán như vậy. Hai là kiểm chứng độc lập. Phán đoán của mô hình phải được kiểm tra lại bằng phương pháp khác. Ba là chuyển sang trạng thái an toàn khi thất bại. Dù mô hình dao động, động cơ vẫn phải có thể dừng lại an toàn. Không có ba thiết bị này thì khó giao động cơ cho trí tuệ nhân tạo.

Trí tuệ nhân tạo không phải là công cụ để ta tin thay cho động cơ. Trí tuệ nhân tạo là công cụ để nhìn động cơ kỹ hơn. Con người mượn con mắt của trí tuệ nhân tạo để hiểu động cơ sâu hơn. Phán đoán cuối cùng và trách nhiệm vẫn thuộc về con người. Trên nguyên tắc đó, trí tuệ nhân tạo làm cho tên lửa an toàn hơn và rẻ hơn.

## Tổng kết

Ta điểm lại phần cốt lõi của chương này. Động cơ tên lửa nhiên liệu lỏng là cỗ máy rất nhanh, rất nóng và nguy hiểm. Muốn dùng động cơ này an toàn, cần giám sát và điều khiển theo thời gian thực. Cách dùng ngưỡng giới hạn trước đây dễ bỏ sót bất thường yếu. Trí tuệ nhân tạo nhìn nhiều cảm biến cùng lúc để lấp khoảng trống này. Mục tiêu là đọc các dấu hiệu yếu trước khi giá trị vượt qua ranh giới nguy hiểm.

Chẩn đoán sớm nhìn cả dòng thời gian và quan hệ giữa các kênh. Kiến trúc kết hợp tích chập thời gian và attention là ví dụ tiêu biểu. Kiến trúc này tìm bất thường sớm từ nhiều cảm biến. Phán đoán phải kết thúc trong vài mili giây trên con chip nhỏ đặt gần động cơ. Tuy nhiên, các chỉ số hiệu năng riêng cho tên lửa hiện vẫn khó xác nhận. Vì vậy, cuốn sách này chỉ giải thích ở mức nguyên lý. Chẩn đoán được dùng cùng với sự giám sát của con người và thiết bị an toàn dựa trên quy tắc.

Bản sao số vẽ ra bên trong buồng đốt vốn không thể nhìn thấy. Nó dùng mạng nơ-ron vật lý tuân thủ các phương trình vật lý. Chỉ với vài cảm biến trên vách, nó khôi phục toàn bộ dòng áp suất. Nó rất hữu ích trong giám sát mất ổn định cháy. Tuy nhiên, nó không thể chứa hết ngọn lửa phức tạp. Nó được dùng để giám sát nhanh và được bổ sung bằng phân tích chính xác. Bản sao số chỉ làm đúng vai trò khi được hiệu chỉnh liên tục bằng tín hiệu thực.

Mạng nơ-ron Bayes đo độ bất định của lão hóa và sai lệch. Nó đưa ra đáp án cùng với khoảng của đáp án. Khoảng hẹp nghĩa là nó tự tin, khoảng rộng nghĩa là nó chưa biết rõ. Thông tin này được dùng để bộ điều khiển quyết định có cần thận trọng hay không. Nó cũng được dùng để quyết định khi nào bảo trì. Tuy nhiên, xác suất của mô hình phải được kiểm chứng bằng dữ liệu thực.

Bộ điều khiển học tăng cường điều chỉnh nhiều van cùng lúc. DLR của Đức nghiên cứu việc này với động cơ LUMEN và bộ thử P8.3. Tài liệu công khai báo cáo sai số trung bình khoảng 1,3 phần trăm trong điều khiển nhiều biến. Một số chỉ số hiệu năng chi tiết chưa đủ xác nhận nên không được dùng ở đây. Muốn dùng công nghệ này trong bay thực tế, vẫn còn nhiều bước kiểm chứng.

Ta nhắc lại một nguyên tắc xuyên suốt chương này. Trí tuệ nhân tạo là công cụ nâng hiệu năng. Nhưng an toàn do quy tắc và con người giữ. Phải có cả hai lớp này thì mới đặt trí tuệ nhân tạo vào động cơ thực. Không tin các con số chưa được kiểm chứng. Cần đọc tách bạch giữa nghiên cứu ban đầu và thành quả đã được kiểm chứng. Thái độ này là con đường đúng khi xử lý một công nghệ nguy hiểm như tên lửa.

Khi bốn công nghệ này được gộp thành một, động cơ tự chăm sóc sức khỏe của mình. Bốn công nghệ trao đổi dữ liệu cho nhau và quay trong một vòng khép kín. Quản lý tự hành này là nền tảng của thời đại tên lửa tái sử dụng. Vì khi dùng một động cơ nhiều lần, trạng thái của nó mỗi lần đều khác. Tuy nhiên, vì an toàn, cần có khả năng giải thích, kiểm chứng độc lập và thiết bị chuyển sang trạng thái an toàn đi kèm.

Các công nghệ trong chương này trả lời những câu hỏi khác nhau. Chẩn đoán hồi hiện nay có gì đang sai. Bản sao số hồi bên trong không nhìn thấy đang ra sao. Mô hình Bayes hồi nên tin phán đoán này đến mức nào. Bộ điều khiển hồi vậy phải làm gì. Khi bốn câu hỏi này gom lại, trạng thái của động cơ và cách đáp ứng nối thành một mạch. Quản lý động cơ bằng trí tuệ nhân tạo tốt phải xử lý cùng lúc bốn câu hỏi này.

Cuối cùng, ta nói rõ thái độ đối với tài liệu trong chương này. Nhiều nghiên cứu mới mà chương này nêu còn ở giai đoạn ban đầu. Nhiều bài báo là nghiên cứu sơ bộ chưa qua bình duyệt. Các ví dụ động cơ của doanh nghiệp chỉ cho thấy dòng chảy công nghiệp. Chúng có trọng lượng khác với thành quả đã được kiểm chứng học thuật. Vì vậy, chương này đặt trọng tâm vào các giá trị và nguyên lý đã được xác nhận. Các con số chưa xác nhận được loại bớt. Mong rằng độc giả hiểu đúng hướng đi của công nghệ. Trí tuệ nhân tạo là công cụ trong hành trình dài làm cho tên lửa an toàn hơn và rẻ hơn.

### Tài liệu căn cứ liên quan

- 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., & Freiburger, 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., & Oswald, 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., Aljaseem, 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.res.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., Oswald, 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>

# Ủng hộ cuốn sách Thư viện AI này

Nếu cuốn sách này hữu ích với bạn, bạn có thể ủng hộ các bản dịch tiếp theo, các ấn bản công khai và những dự án tri thức mở về trí tuệ nhân tạo qua PayPal.

PayPal

<https://paypal.me/kimkjkj>



Quét mã QR hoặc mở đường dẫn ở trên.

Luật sư Kim Kyung-jin