ARTIFICIAL INTELLIGENCE IN HORTICULTURE

Kim Kyung-jin, Attorney at Law

This book was organized and produced with artificial intelligence. It brings together materials collected from major languages around the world.

Contents

Chapter 1

Computer Vision and Deep Learning-Based Crop Pest and Disease Diagnosis

1.1

Precision Detection and Diagnosis of Crop Pests, Diseases, and Weeds

The LaserWeeder, built by the American company Carbon Robotics, burns away more than 100,000 weeds per hour with lasers. Since its launch in 2022, the robot has eliminated more than 10 billion weeds in total; its revenue exceeded $100 million in the first quarter of 2026; and about 100 farms in 15 countries are using it. During the same period in South Korea, an AI diagnostic app developed by the Rural Development Administration has been identifying 31 crops and 182 diseases and pests.

We have entered an era in which cameras and algorithms are taking over work once done by people turning over leaves one by one to inspect lesions and pulling weeds by hand.

The yield and quality of horticultural crops support both national food security and farm income. Yet as the climate changes, pathogens acquire new mutations and spread into unfamiliar regions, while weeds compete with crops for the same sunlight, water, and nutrients. For many years, the human eye stood on the front line of this battle. In English, the practice is called visual scouting. This visual inspection, in which an experienced agronomist or grower walks through a field examining leaves, remains widespread. But there is a limit to how many rows a person can inspect in a day, and different people may judge the same lesion differently.

As fields grow larger, the gaps in this method widen, and disease spreads silently through them.

When weeds begin growing beside a crop, an invisible struggle unfolds beneath the roots every day. During the brief intervals when rainwater seeps into the soil and fertilizer dissolves, weed roots repeatedly draw in water and nutrients before crop roots can reach them. It takes weeks before the difference becomes visible in leaf size, but by then the crop reaches harvest without having grown as much as it should. Whether the problem is disease or weeds, by the time the symptoms become obvious, it is often already too late to intervene.

The broad movement responding to these problems is called precision agriculture. It marks a shift away from applying the same treatment across an entire field and toward intervening only as much as needed for each small zone or individual plant. The precise detection of diseases, pests, and weeds is the first step in that transition. A precise treatment aimed only at the affected location is impossible when no one knows what is present or where it is.

Precision diagnostic technology combining computer vision and deep learning has emerged to close this gap. Instead of a person, an algorithm distinguishes diseased areas from healthy ones and crops from weeds in a single camera image. At its foundation is the convolutional neural network, or CNN. It scans the colors, patterns, and curves along lesion margins in each small image patch, then stacks the results layer by layer and brings them together into a single judgment. With transfer learning, a model already trained on another large image dataset, such as ResNet, EfficientNet, DenseNet, or MobileNet, is retrained on crop images.

This can produce a fairly accurate classifier with only a few thousand photographs, a volume that farms can realistically collect. The principle resembles the way someone who already knows one language can learn another more quickly.

In a fast-moving field environment, speed matters as much as accuracy. Models in the YOLO (You Only Look Once) family scan an entire image just once, draw boxes around lesions, pests, or weeds, and label each one at the same time. True to the name, the method finishes after a single look and can keep pace with video streaming from a camera in real time. Even in a cluttered canopy where leaves overlap, shadows fall, and soil and straw fill the background, recent generations from YOLOv5 through YOLOv8, YOLOv9e, YOLOv10, and YOLOv11 locate individual lesions, single insects, and individual weeds without missing them.

CNN and YOLO share a weakness. Because they divide an image into small patches, they can miss relationships between patches, broad patterns extending across an entire leaf, or changes in the overall form of a plant. The Vision Transformer, abbreviated ViT, and its relative, the Swin Transformer, divide an image into 16×16-pixel patches and calculate all the relationships among them. This calculation, known as the self-attention mechanism, allows the model to take a wide view and determine whether a lesion in one corner of an image and another in the opposite corner come from the same disease.

A hybrid architecture that places attention modules such as CBAM or CAB alongside a CNN gives one system both an eye for fine lesion patterns and an eye for the whole leaf. It is as if one person had fingertips tracing the texture up close and eyes taking in the entire picture from a distance.

Numbers reveal how accurately these models perform. EfficientNet-B0 and B3 variants that identify grape leaf diseases recorded 99.5 percent accuracy and an AUC-ROC of 1.00 on validation data. A model combining DSOCA-QSO, an optimization technique inspired by the movement of quokka groups, reached 99.9 percent accuracy and an F1 score of 99.35 percent in classifying vegetable leaf diseases. TL-DCFP-CNN, which selects features according to context, maintained 98.09 percent accuracy while handling five crops at once: cotton, maize, rice, wheat, and sugarcane.

In real-time detection, YOLOv11 delivered about 94 percent accuracy in detecting tomato leaf diseases. The smaller YOLO-NAS maintained 96 percent accuracy in weed detection in tropical fields while processing one image in 25 milliseconds. That is less than one-tenth the time it takes a person to blink. The tropical fields where YOLO-NAS was tested differ from temperate fields in light and humidity. The sun stands high, shadows are short and dark, and the air is humid. Its ability to maintain accuracy in those conditions suggests that a model built in one climate zone may work in a completely different one.

Unmanned aerial vehicles and ground robots move continuously across fields taking photographs. If a judgment arrives even half a beat late, the robot has already passed the lesion. Field performance therefore depends not only on how accurately a system identifies disease, but on how quickly it reaches that judgment. A difference of milliseconds in processing a single image ultimately determines how much land one robot can cover in a day.

When crops and weeds grow with their leaves touching, bounding boxes are not enough. Segmentation models take over, assigning every pixel to either crop or weed. PSPEdgeWeedNet combined a separate edge-detection branch with a conditional random field to draw a clear boundary between crops and weeds in peanut fields. CAMDWS, a dual-branch architecture that also reads context, identified rigid ryegrass (Lolium rigidum) and wild radish (Raphanus raphanistrum), two major weeds in Australian wheat fields, at IoU values above 91 percent and 90 percent, respectively. IoU is the proportion by which the actual and predicted areas overlap.

ViT achieved 99.8 percent accuracy on PlantVillage data. HybridViT-CAB, which combines ViT with a convolutional attention block, secured an AUC of 0.97 and 91 percent accuracy with a lightweight footprint of only 412,000 parameters and 1.57 megabytes, processing one stage in 141 milliseconds. An airborne diagnostic system using Swin Transformer as its backbone and adding ConvNeXt-V2, SE-BiLSTM, and a deep Q-learning recommendation layer recorded 99.30 percent accuracy on the SAR-CLD-2024 cotton dataset. The center of research appears to be shifting toward models that remain accurate while becoming lighter.

Even the finest algorithm is constrained by what its camera captures. RGB sensors record wavelengths similar to those seen by the human eye. They are cheap, light, and already built into phones and drones, making them standard equipment for symptom classification and weed detection. Yet visible light has the same limits as human vision when it comes to detecting early infection developing inside a leaf before symptoms surface, or distinguishing a green weed that is almost the same color as the crop. Multispectral and hyperspectral imaging fill this gap.

Sensors that separately measure reflectance in selected bands, such as blue at 475 nanometers, green at 560 nanometers, red at 668 nanometers, red edge at 717 nanometers, and near-infrared at 840 nanometers, turn those readings into indices such as the normalized difference vegetation index (NDVI), green normalized difference vegetation index (GNDVI), red-edge NDVI, and soil-adjusted vegetation index (SAVI). They detect early stress caused by chlorophyll breakdown or water loss before visible symptoms appear.

An RGB-D sensor adds three-dimensional positional information by attaching a distance value to every pixel in a color image. This depth information produces the coordinates that tell the fingers of a harvesting robot or the laser of a weeding robot exactly where to aim. Unmanned aircraft carrying these sensors fly at low altitudes between 2 and 50 meters and scan an entire field at resolutions in which one pixel represents 0.3 to 5 centimeters. Ground robots and handheld mobile devices fill the remaining gaps.

When these readings are layered with Internet of Things sensors measuring air temperature, humidity, soil moisture, and rainfall, as well as weather records, multiple layers of data across time and space flow into one pipeline. The collected data move beyond diagnosis into prediction. Overlaying the temperature and humidity trends of the past several days with past disease records can warn of the moment when infection risk starts to rise, before lesions are visible. The technology moves one step beyond finding visible disease to identifying disease that cannot yet be seen.

Collected photographs cannot be used immediately. They are cropped to fixed dimensions such as 224×224, 512×512, or 736×736, and their pixel values are normalized to a set range. To compensate for changing sunlight direction and uneven numbers of photographs among disease categories, images are rotated, flipped, tilted, and blurred. Generative models such as GAN, cGAN, and DoubleGAN create synthetic photographs of diseased leaves and weeds, increasing the volume of training data. A conditional generative adversarial network, or cGAN, takes the type of disease as a preset condition and draws a new synthetic lesion image corresponding to it.

One side tries to create an image that looks real while the other tries to determine whether it is fake, and both improve through the contest. Functions such as cross-entropy loss, class-weighted loss, and boundary loss measure how far the model is wrong, while optimization methods such as Adam and AdamW gradually reduce the error.

Once trained, the model must ultimately run inside a small machine out in the field. Palm-sized computers such as the NVIDIA Jetson Xavier NX, Jetson Nano, and Raspberry Pi, and FPGA accelerators such as the Xilinx Zynq-7000, do the work. Quantization rounds finely spaced decimal values into integers, reducing the number of bits handled by each operation and cutting both processing time and memory use. Precision declines slightly, but computation becomes far faster. Pruning removes connections that have little effect on the result, like cutting unneeded branches from a tree.

Converting a model for TensorFlow Lite or TensorRT can cut operating time sharply. FPGA-accelerated models have reportedly delivered inference 43.2

times faster than RGB-only models, latency below 10 milliseconds, and an 88 percent reduction in power consumption. More recent systems process one image in 42.6 milliseconds, enabling real-time diagnosis on phone-class devices. Ultralight models such as PMVT, with 1.9 million parameters, and pruned transformers now run directly on agricultural drones and compact devices using TensorFlow Lite, ONNX Runtime, and lightweight gRPC communication.

Several measures determine whether a model built this way is useful. Accuracy is the proportion of all judgments that are correct. Precision is the proportion of cases classified as diseased that truly are diseased. Recall is the proportion of actually diseased leaves that the model catches without missing. Some diseases have thousands of photographs, while rare diseases may have only a few dozen. When image counts differ this sharply among diseases, the F1 score combines precision and recall to show whether the model is biased toward common diseases.

AUC-ROC shows performance across multiple thresholds, while a confusion matrix places the true disease name beside the model's prediction in a table, exposing exactly which disease was mistaken for which. Misclassifying a diseased leaf as healthy gives infection time to spread across the field. For disease and pest diagnosis, recall therefore deserves more weight than precision. It is safer to be slightly over-suspicious than to miss a case.

Detection and segmentation, which must locate a target as well as identify it, use different measures. IoU shows how much the region drawn by a model overlaps the true lesion, and mAP is the mean precision measured across multiple thresholds. It resembles stamping two seals in the same place and checking how closely their impressions overlap. Frames processed per second and inference time determine whether a model can operate at actual field speed. Yet even excellent scores are worthless if the model has no value in the field.

Agricultural measures such as weed removal rate, crop damage rate, herbicide use per hectare, and spraying efficiency must therefore be examined alongside them. Whether a computer predicts correctly and whether a farm's costs and yields actually improve are different questions. A model that cannot answer the second question will struggle to take root in the field, however high its accuracy.

Behind the numbers are cameras and apps operating in real fields. Diagnostic systems pairing mobile apps with drones are spreading widely, centered on valuable horticultural crops such as tomatoes, grapes, cotton, peppers, and strawberries.

Models running YOLOv8, YOLOv11, EfficientNet, or MobileNet identify diseases directly in the field: early blight and leaf mold in tomatoes; mosaic viruses that distort leaf veins; yellow leaf curl virus, which turns leaf margins yellow and curls them upward; downy mildew, black rot, and esca, which dries the tissue between grape leaf veins into yellow stripes; aphids and two-spotted spider mites on cotton; and leafminers that tunnel into leaves. A grower takes one picture in front of an affected plant, then chooses the next action after a disease name and control method appear on the screen seconds later.

Trials in plots planted with tomatoes, potatoes, and maize have shown what happens after diagnosis. An early-stage fungicide was sprayed only in areas the deep-learning model classified as diseased, and irrigation was adjusted separately in those areas. Diagnosis did not end as a number on a screen; it led to a decision that changed the actual amounts of chemicals and water. Yet many trials remain confined to one field and one season, prompting calls for larger datasets spanning multiple plots, seasons, and light levels.

These diagnostic tools are powerful in regions where expert help is not readily available. The number of farms and fields one agronomist must cover varies widely by region, but it is too large anywhere to visit them all on foot. The ability to receive a diagnosis close to an expert judgment immediately through a single phone appears to be one reason the technology is spreading. Its role is less to replace people than to fill the times and places their eyes cannot reach.

The practice of using cameras to distinguish crops from weeds and treating only the necessary spots is called site-specific weed management, or SSWM. Its central principle is the shift from spraying an entire field at the same rate to selecting only the places where weeds are present and treating only those areas. Drawing that boundary accurately is difficult. The younger the crops and weeds are, and the closer they grow to one another, the more indistinct the boundary becomes.

Separating crops from weeds is harder than finding disease. Both are green, both grow, and their leaf shapes can even resemble one another when they are young. An experiment in a 120-hectare Australian wheat field addressed this problem through time. A low-altitude unmanned aircraft photographed the same field at eighteen different times, and a spatiotemporal-spectral fusion algorithm called AMSTF read the images as overlapping layers. The daily growth-rate index was 0.025 for weeds and 0.015 for wheat, a difference clearly expressed in the time-series spectral indices.

After CAMDWS segmentation was added, the misclassification rate at the crop-weed boundary fell from a range of 9.8 to 12.3 percent to 5.4 percent. One subtle clue, the difference in growth speed, became the key to redrawing the boundary.

The targeted control system called See & Spray mounts a camera above each nozzle and an ultra-high-speed electronic valve below it. The equipment was developed by Blue River Technology, a subsidiary of John Deere. In real time, it recognizes the problem known as green-on-green: green weeds hidden among green crops. Even to the human eye, green weeds among green leaves often do not stand out at a glance. Depending on the density and type of weeds, the system adjusts droplet size between 150 and 250 micrometers and the spray rate between 0.12 and 0.35 liters per minute.

Compared with conventional broadcast spraying, which distributes herbicide evenly across the entire field, it cuts herbicide use by 31 to 32 percent, bringing the rate down to about 0.68 liters per hectare. The crop avoids the chemical, and the total amount released onto the soil falls as well.

Unlike this chemical method, a path that uses no chemicals at all has entered commercial deployment. An autonomous robot carrying the DIN-LW-YOLO detection algorithm and an RGB-D camera locates in three-dimensional coordinates the growing point of a newly emerged weed seedling between crop plants. The growing point is the topmost part from which the plant produces new growth. The robot fires a momentary high-powered laser at that point and destroys it with heat. It kills the weed with light instead of droplets of herbicide.

The LaserWeeder developed by Carbon Robotics in the United States is a leading example of bringing this method to commercial scale. It removes more than 100,000 weeds per hour and has eliminated more than 10 billion in total since its 2022 launch. Its annualized revenue surpassed $100 million in the first quarter of 2026 and was later reported at as much as $177 million, a scale without precedent for an agricultural robotics company. Robots have been deployed in 15 countries and are used by more than 100 growers across North America, Europe, and Australia.

In February 2026, the company released a large plant model trained on 150 million labeled plant images. It said the custom training that had once taken weeks whenever a robot encountered a new field or unfamiliar crop had been cut to minutes. A second robot no longer has to relearn from the beginning every combination of weeds, soil, and light experienced by the first. The experience accumulated by one robot can be carried directly to the next robot and the next field, suggesting that laser weeding has reached a threshold between being a tool for one farm and becoming a common model shared by many farms.

Returning to that greenhouse at dawn, the grower is now more likely to be holding a phone than a flashlight. The hand that once turned over leaves and gauged lesions by eye is becoming a hand that takes one picture and checks the disease name on a screen. In South Korea, the Rural Development Administration has developed and operates its own app for this purpose. When a grower photographs and uploads a diseased leaf or fruit, AI immediately identifies the disease and recommends a suitable control chemical. It currently covers 31 major crops and 182 diseases and pests, with a plan to expand by 2030 to 139 major crops grown in South Korea.

In Korea's second-generation smart farms, cloud-based AI reads growing conditions and crop biometric data in real time and detects five diseases, including late blight and powdery mildew, as well as two pests, including leafminers. The agency also operates an agriculture-specific AI assistant called “AI Isagi,” connected to 17 information systems, and is preparing a new version that answers spoken questions and even diagnoses farm management.

The camera that finds disease, the camera that picks out weeds, and the equipment that sprays chemicals or fires a laser in response ultimately form one continuous flow. It is a transition from treating the whole field in the same way to distinguishing each leaf and fruit and prescribing a different response for each. The cameras and robots now being tested in fields stand on that boundary. How widely this movement spreads will depend on accuracy, cost, and the grower's trust.

Despite these results, a field is not a laboratory. Datasets commonly used to train deep-learning models, such as PlantVillage, consist of images with clean backgrounds and even lighting. In a real field, however, the angle of sunlight changes with season and time of day, shadows deepen, leaves shake and overlap in the wind, dust settles on lenses, and soil and dry stems clutter the background. When a model trained on laboratory images is transferred directly into such a field, its accuracy drops visibly in what is known as domain shift. Once the imaging environment changes, the model mistakes the same disease for something else.

A difference in the color temperature of the light, the angle of view of the lens, or the direction of the shadow on a leaf can destabilize the judgment. No matter how many images are collected, what the model learned cannot be used unchanged in a field if the environment in which those images were taken differs from the real one.

Labeling itself is another bottleneck. Drawing bounding boxes and naming every pixel requires an agronomy expert and can take several minutes per image. Researchers are studying semi-supervised learning, which learns from a small number of labels; self-supervised learning, which generates its own answers as it learns; and few-shot learning, which recognizes a new disease after seeing only a few images. Large vision-language models and generative AI that produce labeled synthetic images in bulk have emerged as another route around this bottleneck.

More sensors mean more computation. Combining visible-light images in real time with shortwave-infrared and hyperspectral data and Internet of Things readings of soil and weather begins with the difficult task of aligning data streams arriving at different speeds. Images sent by a drone and readings sent by a ground sensor differ in both measurement interval and update rate, so a separate process is needed to match the same location at the same time. A battery-powered compact device must then run detection, inference, path planning, and spray-rate control at once without delay.

Solving that problem requires both shrinking the model and adding dedicated hardware such as an FPGA or NPU.

The final challenge is trust. A deep-learning model is close to a black box: it does not explain why it reached a judgment. Before growers and agronomists can trust that judgment and apply a control chemical, they need to see which part of the leaf caused the model to classify it as diseased. Research that embeds Grad-CAM, SHAP, or causal attention mechanisms in the model and visualizes the contribution of each lesion area appears to be the next step toward building that trust. No matter how accurate the technology is, it will be difficult to accept in the field if the grower doubts its judgment.

Sources and References

Academic Papers

1. Adewumi, A., & Saraswat, D. (2026). Comprehensive review of artificial intelligence and edge computing for precision weed control. Artificial Intelligence and Applications. DOI: 10.47852/bonviewAIA62028867.
2. Kakde, S., Poyam, P., et al. (2025). Design of an integrated spatiotemporal deep learning framework for autonomous precision weed management. International Journal of Applied Mathematics, 38(2s), 1340-1363.
3. Latif, A., Jati, H., Surjono, H. D., & Yusuf, M. (2026). Drone-assisted deep learning weed detection for sustainable agriculture and environmental resilience. IAES International Journal of Artificial Intelligence (IJ-AI), 15(2), 1428-1440. DOI: 10.11591/ijai.v15.i2.pp1428-1440.
4. Mathew, M. P. (2025). A comparative deep learning framework for grape leaf disease classification using EfficientNetB0, InceptionV3, and Xception. Discover Applied Sciences, 7, 1222. DOI: 10.1007/s42452-025-07457-5.
5. Moldvai, L., & Nyéki, A. (2025). Innovative computer vision methods for tomato (Solanum lycopersicum) detection and cultivation. Discover Applied Sciences, 7, 975.
6. Pai, D. G., Balachandra, M., & Kamath, R. (2025). Deep learning approach for crop-weed segmentation in peanut cultivation using PSPEdgeWeedNet. Scientific Reports, 15, 43032. DOI: 10.1038/s41598-025-24174-6.
7. Yenna, G. R., Kumari, N. V., Chouhan, A. S., Umar, S., & Macha, S. R. (2026). HybridViT-CAB: a vision transformer and convolutional attention network for precision weed detection in agricultural systems. Scientific Reports, 16, 18473. DOI: 10.1038/s41598-026-48652-7.
8. FPGA-accelerated IoT deployment of a causal attention multi-modal deep learning network for precision crop disease monitoring. (2026). Environment Conservation Journal.
9. AgroDualNet: a dual deep learning-based crop disease forecasting and fruit ripening detection framework. (2026). Scientific Reports.
10. Precision agriculture at scale: early disease detection using vision transformers. (2026). International Journal for Research in Applied Science & Engineering Technology (IJRASET), 14(4).
11. A novel approach for crop disease detection using transfer learning-based deep convolutional neural network. (2026). Neural Computing & Applications, 38, 500. DOI: 10.1007/s00521-026-12216-7.
12. Pest identification in crop fields using convolutional neural networks (CNNs): a deep learning approach. (2026). Volume 14, Issue 3.
13. PMC13066816. (2025). National Library of Medicine, PubMed Central. https://pmc.ncbi.nlm.nih.gov/articles/PMC13066816/
14. Article PII S2095809925006769. (2025). Elsevier ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2095809925006769

Lectures and Video Materials

1. C-DAC. (2025). AI in Agriculture: Transforming India's Krishi Landscape. YouTube. https://www.youtube.com/watch?v=lV3_VnWiBpM

2. RMCG. (2025). AI unpacked: Practical tools for vegetable growers. YouTube. https://www.youtube.com/watch?v=G2ytLXLW7hw

3. Grounded AI. (2025). How AI Is Transforming Farming: Smarter Cows, Crops & Pest Control. YouTube. https://www.youtube.com/watch?v=bxlF2MpXsB0

4. SyncTellect Automation Hub. (2026). End-to-End AI/ML Project Series | Day 1 - Project Kickoff & Crop Disease Detection Overview. YouTube. https://www.youtube.com/watch?v=SNqOcVib3BY

5. Royal Society of Biology. (2026). Série Santé des Plantes | Viticulture numérique : L'avenir de la santé de la vigne. YouTube. https://www.youtube.com/watch?v=5iXPOMxl6bU

Articles and Reports

1. Carbon Robotics. (2026). LaserWeeder Results. Carbon Robotics. https://carbonrobotics.com/results

2. Hypersinc. (2026). Carbon Robotics crosses 100M revenue on AI LaserWeeding platform. Hypersinc. https://hypersinc.com/article/carbon-robotics-crosses-100m-revenue-on-ai-laserweeding-platform-1775055784166

3. Design News. (2026). LaserWeeders killed billions of weeds in the last two years. Design News. https://www.designnews.com/automation/laserweeders-killed-billions-of-weeds-in-the-last-two-years

4. Rural Development Administration. (2026). AI-Based Image Diagnosis and Treatment Service for Crop Diseases and Pests. Rural Development Administration. https://www.rda.go.kr/board/board.do?mode=view&prgId=day_farmprmninfoEntry&dataNo=100000806197

5. Smart Farm Korea. (2026). Artificial Intelligence-Based Crop Growth Management Prescription Service. Smart Farm Korea. https://www.smartfarmkorea.net/analyService/internal/aiAdvice.do?menuId=M11030102

1.2

Fruit Maturity Assessment and Growth and Yield Prediction

In a large commercial greenhouse, a trial mounted a camera on a pipe-rail trolley, counted bell peppers at each stage of maturity, and examined how those counts related to total yield. The trolley moved at 0.3 to 0.5 meters per second along rails suspended near the greenhouse ceiling. The camera hanging beneath it faced rows of bell peppers, read the surface colors of fruits visible between the leaves, and converted the degree of red ripeness into numbers. The trolley-mounted counting system maintained stable detection performance throughout the trial.

Knowing when a fruit is ripe carries more money than it may seem. Picked too early, fruit is sold before its sugar content has risen. Picked too late, it softens or spoils and is discarded during distribution. For many years, this judgment was left to workers' eyes and hands. They walked through fields or greenhouses looking at colors and touching fruit, and because standards differed slightly from person to person, judgments of the same fruit often conflicted. An exact reading of sugar content, acidity, or firmness required chemical analysis: cutting open the fruit, squeezing out juice, and placing it in an instrument.

That method destroys the fruit itself. Cutting open a few samples cannot reveal conditions across an entire field, so farms often spent labor and time only to end up with an answer little better than an estimate. A wrong judgment does not end with one mistake. If an underripe fruit is missed during the first harvest, it is often no longer marketable by the next and is thrown away.

A single camera is actually being asked more than one question. Is this fruit fully ripe today? Is the whole tree now flowering, or are its fruits enlarging? How many kilograms will this field produce in the future? These are different questions. Maturity is judged from an individual fruit, the growth stage from the course of the whole tree or field, and yield by counting both and converting them into a future number. Yet all three begin with images captured through the same lens. Recent research is therefore moving away from answering each question separately and toward handling all three at once with a single camera and one pipeline.

There is a reason a camera can estimate ripeness from color alone. As fruit ripens, chlorophyll in the skin breaks down and the green fades. Pigments such as carotenoids and anthocyanins take its place, coloring the fruit yellow, orange, or red. Invisible changes take place inside at the same time. Total soluble solids (TSS), the sugars dissolved in water, increase; titratable acidity (TA), which produces sourness, falls; and the proportions of starch and water change.

An RGB camera records the visible spectrum, the same 400-to-700-nanometer band seen by the human eye, and serves as the basic tool for reading changes in external appearance: hue, texture, and shape.

An RGB-D camera adds distance information that color alone cannot provide. An ordinary camera captures flat color information, while an RGB-D camera attaches one additional distance value to every pixel. It resembles a map that once showed only latitude and longitude and now includes elevation. With this depth value, the camera calculates a fruit's front-back and left-right position, radius, and volume. It can also detect fruit half hidden behind leaves or branches. Leaving occlusion uncorrected produces an undercount, which pulls the entire yield forecast downward.

Devices such as the Intel RealSense D435i, commonly used by patrol robots in greenhouses, perform this role.

Collecting the data also demands considerable work. Camera images contain noise, which is removed with Gaussian and Wiener filtering. A technique called CLAHE evens out contrast between shaded and bright areas, followed by min-max normalization that brings pixel values into a fixed range. If there are too few images, the existing ones are rotated, flipped, brightened or darkened, or partially masked to increase their number. More recently, techniques such as CGAN and SMOTE have produced synthetic fruit images and even chemical composition measurements that were never actually captured, adding them to the training set.

These augmented images must not be mixed before the data are split into training and validation sets. Splitting first and altering only the training set prevents inflated evaluation results later.

Visible light reaches its limit with fruit whose skin never changes color. The Thai mango cultivar Nam Dok Mai Si Tong (NDMST) is one example. Its skin remains uniformly yellow throughout ripening, making it difficult to judge harvest time from photographs alone. Multispectral and hyperspectral cameras are used in such cases to measure near-infrared wavelengths above 700 nanometers, which the human eye cannot see. Near-infrared light responds sensitively to changes in moisture and chemical composition inside the flesh, revealing what is hidden beneath the skin. Researchers used this principle to develop a system called AFMG-DLFF.

It combines 1,536-dimensional external features extracted from RGB images with a 512-dimensional vector encoding internal conditions, including the BrimA index, calculated by subtracting acidity from sugar content, to form a 2,048-dimensional vector. K-means clustering establishes three stages, unripe, half-ripe, and fully ripe, while glowworm swarm optimization, inspired by the way groups of glowworms gather by following one another's brightness, tunes settings such as the learning rate and batch size. The resulting fusion model achieved grading accuracy of 97.86 percent.

Under the same conditions, a model using image information alone reached only 88.16 percent, while a model fusing only images reached 88.73 percent. Adding internal data sharply raised accuracy.

Ripeness can also be measured through heat rather than light. As a fruit ripens, the difference between its surface temperature and the surrounding air gradually narrows, and a thermal imaging sensor reads this minute difference directly. Oil palm plantations combine this principle with near-infrared measurements and color analysis to estimate oil and free fatty acid (FFA) levels inside a fruit bunch without cutting it open. A low-cost ESP32-CAM module hung beneath a tree counts naturally fallen loose fruit in real time and reports it to a web dashboard connected to a YOLOv8x recognition model.

The system has recorded an F1 score as high as 0.94, a balance of precision and recall, and helps prevent harvest timing from being missed.

The kind of eye a farm acquires depends on its budget and crop. A low-cost RGB camera is enough for a crop whose ripeness can be judged by color. A crop with many fruits hidden by leaves requires an RGB-D camera with depth information. Fruit such as mangoes, whose exterior and interior do not change together, needs more expensive equipment such as near-infrared or hyperspectral sensors. Crops that require readings of fine chemical components such as oil or fatty acids may add a thermal sensor as well. For now, the practical sequence appears to be starting with one inexpensive camera and adding sensors one at a time as the need arises.

Back at the trolley near the greenhouse ceiling, the camera's work does not end with taking a picture. AI draws a rectangular box around every fruit in the image or goes one step farther, tracing each fruit's outline pixel by pixel and cutting it out from the background. The problem is that when the trolley passes the same section again or the camera angle shifts even slightly, the same fruit may be counted twice or three times. Non-maximum suppression prevents this by keeping the one overlapping box with the highest confidence score and deleting the others.

An object-tracking method such as DeepSORT is then added to follow whether a fruit in the current frame is the same individual seen in the previous frame. These two devices allow the trolley to complete one circuit of the greenhouse and report a fruit count with duplicates removed.

Cluster tomatoes require a different approach from individual tomatoes. One greenhouse mounted an Intel RealSense D435i camera on a patrol robot. The camera was fixed at a height of 1.2 meters, a vertical tilt of 30 degrees, and a horizontal rotation of 15 degrees, with the distance between lens and fruit set at 0.5 to 1 meter. It collected video at a resolution of 1280×720 and 30 frames per second.

As the robot moved along the rail at 0.3 to 0.5 meters per second, AI divided maturity into three stages according to the proportion of red fruit within a cluster: early, when 10 to 30 percent had turned red; middle, at 30 to 70 percent; and late, above 70 percent. The LampCT-YOLO model assigned to this task adds an attention mechanism called SegNeXt to a YOLOv10 backbone, tuning it to respond more sensitively to color information. It suppresses unnecessary background responses from greenhouse frames and soil floors, then uses LAMP channel pruning to reduce its size enough for real-time detection on the robot's compact computing device.

A closer look at this robot shows an edge-computing unit such as an NVIDIA Jetson AGX Orin or Xavier NX on its body, with a RealSense camera and a two-dimensional LiDAR sensor mounted at the front. LiDAR emits light and measures the time it takes to return, drawing a dense map of distances to nearby obstacles and helping the robot remain on greenhouse rails and aisles. Using the camera and LiDAR together, the robot does more than grade fruit maturity in real time. It extracts the three-dimensional coordinates of every ripe fruit.

Those coordinates are needed so that, in the next step, a robotic arm or a person can move to the exact location and pick it.

For a farm, using separate devices to inspect disease and classify maturity is cumbersome. AgroDualNet was developed to reduce that burden by placing both tasks in one system. Its disease branch combines CBAM and ResNet50 with the SMO optimization technique and achieved 99.6 percent accuracy on the PlantVillage dataset of leaf-disease images. Its maturity branch uses the much lighter MobileNetV2 and YOLOv8 to distinguish four stages, unripe, half-ripe, fully ripe, and softening, and runs in real time even on a small battery-powered device.

Receiving two judgments from one camera and one program means the farm does not need two sets of equipment.

Pear growers in Japan faced a different problem: explainable AI. Even if AI can assign grades, it is difficult to earn trust in the field if it cannot explain why it assigned them. Researchers added a global context block to a Mask R-CNN neural network. The block scans the whole image first and supplies advance hints about areas likely to contain defects, allowing the network to cut out blemishes and stains on the pear surface precisely at the pixel level. The proportion of the entire surface occupied by a defect is then divided into fuzzy degrees such as large, somewhat large, somewhat small, and small.

Those degrees enter a fuzzy inference system that imitates expert judgment rules and produces the four grades used in the real market: premium, superior, good, and off-grade. Because the system imitates the ambiguity of human judgment instead of answering with one rigid number, the grower can follow the evidence for the grade given to the pear.

Drones take over when the question is when to harvest an entire broad field. When a Mavic 2 Pro flies at about 7 meters and photographs a field where tassels, the male flowers at the top of maize plants, are growing, it needs only five minutes to cover one hectare. A tassel-detection model based on YOLOv5 sorts tassels into three stages, not yet flowering, beginning to flower, and fully flowering. It then calculates a growth index from the difference between the proportions of fully flowering and not-yet-flowering tassels among all tassels.

This index showed a strong negative correlation with the date of silk emergence, with a correlation coefficient as distinct as −0.943.

Beginning from the day maize silks emerge, the calculation shifts to adding daily temperatures. Regardless of whether each day is warm or cool, the maize must accumulate a set amount of heat before moving to the next stage. The 恵味Star cultivar begins harvest when accumulated temperature reaches 500 degrees Celsius and finishes at 560 degrees, while Honey Bantam 20 begins at 450 degrees and ends at 510 degrees. Each cultivar has its own standard. The resulting optimal harvest window lasts about five days, and predictions made this way one month before harvest proved correct about 70 percent of the time.

In a similar study predicting the date of first bloom in apple trees, a model called BiLSTM-MhA combined daily temperature records with elevation, latitude, and longitude. It looked 15 to 20 days ahead while limiting error to just over one day, or 1.34 days precisely.

Field validation followed to determine whether this forecast was actually useful in sweet-corn fields. Testing predictions across more than eighty fields confirmed the 70 percent accuracy once again. There was no statistically significant difference in yield between a one-pass harvest of the entire field on the predicted date and selective harvesting in which ears were picked across several days according to their individual ripeness (p-value 0.07). If the forecast is right, workers do not need to return repeatedly to select ears, and a machine can clear the whole field in one day without a yield penalty.

For such a forecasting system to have practical value, its results must reach the farm manager immediately. In sweet-corn fields, drone-image analysis connects automatically to a regional weather-data server, placing an expected harvest date on the screen without anyone having to perform a separate calculation. The web dashboard at an oil palm plantation works the same way. From an office, a manager sees in real time which trees have loose fruit accumulating beneath them and decides the order in which to send workers. Work once done by walking the ground is moved to a single screen.

This automatic connection is the final link that turns a laboratory model into a tool for an actual farm.

In a photograph of a strawberry field, small red fruits appear densely overlapped between the leaves. For fruits such as strawberries and peaches that are small and grow close together, separating individuals is difficult in itself. A recent model called FruitQuery used a transformer architecture to detect the maturity of strawberries and peaches growing in fields. It recorded mean average precision of 67.02 percent with only 14.08 million parameters, making it lighter and more accurate than YOLOv8, YOLOv9, and YOLOv10.

Strawberry-field images impose a heavy burden because people must mark the correct answer on each one by hand, and annotation itself becomes a training bottleneck. SSEFNet, released in 2026, reduced that burden through semi-supervised learning, training on a few labeled images together with unlabeled ones. It achieved 91.1 percent precision even on a complex, tangled strawberry dataset. Researchers at the University of Florida went farther and built a virtual strawberry field inside a computer, then trained the model there before taking it into a real field.

The resulting model found fruit in real fields with 92 percent accuracy and measured fruit diameter with an error of only 1.2 millimeters, a level suitable for direct use in commercial grading. As the cost of collecting and annotating image data falls, multisensory fusion that combines images with sensors measuring temperature, humidity, and ethylene gas concentration appears likely to become the next established stage.

Yield prediction looks easy, but it is not. Fruit visible in a photograph is only what could be seen at the moment the image was taken. Other fruit remains hidden beneath leaves, while new flowers may set fruit over the next several weeks. A sound forecasting model therefore looks beyond the number of fruit visible today and calculates with the tree's growth stage, accumulated temperature, and harvest records from the same period in past years. The camera has seen only one present scene. Yield forecasting adds time to that scene and draws the future.

Some systems connect counting, quality grading, and yield calculation in a single line. The FruitNet pipeline first uses YOLOv8 to find fruit in an image in real time and produce an initial count. It then combines a deep convolutional neural network with a support vector machine to divide the fruit into four quality grades. Finally, it places the counts and size information for each grade into a statistical method called random forest regression to calculate the field's final yield.

Tests on the MinneApple dataset, which contains more than 18,000 apple images and over 40,000 individual annotations, produced training accuracy above 98 percent and validation accuracy ranging from 94 to 95 percent. When fruit overlaps densely and becomes occluded, a method for resolving overlapping boxes and a regression module that examines several variables together prevent duplicate counting.

The same camera information that counts and grades fruit is used to plan the path a robotic arm will follow to pick it. After an RGB-D sensor such as Kinect or RealSense locates the fruit, AI combines three factors into one regression equation: a confidence score showing how reliable the judgment is, whether the gripper has two or three fingers, and how far the fruit is tilted from vertical. When this value is placed directly into the cost calculation of the A* pathfinding algorithm, the robotic arm no longer reaches blindly along the shortest route.

It plots a path that considers both a favorable gripping angle and fruit whose location is certain. In tests with oranges, the coefficient of determination was 0.89, positional error was 3.2 centimeters, and percentage error was 5.4 percent. In tests with strawberries, the coefficient of determination was 0.82, positional error was 4.7 centimeters, and percentage error was 7.9 percent. After this calculation was added, the distance traveled by the robotic arm to pick oranges fell by 14 percent and the distance to pick strawberries by 11 percent, while the precision of grasping the correct position rose by 12 to 15 percent.

Set out in this way, the direction becomes visible. Models that classified only maturity and those that counted only fruit were once separate. Now models are appearing in succession that combine disease and maturity, as AgroDualNet does; connect count, quality, and yield in one line, as FruitNet does; and place perception and motion planning in one body, as an RGB-D robotic arm does. The path ahead appears to be answering more questions with one camera and one program. It is a plausible direction because it reduces the burden on farms of buying and managing multiple pieces of equipment separately.

Instead of arranging several cameras, one camera performs several roles at once.

Several common measures show how well these differently constructed models perform. Accuracy is the proportion of all judgments that are correct. Precision is the proportion of fruit AI calls ripe that truly is ripe, while recall is the proportion of truly ripe fruit that AI finds without missing. The F1 score, which considers precision and recall together, is useful for showing how well a model catches small classes such as unripe or spoiled fruit. The accuracy of bounding boxes is measured by IoU, the proportion of overlap between a box drawn by AI and the true fruit location.

mAP averages performance across several confidence thresholds and summarizes the overall capability of a detection model. Devices that must operate in real time, such as robots and patrol cameras, are evaluated with FPS, the number of images processed per second. Predictions that answer with a number, such as yield or sugar content, use other measures. Mean absolute error and root mean square error show the distance between predicted and actual values, while mean absolute percentage error converts that distance into a percentage.

The coefficient of determination expresses, on a scale from 0 to 1, how much of the actual variation in yield or sugar content the model explains; values closer to 1 indicate more precise prediction. The freshness score gives the percentage of the total fruit surface that is free of defects. Papers therefore report several measures side by side instead of presenting accuracy alone, preventing the selection of a single metric that makes a model look favorable.

All these figures eventually appear in one table on a farm-office wall or tablet screen. It lists, for each zone, the number of fruits picked today, the proportion in each grade, and the number expected to ripen fully over the next week. The manager uses the table to decide how many people to send to each zone tomorrow and how many boxes to prepare. A handful of numbers changes a person's daily schedule, so a one-percentage-point difference in model accuracy is not a small matter in the field.

The picture does not end at harvest. It continues into logistics. Knowing in advance how many boxes will come in today allows the farm to decide a day ahead how many refrigerated trucks to call and how many workers to assign to the packing house. If the prediction is wrong and volume suddenly surges, boxes pile up in front of cold storage and some cannot be cooled in time, lowering their quality. Maturity classification and yield prediction are technologies inside the field, but they are also variables that reshape logistics planning beyond it.

What matters to a farm owner is consistent accuracy. A model that performs well only on a few images captured on a clear day but fails on cloudy days or near evening has no value in the field. A number that happens to look good once is entirely different from the same level appearing every morning. The fact that a trolley-mounted counting system circling a greenhouse each day maintained steady performance throughout a trial is as important as its peak accuracy figure.

Moving such a system unchanged into the field exposes problems that were invisible in the laboratory. Outdoors, the strength of sunlight changes by the hour, strong direct light creates deep shadows, and branches and leaves repeatedly overlap and separate in the wind. Dust on a lens or a film of moisture can sharply reduce recognition performance when the system depends on a single camera. Resisting such environmental changes will require multiview imaging from several angles at once, followed by three-dimensional fusion that combines the resulting images into one spatial point cloud.

Fruit whose exterior and interior tell different stories remains a major obstacle. In fruit such as mangoes, whose skin color does not change through ripening or whose skin is thick, photographs alone cannot reveal sugar content, acidity, or fatty acid levels. Research must advance both fusion methods that add a small number of chemical measurements to inexpensive RGB images and efforts to make near-infrared and hyperspectral sensors themselves smaller and cheaper. Without an eye that sees exterior and interior together, even sophisticated AI cannot read the whole story beneath the skin.

Equipment used on actual farms is not as well provisioned as a laboratory computer. High-performing vision transformers and large convolutional neural networks respond slowly on small robots and inexpensive IoT modules with limited batteries and graphics processors. Work therefore continues on lightweight techniques: quantization, which converts values to lower precision to reduce computation; structural pruning, such as LAMP channel pruning, which cuts unnecessary connections; and knowledge distillation, which transfers the judgments of a large model into a small one.

AI used in the field is often the model that remains practical within the available hardware.

The training method itself has a common trap. Some studies inflate the data by rotating, flipping, and changing the brightness of images, and only afterward split them into training and validation sets. This creates data leakage, because altered images of the same fruit enter both sets. The validation score is then inflated above the model's real ability. To prevent this, the training and validation sets must be separated first, data augmentation must be applied only to the training set, and the validation set must remain untouched in its original form.

Reported figures from studies that follow this principle cannot be compared directly with those from studies that do not.

Sources and References

Academic Papers

1. Bijwe, K. B., & Gadicha, A. B. (2026). FruitNet: a deep hybrid model for fruit detection and yield estimation. Bulletin of Electrical Engineering and Informatics, 15(2), 1150-1160. DOI: 10.11591/eei.v15i2.11112.
2. Kripa, S., & Jeyalakshmi, V. (2026). Automated fruit maturity grading using deep learning with feature fusion. Preprint submitted to Scientific Reports.
3. Moldvai, L., & Nyéki, A. (2025). Innovative computer vision methods for tomato (Solanum Lycopersicon) detection and cultivation: a review. Discover Applied Sciences, 7, 975. https://doi.org/10.1007/s42452-025-07613-x
4. Simatupang, et al. (2026). YOLOv8-based IoT system for oil palm harvest readiness identification through loose fruit detection. Journal of Deep Learning, Computer Vision and Digital Image Processing, 15(2), 237-249.
5. Zeeshan, S., et al. (2025). Data-driven trajectory optimization in robotic fruit harvesting via deep learning-based perception and actuator-aware motion planning. Engineering Research Express, 7, 045210.
6. Deep learning for ripeness grading of oil palm fresh fruit bunches: A comprehensive review. (2026). Trends in Sciences, 23(8), 13227.

7. Deep learning-based detection of strawberry fruit and ripeness in smart farming: First review of architectures and methodologies. (2026). Smart Agricultural Technology.

8. AgroDualNet: a dual deep learning-based crop disease forecasting and fruit ripening detection framework. (2026). Scientific Reports. https://www.nature.com/articles/s41598-026-49002-3

9. 大澤央, 齋藤正博, 八木岡敦. (2026). Development of a Technique for Predicting the Optimal Harvest Time of Sweet Corn Using Object-Detection AI. 『農業食料工学会誌』, 88(2), 85-91.

10. 大澤央, 齋藤正博, 八木岡敦. (2026). Development of an AI Technique for Predicting the Optimal Harvest Time of Processing Sweet Corn. 『農業食料工学会誌』, 88(3), 123-128.

11. 劉恩奇, 劉淼, 楊貴軍 et al. (2026). Bi-LSTM Prediction of the Apple First-Flowering Date Based on Enhanced Time-Series Temperature Features. 『智慧農業(中英文)』, 8(2), 84-96.

12. 趙麗成 et al. (2026). Development of a Cluster-Tomato Maturity Detection Algorithm Based on Improved YOLOv10 and LAMP Channel Pruning. 『智慧農業(中英文)』, 8(2), 133-146.

13. A Study of a Pear-Grading Method Using Deep Learning with Global Context Information and Fuzzy Inference. (2026). 『農業情報研究』, 35(2), 65-78.

14. Deep learning-based fruit counting and yield estimation for sweet pepper using a pipe-rail trolley-mounted camera system in a commercial greenhouse. (2026). Horticulturae, 11(7), 718. https://doi.org/10.3390/horticulturae11070718

15. FruitQuery: a transformer-based model for strawberry and peach ripeness detection in orchard conditions. (2026).

16. SSEFNet: a semi-supervised learning model for strawberry ripeness detection in complex field datasets. (2026). arXiv:2602.15656. https://arxiv.org/pdf/2602.15656

17. Virtual strawberry field simulation for training fruit detection and sizing models. (2026). University of Florida.

Lectures and Video Materials

1. CSUMB College of Business. (2025). 2025 Ag Jobs: Bringing a Greater Vision, Harvesting Data: Careers in Agriculture. YouTube. https://www.youtube.com/watch?v=rgshB8Jh2ZE

2. Street Science with Fun. (2025). AI/ML-Powered Mango Detection & Harvesting Machine. YouTube. https://www.youtube.com/watch?v=Dy3xDC8rElw

3. Washington State Academy of Sciences. (2025). AI for Post Harvest. YouTube. https://www.youtube.com/watch?v=hJlOrgNlQG4

4. Solvi. (2026). Webinar: Validating AI-Based Plant Counts with Real Field Data. YouTube. https://www.youtube.com/watch?v=B0KuJdRtmaQ

5. Greg Bronevetsky. (2025). Predictive Pattern Recognition of Plant Growth Traits in Simulated and Controlled Environments. YouTube. https://www.youtube.com/watch?v=uKepfZAnKbk

Chapter 2

Horticultural Robots and Autonomous Field Operations

2.1
Harvesting Robots and End-Effectors for Fruit and Vegetable Crops and Mushrooms

As of April 2026, a cherry tomato harvesting robot capable of picking 31 kilograms a day by itself had entered commercial production in Japan. That same month, Israeli agtech company DailyRobotics deployed an open-field strawberry harvester at a farm in California, while the robot-as-a-service snack tomato harvester made by Japanese startup inaho reduced farms' labor demand by more than 45 percent. Horticultural harvesting robots face three major tasks: seeing where fruit is, gauging how much force can be applied without damaging it, and immediately correcting a grip when the first attempt is wrong.

For crops such as tomatoes, strawberries, and button mushrooms that must be picked one by one by hand, harvesting accounts for anywhere from half to as much as 70 percent of total production costs. As rural workforces continue to shrink and age, expectations for robotic arms that can replace human hands grow each year. Yet across currently commercialized harvesting robots, the fruit-picking success rate remains around 66 percent, and roughly one in every twenty harvested fruits bears a surface wound or pressure mark.

People pick fruit by sight and gauge force through their fingertips, but robots must recreate those senses with sensors and computation. In robots that harvest fruits, vegetables, and mushrooms, therefore, the design of the hand that actually grasps the crop, the end-effector, and the way it controls force determine the capability of the entire robotic arm.

The shape of a harvesting robot's arm varies with the planting method and degree of occlusion. In trees thick with overlapping branches and leaves, an articulated arm with five or six joints has an advantage. Just as a person rotates the shoulder, elbow, and wrist in sequence to slip a hand through a narrow opening, a five- or six-jointed robotic arm finds gaps between leaves and reaches fruit hidden deep inside. By contrast, a parallel robotic arm is better suited to crops lined up along pipe rails in a greenhouse, where it need only move within a fixed cylindrical workspace.

Built from several short links that support one point together, a parallel arm is structurally rigid and quick to respond. Even in a confined space roughly 300 millimeters in radius and 300 millimeters high, it moves rapidly without wobbling and removes fruit by suction.

There is little room for broad vertical arm movement in mushroom growing rooms, where shelves are stacked in tiers no more than 250 millimeters apart. The SCARA, or Selective Compliance Assembly Robot Arm, architecture has become established in such settings. A SCARA arm rotates rapidly from side to side over a flat plane and rises or falls only slightly when needed, allowing it to slide into the narrow gap between shelves almost as if laid flat and pick mushrooms. One arm is not enough for a crop such as an apple tree, where fruit hangs in every direction.

The apple-harvesting robot known as MARS 3.1 moves two arms independently, reaching in simultaneously from both sides to divide up the fruit in the shallow zone just inside the tree's outer surface without the arms colliding.

There are three main ways to build the fingertips that actually touch the fruit, namely the end-effector. A rigid mechanical gripper whose three fingers close through a screw-driven actuator is good at pushing fruit toward the center and bearing its weight; if needed, a small blade or rotary cutter can be mounted at the fingertips to sever the stem immediately. Because its fingers are made of metal or hard plastic, however, force tends to concentrate at a single point on the fruit surface, easily leaving pressure marks or wounds.

A vacuum-suction mechanism replaces fingers with a suction cup pressed against the fruit surface, evacuates the air inside to create negative pressure, and then pulls or twists the fruit free. It works well on fruit with smooth surfaces or thick skins, but when a surface is bumpy or covered in fine hairs, air leaks through the gap between the cup and the surface and the suction quickly fails.

The flexible, compliant gripper was developed to remedy the weaknesses of these two approaches by making its fingers from soft materials such as silicone or hyperelastomers. It bends by controlling how much air fills or leaves multiple small chambers divided within each finger, or it uses a structure known as the

Fin Ray Effect. As in a fish fin, supports cross diagonally so that pressing one side causes the other side to curve naturally around the object. Whether the fruit is round or irregular, these fingers reshape themselves to match its surface.

Force therefore spreads broadly instead of concentrating at one fingertip, fundamentally reducing fruit damage. The softer the hand becomes, the closer the robot comes to the seemingly contradictory goal of gripping harder while injuring the fruit less.

No matter how soft the hand, it is useless without vision and touch. A robot must simultaneously know exactly where the fruit is, where to cut its stem, and how much force is currently acting on its fingertips. The eyes for localization include RGB-D cameras that capture color and depth together, such as the Intel RealSense D435i and D405 and Microsoft Kinect v2, along with stereo cameras and 3D laser scanners.

The system reads ripeness and texture in the color image, overlays depth data to form a point cloud, and then calculates the fruit's front-back, left-right, and vertical position, as well as its roundness and volume, in the robotic arm's coordinate system. When harvesting button mushrooms, the 3D scanner measures the cap's diameter and height, leaves mushrooms smaller than the specified shipping grade untouched, and selects only fully grown ones for picking.

Once vision has found the location, fingertip sensation comes next. A thin film pressure sensor attached inside the finger, a torque sensor that reads back the force on the motor, an ultrasonic distance sensor, a 23-bit encoder for precise position readings, and sensors that measure pneumatic and hydraulic pressure operate as a set. The pressure sensor monitors in real time whether the force applied to the fruit exceeds the crop's tolerance, while the ultrasonic sensor catches the instant the fruit slips slightly within the hand and immediately readjusts motor force.

When fruit is hidden within a canopy dense with leaves and branches, finger vision using an additional small camera at the fingertip works alongside view-replanning that repositions the arm and captures another image. If the view becomes obstructed during approach, the robot immediately changes angle and finds a clear opening again.

Deep-learning detection models serve as the eyes that locate the fruit and determine where to cut. Recent members of the YOLO, or You Only Look Once, family, including YOLOv8, YOLOv10, and YOLOv11, are widely used. LampCT-YOLO, refined specifically for tomatoes, uses YOLOv10 as its backbone and adds an attention technique called SegNeXt. This method accompanies segmentation, the process of tracing and cutting out an object's boundary pixel by pixel, and gives greater weight to color cues that indicate tomato ripeness even amid the cluttered background of intermingled greenhouse leaves.

LAMP pruning is added to reduce computation, allowing the model to run directly on a small field controller without an expensive server. To find the precise cutting point, the Swift Transformer V2 model is combined with traditional morphological image processing, which refines pixel regions by shape, to segment the fruit, calyx, and peduncle separately.

To be useful in practice, a trained model must run not on a full-sized laptop but inside a small box mounted beside the arm. Models streamlined into ONNX or TensorRT format are transferred to field controllers such as the NVIDIA Jetson Xavier NX and AGX Orin, Raspberry Pi 5, or an Omron PLC. The time such a robot takes to recognize and pick one fruit, its harvesting cycle, varies by crop and hand design. A SCARA arm takes less than 10 seconds per button mushroom; MARS 3.1 takes 7 to 9 seconds for one arm to pick an apple; strawberries take about 4 to 11 seconds and oranges about 7 to 11 seconds.

Several numbers are used to gauge a robot's capability. How accurately the camera finds fruit and peduncles is expressed as mean Average Precision, or mAP, while the difference between the detected and actual positions is measured as an error in centimeters. Yet the numbers that matter most to people working in the field are different: the proportion of attempts that actually succeed in picking fruit, the proportion of harvested fruit with surface damage, the enclosure ratio showing how broadly the fingers wrap around the fruit surface, and whether the grip stayed within the crop-specific safe force range.

Safe gripping force is set at 8 to 12 newtons for strawberries, close to 80 newtons for oranges, and 6 to 8 newtons for mandarins. The scale becomes easier to grasp by imagining the difference between lightly pressing an apple with a fingertip and firmly holding a child's hand. A preservation rate recording whether branches or peduncles were unnecessarily broken is measured as well.

Because a tomato has soft skin but a firm connection between stem and fruit, the hand must be configured somewhere between crushing it by pressure and failing to detach it by pulling. A flexible gripper with three fingers made from Dragon Skin 30 silicone and four air chambers wraps around the fruit at an air pressure of 28.293 kilopascals, a value found by a genetic algorithm. A genetic algorithm generates many trial combinations and repeatedly recombines the better performers to approach the best answer, eliminating the need for a person to adjust each value by hand.

Tomatoes held by this refined hand showed no sign of spoilage even after more than nine days, and in a dynamic picking test it succeeded on about 97 out of 100 fruits.

Another approach directly imitates human hand movements. Gripping a tomato with three fingers and twisting it through 360 degrees detaches only the fruit without damaging the peduncle, just like the unconscious wrist motion used to pick a tomato at the table. When the peduncle is cut instead, the Swift Transformer V2 model finds the exact cutting position. A patrol harvester equipped with LampCT-YOLO travels greenhouse rails all day, instantly classifying the ripeness of cluster tomatoes, which bear multiple fruits in a bunch, into three stages: green, orange, and red.

Strawberries have delicate flesh that bruises or softens under even slight force. Their design therefore focuses more than that of any other crop on reducing both contact area and force. A flexible silicone gripper using two rather than three fingers minimizes the surface touching the strawberry and applies only 8 to 12 newtons. Adding a path-planning method that combines regression analysis with A* search improves grasp-position accuracy by 15 percent and reduces arm travel by 11 percent.

Another method avoids touching the fruit body altogether: a camera locates the stem, scissor-shaped fingertips cut only that stem, then grasp it and carry the fruit to a collection box.

In a suction approach, a soft suction cup 15 millimeters deep is placed against the strawberry and generates 2.38 newtons of suction. The pressure actually applied to the fruit is only 20 to 31 percent of the limit that strawberry flesh can withstand, enabling damage-free harvesting. Such hand technology is already moving into real fields. Beginning in April 2026, Israeli agtech company DailyRobotics deployed its Q2 open-field strawberry harvester at farms in California.

Fitted with two robotic arms and soft grippers, the machine places strawberries into transparent packages known as clamshells immediately after picking; the company says it is two to three times faster than a person.

In button mushroom growing rooms, shelves are stacked in multiple tiers at intervals of no more than 250 millimeters, leaving no room for an arm to rise. A low-profile harvesting robot solves this with a SCARA architecture that combines one vertical axis with two rotating horizontal joints and slides between the shelves in a low, nearly horizontal posture. Its vertical axis carries a precision reduction device called a strain-wave gear, a belt, and a counterweight together, nearly eliminating jolting and positioning the arm within an error of 1 millimeter.

A 3D scanner first measures the mushroom cap diameter, and only mushrooms meeting the specified shipping grade are extracted with a suction cup or flexible gripper. The entire process takes less than 10 seconds per mushroom.

Because an apple tree bears fruit in multiple layers from front to back, a single arm cannot reach deeply concealed apples. MARS 3.1 moves two arms independently and reaches in from both sides at once to the shallow zone just inside the tree surface. Each arm handles one fruit in 7 to 9 seconds while keeping the proportion of bruised fruit below 1 percent. While one arm picks fruit on a front branch, the other drives directly toward fruit deeper inside to save time. Tests that combined a Fin Ray Effect fingertip structure, motor-torque feedback, and ultrasonic slip detection reported damage rates as low as 0 percent.

Oranges have thick, tough skins and can tolerate ample force. A three-finger silicone gripper designed to apply as much as 80 newtons, combined with A* path search, reduces arm travel by 14 percent and improves grasp-position accuracy by 12 percent. Cucumbers, by contrast, have bumpy and strongly curved surfaces, making it difficult for a single suction cup to seal without gaps.

A pleated suction cup based on paper folding, or origami, uses a concavely curved structure to fold and unfold itself to match the cucumber surface.

By expanding the suction-contact radius to 10.66 millimeters, it achieved an in-field harvesting success rate of 86.2 percent. Because skin thickness and curvature differ from fruit to fruit, it is still too early to expect a single end-effector to harvest every crop.

These fingertip technologies are leaving the laboratory bench and entering the daily routine of actual farms. In Japan, after nearly three years of field modification and refinement, a cherry tomato robot was independently harvesting 31 kilograms per day as of April 2026. Cases in which a fruit and vegetable harvesting robot is permanently installed and continuously operated on a commercial farm remain rare worldwide.

The snack tomato harvesting robot developed by Japanese startup inaho adopted a robot-as-a-service rental model rather than outright purchase, allowing farms to reduce labor by more than 45 percent without a large capital outlay. Only five farms piloted such rental robots in 2026, but in the same year TTA-ISO introduced a fully automated tomato harvesting robot requiring almost no human handling and was preparing for its first delivery. In strawberries, a publicly available dataset collecting field measurements of fruit position and orientation has enabled multiple research teams to compare their robots by the same standard.

Such cases, in which one machine independently fills a set daily quota, suggest that fruit and vegetable harvesting robots have moved beyond experimental operation and into actual calculations of revenue and labor costs.

Despite this level of refinement in fingertip technology, several barriers remain before robots can fully replace human hands. In an irregular canopy where leaves, branches, and fruit overlap in layers, a single camera struggles to find every fruit center and stem position. If part of the view is blocked, the arm may collide with the wrong object or simply pass unseen fruit. Further refinement of techniques that merge point clouds from multiple viewpoints, and of active imaging that autonomously changes angle according to the degree of occlusion, is needed to lower this barrier.

Soft fingers carry challenges of their own. Flexible silicone grippers excel at preventing fruit damage, but they respond slowly while filling and releasing air, have limited strength for lifting heavy fruit, and stretch or crack with prolonged use. Variable stiffness, in which the hand normally wraps softly around an object but particles cluster and make it momentarily rigid only during gripping, is seen as a promising answer. Research at multiple institutions is transferring directly into the fingers the principle by which a bag filled with coffee grounds becomes rigid when its air is evacuated.

How companies resolve this contradiction is likely to become a yardstick separating harvesting-robot leaders over the next several years.

Cost is another obstacle. A harvesting robot equipped with precision reducers, 3D sensors, and multiple motors costs more than a small farm can readily afford. In the dust, vibration, and humidity of fields and greenhouses, moreover, sensors soil easily and components fail frequently. Standardized modules that can be replaced as complete units, together with maintenance systems that permit immediate field repairs, will be needed before the technology can spread widely.

The final challenge is making the robot's reasons visible to people. Beyond combining images, depth information, and fingertip pressure signals inside the controller without delay, the system must show growers on a screen why the robot gripped that location and applied that amount of force. This technology is called Explainable AI, or XAI. Farmers need such explanations before they can trust a robot with an entire field. No matter how good its force control and visual recognition become, without a window through which people can inspect its decision process, the robot remains an unfamiliar worker.

Placing the six crops side by side reveals a clear divide in end-effector design. For crops such as strawberries and tomatoes, whose skins are delicate and whose market value is high, gentle end-effectors based on air chambers and suction have been developed first. For oranges, whose thick skins withstand substantial force, stronger end-effectors followed. Cucumbers are common but have irregular surfaces that make suction itself difficult, so separate solutions such as origami structures emerged later.

A task people naturally distinguish through fingertip sensation becomes a completely different blueprint for every crop when entrusted to a robot.

A parallel robotic arm hovers low above a safflower field and darts back and forth, while each pass of its fingertips makes a handful of threadlike florets disappear. Even with the ability to locate a crop visually and gauge force at its fingertips, the arm wastes time retracing its path if it chooses the wrong harvesting order. Deciding that order is the robot's brain, and the problem is most conspicuous in safflower fields. One safflower head contains a dense cluster of twenty-five to thirty-one slender florets. The robot cannot simply follow whichever one comes into view; it must first plot the shortest complete route.

Borrowing the shortest-route formulation known as the Traveling Salesman Problem, or TSP, a pointer network trained through reinforcement learning receives the floret positions and immediately rearranges the visiting order by weighting nearby targets more heavily. When a parallel robotic arm follows this sequence and gathers the florets with a suction-equipped end-effector, its travel distance falls by 11.5 to 20.17 percent compared with visiting them in the order they appear.

The pathfinding improvements described earlier for strawberries and oranges arise from similar computation. When deciding where the arm should move next, the robot considers the camera's confidence in the fruit it has just seen, the type of gripper at the fingertip, and how oblique the approach angle is. Feeding these values into an equation prepared through regression analysis estimates the cost of the next step, and A* pathfinding then draws the actual route from that estimate.

It detours around low-confidence directions, avoids angles likely to damage the end-effector, and bends the path in advance so that it approaches the fruit at an angle as close as possible to a right angle. Merely translating into computation the intuition people acquire by eye markedly reduces the time the arm spends wandering.

Before the camera acquires that confidence, it undergoes practice. If trained on photographs exactly as captured, a robot recognizes only frontal views taken on clear days. The training images are therefore deliberately disturbed. They are rotated, brightened or darkened, and mixed with fine noise so that the model experiences in advance the shadows between leaves and the dim light of cloudy days. After several such processed images pass through the learning method called Adam, the robot is less easily confused when it encounters unfamiliar angles and light in the field.

The capability of the resulting brain and end-effector is not fully described by visual accuracy alone. Performance is calculated under both loose and strict thresholds to see how consistently the camera is correct, while coefficient of determination and error rate separately measure how closely path predictions from regression analysis match actual paths. Tests also measure how long a small field controller takes to answer after seeing one image and how many images it processes per second. Percentage reductions in path length and time must be recorded alongside these figures to compare the performance of different robots.

Numbers reported in papers and numbers felt on a farm belong to different layers. A several-percentage-point increase in enclosure ratio or a reduction of ten-some percent in path length shows only the improvement over a previous method under identical conditions. A test plot is a place where row spacing and lighting have been arranged in advance, but a real farm is far less compliant. Figures for kilograms harvested per day and percentage reductions in labor emerge only after a robot has stood on a real farm and operated for many days. Few robots reported to date present both layers of numbers together.

There is a reason the strawberry dataset includes orientation as well as position. Knowing only where a fruit is does not suffice; the robot must also know which way its stem points so the end-effector can approach at a straight angle and pick it without damaging the peduncle. Publishing such pose information collected from multiple fields saves new research teams the effort of photographing a dataset from scratch and allows them to begin testing end-effector designs immediately.

The emergence of rental rather than purchased harvesting robots is tied to these circumstances. It is an attempt to relieve the burden of buying an entire robot carrying precision reducers, a 3D scanner, and a computing unit. Yet conditions change each time a robot is moved to a new field. Because greenhouse structures, growing methods, and planting intervals differ from one farm to another, the camera must be recalibrated and fingertip force values measured anew each time. The ability to pick one fruit without damage and the ability to transplant that capability unchanged to another field are separate challenges.

From the farm owner's perspective, the calculation is plain. A farm can rent and run one robot for several days during a season when workers are difficult to find, then increase the number next season if the damage rate and daily harvest are acceptable. The opportunity to test the machine first in one's own field before buying it is the greatest difference from the former need to make a one-time decision on an expensive machine.

Even these painstakingly trained models recognize only one crop at a time. Vision that distinguishes tomato ripeness uncannily well must learn from the beginning when taken to a strawberry field. An end-effector designed not to injure delicate strawberry flesh is likely to drop heavy fruit if mounted on an apple harvester. People naturally adjust sight and touch as they move between fields, but a robot must retrain its eyes and reshape its fingertips whenever the crop changes.

To become more like a genuine worker capable of moving among crops, it needs technology that unites crop-specific eyes and hands while allowing them to perform immediately in different fields.

If refining vision and fingertips was the challenge of the past several years, fitting the robot with a brain that decides its own visiting order, measuring the results by a common standard, and transferring skills learned in one crop and field to another are challenges that have only begun. Field trials over the next several years will determine whether computation validated in safflower fields also works in strawberry greenhouses or between apple trees. Fingertips, vision, and brains are now being validated separately in different crops and fields; they have not yet been integrated within a single robot.

Sources and References

Academic Papers

1. Feng, K. (2026). Structural Forms and Crop Adaptation Mechanisms of End-Effectors for Fruit and Vegetable Harvesting. Automation and Machine Learning, 7(2), 60-67. DOI: 10.23977/autml.2026.070207.

2. Zeeshan, S., Malik, M. A. I., et al. (2025). Data-driven trajectory optimization in robotic fruit harvesting via deep learning-based perception, gripper configuration, and fruit morphometrics. Engineering Research Express, 7, 045210.

3. Minorowicz, B., Patalas, A., & Łuczak, A. (2026). Design of a SCARA Robotic Arm for Mushroom Harvesting. Acta Mechanica et Automatica, 20(2), 366-378. DOI: 10.65731/ama/2026-0037.

4. Li, M., Wang, X., Liu, Z., Duan, M., & Jin, Z. (2026). DEMA-3D TSP: An Enhanced Reinforcement Learning with DEMA Attention in Sequence Optimization for Safflower Picking Robot. Smart Agriculture, 8(2), 200-219. DOI: 10.12133/j.smartag.SA202506004.

5. Zhao, L., Lu, X., Wu, Q., Ren, N., Zhou, L., Cheng, Y., Hu, A., & Qi, C. (2026). An Improved YOLOv10-Based Tomato Ripeness Detection Algorithm with LAMP Channel Pruning. Smart Agriculture, 8(2), 133-146. DOI: 10.12133/j.smartag.SA202507045.

6. Moldvai, L., & Nyéki, A. (2025). Innovative computer vision methods for tomato (Solanum Lycopersicon) detection and cultivation: a review. Discover Applied Sciences, 7, 975. https://doi.org/10.1007/s42452-025-07613-x

Lectures and Video Materials

1. World AgriFood Innovation Conference (WAFI). (2026). WAFI2026 Global Agrifood Innovation Roadshow Launch Event & Inaugural Roadshow. YouTube. https://www.youtube.com/watch?v=U5CReIR_Rrg

2. World AgriFood Innovation Conference (WAFI). (2026). Australia Agrifood Innovation | WAFI2026 Global Agrifood Innovation Roadshow. YouTube. https://www.youtube.com/watch?v=ITnVRq1Tioo

3. C-DAC. (2025). AI in Agriculture: Transforming India's Krishi Landscape. YouTube. https://www.youtube.com/watch?v=lV3_VnWiBpM

4. Street Science with Fun. (2025). AI/ML-Powered Mango Detection & Harvesting Machine. YouTube. https://www.youtube.com/watch?v=Dy3xDC8rElw

5. Department of ECE, MGR Educational and Research Institute. (2025). Expert Talk on An Overview - Role of Computer Vision in Autonomous Navigation for Smart Agriculture. YouTube. https://www.youtube.com/watch?v=GVZ46DY2Tao

6. UC Agriculture and Natural Resources. (2025). 2026 Farm Robotics Challenge - IP 101. YouTube. https://www.youtube.com/watch?v=JA1greBXSoA

7. HortiQD. (2026). HortiQD 4th Webinar: The Data Bridge: Connecting Farmers and Robots for Smarter Agriculture. YouTube. https://www.youtube.com/watch?v=7W8NbMdhuss

8. Farm Foundation. (2025). Farm Foundation Forum: Automation for a Resilient Future for Agriculture. YouTube. https://www.youtube.com/watch?v=aWWqVnNkhoQ

Articles and Online Materials

1. AgFunderNews. (2026). DailyRobotics gears up for commercial launch in California in 2026 with robotic strawberry harvester. AgFunderNews. https://agfundernews.com/dailyrobotics-gears-up-for-commercial-launch-in-california-in-2026-with-robotic-strawberry-harvester

2. HortiDaily. (2026). Japanese agri-tech startup puts cherry tomato harvesting robot into routine production use. HortiDaily. https://www.hortidaily.com/article/9842754/japanese-agri-tech-startup-puts-cherry-tomato-harvesting-robot-into-routine-production-use/

3. HortiDaily. (2026). Harvest robot snack tomato commercial ready. HortiDaily. https://www.hortidaily.com/article/9778303/harvest-robot-snack-tomato-commercial-ready/

4. Future Farming. (2026). TTA-ISO introduces fully automated tomato harvesting robot. Future Farming. https://www.futurefarming.com/tech-in-focus/field-robots/tta-iso-introduces-fully-automated-tomato-harvesting-robot/

5. arXiv. (2026). [Field-Measured 6D Pose Dataset and Benchmark Model Preprint for Strawberry Harvesting Robots]. arXiv:2606.11381. https://arxiv.org/pdf/2606.11381

2.2

Route Optimization and Precision Control for Autonomous Robots and Drone Swarms

Hylio received approval from the U.S. Federal Aviation Administration for a single operator to supervise a swarm of three drones at once. Instead of a joystick for directly moving the drones, the operator holds a screen that displays the status of multiple aircraft together. The company's HYL-150 ARES carries about 50 kilograms and sprays more than 20 hectares per hour, while the XAG P150 has AI-based sensing that autonomously detects obstacles.

Behind such approved operations, in which multiple drones divide areas among themselves and coordinate direction and speed, lies a process in which machines calculate routes too complex for people to plan one by one.

This presents a problem distinct from making the fingers at the end of a robotic arm grasp fruit accurately. No matter how capable its hand, if a robot cannot decide where to start and in what order to move, it will pass over the same row twice or wander first to a distant point until its battery runs out. Fruit, weeds, and florets scattered through a field are not arranged in neat lines. A fruit hidden by a branch may suddenly appear when the robot shifts sideways, and a fruit visible moments earlier may be concealed again when leaves move in the wind.

The problem of finding the shortest route among twenty houses a mail carrier must visit in one day is called the Traveling Salesman Problem. Farm robots and drones must solve it not on a flat plane but in three-dimensional space that includes height. This is therefore known as the three-dimensional Traveling Salesman Problem, abbreviated 3D-TSP. Once there are more than twenty-five or thirty target points, the number of possible routes explodes, and long-used search methods such as genetic algorithms often take too long to calculate or stop at an answer that is not truly optimal.

Attempts to solve this problem existed before reinforcement learning. Ant colony algorithms modeled the way ants follow one another's scent trails while searching for food, while particle swarm algorithms imitated flocks of birds flying in formation. Borrowing rules from insect or bird group behavior gave these methods a certain plausibility, but in a real field where targets multiply and obstacles appear from moment to moment, the robot could be left standing indefinitely while computation finished.

Reinforcement learning, which extracts its own rules through numerous trials rather than following rules fixed in advance, succeeded these methods largely because it is impractical to write new rules individually for conditions that differ by field and change by season.

A newer solution combines reinforcement learning with a pointer network. Reinforcement learning provides no answer sheet in advance; it assigns scores to the outcomes of the system's own trials and improves them gradually. A pointer network does not invent a new destination but selects, as if pointing a finger, one of the coordinates already supplied as input. Their combination can be likened to an actor and a critic.

A neural network called the actor plays the performer, proposing the order of coordinates to visit one by one, while a neural network called the critic plays the reviewer, scoring in advance how efficient the proposed route will be as a whole. The actor consults the critic's score and adjusts the next proposal little by little. Repeating this process thousands of times develops an instinct for choosing ever shorter routes.

Several mechanisms refine this judgment. Multi-head self-attention allows each coordinate to be compared simultaneously with every other coordinate for proximity and direction, while distance encoding calculates the actual Euclidean distances between coordinates in advance and supplies them to the network as reference data. Dynamic Exponential Moving Average, abbreviated DEMA, then gives greater weight to recent routes and gradually forgets older ones as it updates the decision. AC-RL-PtrNet, developed in China and installed on a harvesting robot for safflower, a member of the daisy family, implements this combination.

It reads the flow of input 3D coordinates through Long Short-Term Memory, or LSTM, and stabilizes learning by adding layer normalization and residual connections to the critic network. To reduce computation, structural pruning removes unnecessary connections from the LSTM gates and fully connected layers like thinning tree branches, reducing the model's size without greatly lowering decision accuracy.

A harvesting robot fitted with this model underwent field trials in an actual safflower field in Xinjiang, China. Its parallel architecture placed several arms side by side and divided one route into sections handled simultaneously by the arms, raising operating speed. A condition with twenty-five target points was tested in fields planted with the two varieties Yuhong No. 1 and Yunhong No. 5. When a genetic algorithm and this model were run side by side, route length fell by 9.56 percent and operating time by 5.43 percent.

With thirty-one targets in a more densely entangled condition, the difference widened further: route length fell by 20.17 percent and operating time by 29.70 percent. The more numerous and intertwined the targets, the larger the advantage of learned decisions over a human-defined sequence.

Route optimization does not end with deciding visit order. In land leveling and in designing travel routes themselves, a reinforcement-learning technique called Proximal Policy Optimization, or PPO, is used. PPO changes the policy only a little at a time instead of making a large change all at once, preventing learning from suddenly collapsing. It can plan routes that reduce both travel distance and the amount of soil that must be excavated and moved, allowing a robot to move less and use its battery longer.

Leveling the ground and traveling over it may seem like separate problems, but both are ultimately solved within the same reinforcement-learning framework.

Even the best route is useless if the robot does not know its own position precisely. An autonomous ground robot uses RTK-GPS, which corrects satellite-navigation error with a base-station signal, to reduce positional error below 2 centimeters. It maps surrounding terrain in three dimensions with lidar, which measures the return time of emitted laser light, and an RGB-D camera, which adds distance information to color.

Placing two stereo-vision cameras side by side derives depth from the disparity between two viewpoints, making it possible to transform the location of fruit or stems from camera coordinates into the coordinate system in which the robotic arm moves. Ultrasonic distance sensors prevent the robot from colliding with another obstacle or overturning when turning at the end of a row or when its view is blocked.

Many other small, labor-intensive sensors are packed throughout an autonomous robot's body. A sensor measuring the voltage of its 12-volt battery reports remaining power in real time, while a solenoid-valve controller serves as a gatekeeper that opens and closes a nozzle with a single electrical signal. An encoder on a permanent-magnet DC geared motor counts wheel rotations and works backward to calculate how far the robot actually moved, while sensors measuring pneumatic and water pressure ensure that pressure in the spraying system stays within its specified range.

However precisely a route has been planned, if any one of these small sensors fails to do its job, the calculated numbers cannot become real wheel and nozzle movements.

Drones in the sky fill gaps left by robots crawling on the ground. Vegetation indices derived from multispectral images taken by a drone over multiple days are also used to find diseased leaves, but here they serve a different purpose. Reading the same indicators across the entire field produces a prescription map showing, by zone, the probability of weed concentrations. This map is redrawn every few days in step with weed growth and kept current. With this map in hand, the drone moves to the stage of determining the actual route it will fly.

A dual-branch convolutional neural network that reads context jointly distinguishes the boundary where crops and weeds meet at the pixel level. Once the boundary is identified, a stage called AWTPO uses a modified Dijkstra algorithm to find, among candidate routes, the one that consumes less fuel, battery power, and carried chemical weight. Dijkstra's algorithm is an established pathfinding technique originally used to find the fastest transfer route among subway lines passing through multiple stations. Here, field coordinates replace stations and fuel and battery consumption replace transfer time.

The resulting system flies the drone along the selected route and opens the nozzle to release chemicals only while passing a weed patch.

A separate eye mounted in front of the nozzle measures the size of a weed patch in real time. When the weed-covered area exceeds 1.5 square meters, a variable-rate microsprayer called VRMDWN activates. It adjusts droplet size between 100 and 400 micrometers. Its flow rate changes at the same time between 0.12 and 0.35 liters per minute. Dense weed patches receive coarser droplets and more liquid; sparse ones receive finer droplets and less. The work does not end after spraying. New images are taken 7, 14, and 28 days afterward and overlaid on past recurrence records.

Using the resulting data, the PTWRP stage applies the reinforcement-learning method Q-learning to predict where weeds are likely to reemerge.

Predicting where weeds will return is less visible than control itself, but its practical value is considerable. Without a forecast, growers must walk the entire field again every few days after treatment and visually check whether weeds have returned. The prediction map produced by PTWRP narrows the areas requiring reinspection to a handful, greatly reducing the time needed for the next drone flight or human patrol. Weed control changes from a one-off task completed in a single pass into a cycle of repeated prediction and verification.

For drones and ground robots to find the correct route, positions measured in different ways by different sensors must first be aligned to one reference frame. Latitude and longitude from satellite navigation, robot-relative coordinates from lidar and cameras, and joint-based coordinates in which the robotic arm moves all have different origins and directions. Only after coordinate-frame transformations overlay them on a single field map will the weed location detected by a drone coincide with the position where a ground robot actually opens its nozzle.

If this alignment is even slightly wrong, a painstakingly calculated optimal route becomes a wasted trip along the wrong row.

The system was tested in a real wheat field. Drones repeatedly photographed a 12-hectare commercial wheat field for 90 days, from seed emergence until before heading, to run the system. The flight path generated by AWTPO reduced drone flight time to 17.5 minutes. Targeted spraying by VRMDWN lowered herbicide use to 0.68 liters per hectare, 32 percent less than conventional blanket spraying. Yield protected from weed competition increased by 18 to 22 percent.

Compared with applying the same quantity across the whole field, delivering precisely the required quantity only where needed appears to have produced clear differences in both herbicide use and yield.

A route shortened by a few minutes may sound like a small number, but it accumulates differently in a field. A robot's daily working time is bounded by battery capacity and daylight, so time saved on each pass becomes enough by day's end to cover several additional rows. The number of robots and drones one person can supervise in a day ultimately depends on this margin; shorter routes allow the same workforce to manage a larger farm. Reducing abrupt turns and unnecessary repeat travel over the same section also lowers collision risk in fields where people and robots work together.

The precise decisions of individual drones show their real power when extended to swarms moving as a group. One drone traversing a large field alone inevitably leaves areas uncovered before its battery runs out, but several drones dividing the area and flying simultaneously can survey several times as much land in the same period. By 2026, spraying drones were moving beyond specialized equipment used by a handful of farms to become mainstream machines combining high-capacity tanks with artificial intelligence.

The XAG P150 distributed by Pegasus Robotics uses the downdraft from four rotors to disperse chemicals evenly and carries AI sensing that autonomously detects obstacles from 1.5 to 90 meters away. Hylio's HYL-150 ARES is designed for swarm operations, carrying about 50 kilograms and treating more than 20 hectares per hour. Such group flight is possible because onboard computers let individual drones exchange signals while coordinating multispectral cameras, spray tanks, and satellite-navigation signals. Hylio received U.S.

Federal Aviation Administration approval for one operator to supervise a group of three drones simultaneously; under this model, one person can cover thousands of acres in a day. Such approval for drones that autonomously divide areas and coordinate direction and speed demonstrates that one-person supervision of multi-aircraft swarm operations already functions in the field.

People and robots are spending more time working together in the same field. If a person is in the same row while a drone flies low to spray, ultrasonic sensors and vision cameras detect the movement and slow the aircraft or change its route. In the current model, where one operator monitors several drones or robots at once, the status of multiple machines appears on one screen. Catching the signal when any one of them develops a problem becomes the supervisor's new task.

As machines increasingly choose their own routes, the human role appears to be shifting from the hand that directly controls them to the eye that oversees the whole operation.

Ground robots are often more widely used than drones in orchards where trees form dense vertical walls. The electric autonomous spraying robot OrcBOT reads its own position and tree form using RTK-GPS and stereo cameras, while a YOLOv5 neural network determines in real time whether leafy canopy is present. Based on this judgment, a central electronic control unit switches the left and right nozzles independently and sprays nothing in empty sections without trees. A further device applies 10 kilovolts to give the droplets an electrostatic charge.

By the same principle that makes hair cling to a sweater when it is removed in winter, charged droplets penetrate and adhere to the backs of leaves and the interiors of branches. A conventional sprayer tends to deposit chemical only on the windward face of a leaf and miss the far side, whereas electrostatically charged particles carry an additional attractive force that draws them to leaves regardless of orientation.

Results are available from an actual test of this robot in the orchard of Ankara University's Faculty of Agriculture in Türkiye. Sprayed droplets measured 150 to 167 micrometers in diameter and met ASAE S572.1, the fine-droplet standard established by the American Society of Agricultural Engineers. Canopy coverage, the proportion of leaf surface bearing chemical, averaged 55 percent and reached as high as 57 percent. Compared with a conventional tractor sprayer, chemical use fell by 49 to 72 percent, a difference confirmed as statistically unlikely to be due to chance.

Operating cost, which had been $19.70 per hour when a person drove and sprayed with a tractor, fell by nearly 95 percent to $0.76 per hour when the robot took over.

Weeds in a wheat field spread over hectares and pests and diseases in an orchard of individual trees require different spray patterns from the outset. VRMDWN selects and treats only small weed patches scattered across the field, leaving the spaces between them completely untreated. OrcBOT, by contrast, must envelop an entire tree, so its target is large and the spray must reach evenly even behind leaves. Weed control depends on a map that accurately identifies where targets are; tree treatment depends on how evenly chemicals can be delivered inside an already known tree.

Although both technologies look like autonomous spraying from the outside, the problems they must solve have different faces.

Autonomous technology is entering small plots as well as large farms. A small robot driven by a PIC microcontroller performs multiple tasks by swapping modules. With a 250-gram seed hopper, it carries out precision hill seeding, placing one seed at each specified interval. With a 500-milliliter general-purpose pump nozzle, it spot-sprays chemicals only where needed. With a thin mild-steel blade, it mechanically weeds between plants. A compact robot that performs all three tasks with one battery and one microcontroller is emerging as an option for small farms unable to afford large agricultural machinery.

Before entering the field, all these neural networks are pretrained on platforms such as PyTorch or TensorFlow with optimization methods such as Adam or AdamW. The trained model is transferred in ONNX or TensorRT format to a palm-sized device such as an NVIDIA Jetson Xavier NX, Jetson Nano, Raspberry Pi 5, or industrial PLC controller, taking on a lighter form. Solving an actual 3D Traveling Salesman Problem with twenty-five to thirty-one target points takes no more than tens of milliseconds, so there is almost no perceptible pause while the robot waits for computation.

Even when a drone sprays while flying at 5 meters per second, motion planning accounts in advance for the brief delay required to open and close the nozzle.

Several measures together determine whether the resulting system is genuinely useful. Route reduction relative to a conventional algorithm, operating-time reduction, the latency required for an edge device to calculate a route, and energy savings in battery power and fuel evaluate the movements of the robots and drones themselves. Reduced herbicide use per hectare, canopy coverage, droplet size, misclassification rate at crop-weed boundaries, protected yield, and recurrence-prediction accuracy, which measures how accurately the system forecasts weeds returning after treatment, show whether it produced a real agricultural benefit.

Several further numbers verify how well the camera distinguished weeds from crops. Intersection over Union, or IoU, measures the overlap between the weed area drawn by the model and the actual weed area, much like measuring the overlap when two stamps are pressed in the same place. mAP, which averages precision measured at multiple thresholds, resembles repeating that stamping under several different conditions and taking the average.

The F1 score combines precision, the proportion of items classified as weeds that really were weeds, with recall, the proportion of actual weeds that were successfully detected, thereby exposing judgments skewed too far to one side. Excellence in only one of these three layers is not enough for a trusted field system. The three wheels of routing, treatment, and recognition must turn together.

Despite these achievements, many hurdles remain. Most results to date are limited by having been obtained in a single specific field. There is no guarantee that the sequencing instincts learned in a Xinjiang safflower field will transfer unchanged to an apple orchard with entirely different tree spacing and terrain. Because changes in crop height, planting interval, and ground slope alter the shape of the shortest route itself, results from a wider range of crops and terrain must accumulate before route-optimization models can move from tools for one field to tools used broadly across many fields.

Current 3D Traveling Salesman algorithms and flight-path calculations generally assume flat or orderly fields and often fail unchanged in steep terrain or areas crowded with obstacles. When wind exceeds 4.5 meters per second, or on days with fog, rain, or heavy cloud, multispectral image quality deteriorates and drone flight positions can deviate by around 1 meter, leading to misidentified weeds or chemical sprayed on the wrong row. The current practice of choosing clear, windless days risks missing the treatment window itself when bad weather persists.

On such days, a person must still go out, inspect the field visually, and reset the timing of control.

The physical limitations of drones and ground robots are another constraint. The drone used in testing could fly for only about 22 minutes and carry only around 1.2 liters of chemical, so uninterrupted coverage of a large field frequently requires time to replace batteries and refill the tank. During the few minutes in which a person swaps the battery and refills the chemical tank by hand, the drone cannot fly and must sit at the side of the field.

Mobile ground-based charging equipment that travels to replace batteries, and heterogeneous swarm-control technology in which drones and ground robots divide different roles and cooperate, are emerging as the next tasks. Once the cost of charging infrastructure is included, the initial burden of adopting the technology is substantial, especially for larger farms. Only a system in which aerial and ground robots divide the work each does best and automatically bridge the handoff between them will be transferable beyond current test-field scales to far larger farms.

The need to process complex computations in the middle of a field is another obstacle. Reinforcement-learning route planning and vision models that handle multiple tasks at once demand heavy computation and may respond slowly on small battery-powered devices. Quantization and pruning to reduce computation, and dedicated processors such as NPUs and FPGAs, continue to be studied as remedies. The balance between computing resources and accuracy will inevitably differ from farm to farm according to field size and crop value.

The final challenge is trust. Growers still lack a satisfactory way to see why a reinforcement-learning model chose a route and decided to spray only a particular location. If Explainable AI methods such as Grad-CAM or SHAP, which visualize the basis of a decision, are embedded in the system, growers can decide for themselves whether to accept the robot's choice or review it. If treatment records and weed-regrowth locations are accumulated year after year and fed back as knowledge, the system may even track the gradual shifts in soil and weed ecology over several years.

Only after such trust accumulates can farming move from today's semiautonomous model, in which a person gives the final approval, to an autonomous model entirely free of human hands.

Several further tasks must be solved before this technology spreads to large farms. Growers still need a way to see the basis for routes and treatment decisions selected by reinforcement learning, and work interruptions each time batteries or chemicals run out must be reduced. In swarm flight, where multiple drones exchange signals and divide territory, a signal loss can cause aircraft to lose one another's positions, overlap routes, or collide.

The effectiveness of today's supervisory model, in which one operator watches multiple aircraft displays at once, depends on how quickly a disconnected aircraft can be identified and brought to a safe stop.

Sources and References

Academic Papers

1. Adewumi, A., & Saraswat, D. (2026). Comprehensive Review of Artificial Intelligence and Edge Computing for Precision Weed Control. Artificial Intelligence and Applications. https://doi.org/10.47852/bonviewAIA62028867
2. Choudhari, A., Katkar, R., Gupta, M., Wathe, N., Kakde, S., & Poyam, P. (2025). Design of an Integrated Spatiotemporal Deep Learning Framework for Autonomous Precision Weed Detection, Treatment, and Recurrence Prediction in Drone-Based Smart Farming Sets. International Journal of Applied Mathematics, 38(2s), 1339-1366.
3. CropLife. (2026). 7 Agricultural Spray Drones to Watch in 2026. CropLife. https://www.croplife.com/iron/sprayers/7-agricultural-spray-drones-to-watch-in-2026/
4. Gerdan Koc, D., Koc, C., & Comakli, M. (2025). AI-powered autonomous spraying robot for precision orchard applications. Journal of Agricultural Engineering, 56(1) / 2026 LVII:1766. https://doi.org/10.4081/jae.2025.1766
5. Gundi, A., Kulagod, A., Sunny K, H., & Parashuram, B. (2026). Autonomous Agricultural Robot for Seeding, Spraying, and Weeding. International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 15(5), 1308-1311. https://doi.org/10.17148/IJARCCE.2026.155170
6. Li, M., Wang, X., Liu, Z., Duan, M., & Jin, Z. (2026). DEMA-3D TSP: An Enhanced Reinforcement Learning with DEMA Attention in Sequence Optimization for Safflower Picking Robot. Smart Agriculture, 8(2), 200-219. https://doi.org/10.12133/j.smartag.SA202506004
7. Logeshwaran, J., et al. (2026). Embedded Drone Swarm for Precision Agriculture with Vision AI. International Journal of Computer Applications, 187(115), 42-43.
8. The Robot Report. (2026). Drone swarms: How they actually work and what industries should care. The Robot Report. https://www.therobotreport.com/drone-swarms-how-they-actually-work-and-what-industries-should-care/

Lectures and Video Materials

1. Food and Agriculture Organization (FAO). (2025). Travel Around the World with Digital FAO and Agro-Informatics - Sydney. YouTube. https://www.youtube.com/watch?v=tbFEk73816U
2. PhD Navigator. (2025). Sampling Techniques in ICT Research | Energy Optimised Field Robot for Agriculture. YouTube. https://www.youtube.com/watch?v=jJerj_N7ALQ
3. World AgriFood Innovation Conference (WAFI). (2026). WAFI2026 Global Agrifood Innovation Roadshow Launch Event & Inaugural Roadshow. YouTube. https://www.youtube.com/watch?v=U5CReIR_Rrg
4. 茶柱. (2026). 20260120 Smart Agriculture in Tea Plantations: Kawasaki Giken, Ochiai Cutlery, and Terada Seisakusho. YouTube. https://www.youtube.com/watch?v=zmFbOC5RcJU
5. 臉書李載榮. (2025). A88 Private Institute for Information Industry Satellite Precision Agriculture AI Adoption Workshop, Taoyuan: AI Satellite Multispectral Imagery and Drone Applications. YouTube. https://www.youtube.com/watch?v=dtT3bafdKgw

6. Trimble Geospatial. (2026). eCognition Webinar: Deep Learning, UAVs and Precision Agriculture. YouTube. https://www.youtube.com/watch?v=alrseDYV8Lo

7. Society of Precision Agriculture Australia SPAA. (2025). Grain Automate National Webinar: Paving the way to on-farm autonomy. YouTube. https://www.youtube.com/watch?v=nZnsWgvs6-g

8. Agri Coaching Chandigarh. (2026). Digital Agriculture Mission 2026 | AI, Drones, Agri Stack & Smart Farming | Current Affairs. YouTube. https://www.youtube.com/watch?v=aI2rsTZBwMA

Chapter 3

Smart Greenhouses and Controlled Environment Agriculture (CEA)

3.1 IoT-Based Multisensor Monitoring and Intelligent Environmental Measurement

More than 125 million low-power LoRaWAN Internet of Things devices have been deployed across 70 countries, and the related market is estimated to be worth $14.25 billion in 2026. Researchers at the University of Nebraska in the United States reported that real-time monitoring of soil moisture alone reduced water use by as much as 35 percent. Greenhouses require equally dense measurement coverage. In a plastic greenhouse, the temperature at floor level can differ from that at eye level by three or four degrees, while a separate layer of hot air forms near the roof.

Inside a palm-sized white box attached to one of the greenhouse posts, temperature and humidity readings are refreshed every few seconds. Those readings determine whether the vents will open that night, whether the heater will run, and whether the irrigation system will switch on.

Air temperature is not the only variable that must be managed inside a greenhouse. Air temperature and relative humidity, carbon dioxide concentration, and the amount of sunlight (the measure of light in the wavelengths actually used for photosynthesis is called photosynthetic photon flux density) all change together, as do the electrical conductivity and hydrogen ion concentration of the nutrient solution in hydroponics, the moisture remaining in soil or substrate, the temperature of the water tank, and even ethylene gas in a storage facility. These variables do not behave independently.

Opening a vent to lower the temperature releases humidity and carbon dioxide at the same time, while turning on a heater raises more than the temperature: it also causes relative humidity to plunge. Supplying more water raises substrate moisture, but it can also increase the humidity of the air around the roots and help create conditions favorable to mold.

The greenhouse microclimate, meaning the climate within a small enclosed space, must therefore be treated not as a collection of individual switches that can control one variable at a time, but as a nonlinear, time-varying coupled system in which multiple variables constantly push and pull one another.

When this balance breaks down, the damage quickly becomes visible. Just a few days outside the proper temperature range can cause stems to grow excessively or flower buds to drop. If high humidity persists overnight, droplets form on the undersides of leaves and create favorable conditions for fungi and powdery mildew pathogens to multiply. A single failure of fruit set or a single outbreak of disease can eat away at the entire season's yield. The grower's intuition and experience alone are therefore not enough.

Internet of Things (IoT) sensors planted densely throughout the greenhouse must work together with artificial intelligence calculations that read their values and make decisions in real time. If the sensors are the eyes and skin, the calculations behind them are the judgment that weighs the conditions at every moment and chooses the next action.

The problem is that this judgment cannot work properly when a greenhouse has only one sensor. Airflow speed differs between the space directly in front of a ventilation fan and the opposite corner of the greenhouse, while temperatures several degrees apart can be found above a dense crop canopy and in the shade beneath it. The edge near a greenhouse wall and the center experience different weather as well. If a single sensor's representative reading is accepted as the value for the entire greenhouse and used to decide on heating or ventilation, zones far from that sensor may be left unattended to grow leggy or wilt.

Finding a way to handle this spatial nonuniformity remains a central challenge in greenhouse instrumentation design.

The list of sensors that actually goes into a greenhouse is quite long. On the air side, DHT11 and DHT22 temperature and humidity sensors measure temperatures from minus 40 to 80 degrees and relative humidity from 0 to 100 percent, with a WQX-TH temperature and humidity sensor beside them. A DS18B20 water-temperature sensor goes into the water tank. A WQX-L light sensor measures light from 0 to 157,286 lux, while WQX-R and UV sensors measure ultraviolet radiation from 0 to 2,000 watts per square meter.

FC-28 or YL-69 moisture sensors measure the water remaining in soil or substrate, and carbon dioxide concentration in the air is handled by a sensor that uses the degree to which infrared light is absorbed by a gas, namely a nondispersive infrared (NDIR) sensor. A sensor that measures both electrical conductivity (EC) and hydrogen ion concentration (pH) is inserted into hydroponic nutrient solution, while a separate sensor detects the ethylene gas released by fruit during the cold-chain distribution of harvested produce.

A single sensor costs only a few thousand to a few tens of thousands of won, but together they form a net that can read the greenhouse's every breath.

How power is supplied is another issue that determines the lifespan of a sensor node. In places where it is difficult to extend a new power outlet, such as on a greenhouse roof or beside the substrate, nodes equipped with a small solar panel and storage battery are commonly used. Their circuits normally sleep, waking only briefly at scheduled times to take a reading before going back to sleep, allowing them to last for months on a single battery charge.

Actuators that perform physical work, such as irrigation pumps and ventilation fans, are connected to commercial power from the outset, while only the sensor nodes beside them operate separately on low-power batteries. To save even the energy consumed by a transmission, readings are often collected for a set period and sent together instead of being transmitted every time they are taken. If this compromise is poorly calibrated, the battery runs down sooner than expected and the sensor may go dead at the very moment a decision is needed.

Eyes are needed outside the greenhouse as well. When a JNGW100 weather station and WQX spectral sensors are installed outside, they can deliver outdoor air temperature and humidity, dew point, and solar irradiance at one-minute intervals. Overseas open datasets are brought in in real time as well, including the gridded agricultural weather data produced by Japan's agricultural research organization, WeatherAPI, the US National Aeronautics and Space Administration's NASA POWER, and the United Nations Food and Agriculture Organization's FAO CLIMWAT. Forecast values for the coming several days are then fed into the control model.

Because sensors inside the greenhouse cannot warn in advance of an approaching cold snap or heat wave, comparing outdoor weather data alongside indoor sensor readings is increasingly becoming standard procedure.

Knowing even one hour earlier when to shut the vents can make the difference between life and death for a crop; learning only after the reading has risen may be too late.

Nor are all the numbers drawn from the air. A method known as the “speaking plant approach” reads signals emitted by the crop itself. Multiplying the difference in carbon dioxide concentration between air entering and leaving the greenhouse by the airflow rate makes it possible to calculate, every five minutes, how much carbon dioxide the entire crop canopy absorbs through photosynthesis and how much it releases through respiration overnight.

In the same way, multiplying the difference in water-vapor concentration by the airflow rate allows the amount of water released by the leaves and the extent to which the stomata, the tiny pores on the leaf surface, are open to be estimated without touching a single leaf. If a thermometer is a tool for measuring the air in a room, this method turns the crop itself into a living instrument.

Sensor readings travel through several layers. At the front line in the field are an STM32 or ESP32, an ESP8266 module with Wi-Fi, or a combination of an Arduino Uno and Arduino Nano that divide the work between two devices. When the Arduino Uno collects sensor readings and directly operates the actuators, the Arduino Nano beside it communicates with the web server and draws graphs on the display. Dividing the work this way allows one side to keep going if the other stops. Nodes within the greenhouse exchange values over RS485, a wired method long used in industry, or over a ZigBee wireless sensor network.

To push the values up to a central server or the cloud, communications networks such as 4G DTU, GPRS, LPWAN, and Wi-Fi are used in turn. As a result, a small display at the greenhouse, an office computer, and the grower's mobile app all receive the same numbers at the same moment.

RS485 wiring, long proven in industrial settings, carries readings reliably without being disturbed by moisture or electromagnetic interference, but it brings the inconvenience of having to reroute cables whenever the greenhouse layout changes. Wireless sensor networks such as ZigBee require no wiring and suit greenhouses in which cultivation methods change frequently, but obstacles that block radio waves, including steel frames and water tanks, can cause readings to be interrupted or arrive late.

In practice, the two methods are therefore commonly combined rather than standardized on one: core sections are wired, and only sections that change frequently are supplemented with wireless connections.

Among these technologies, the low-power long-range network known as LoRaWAN, a form of LPWAN, has moved within a few years from laboratory technology to industrial-scale infrastructure. As of 2026, the global LoRaWAN IoT market is estimated at $14.25 billion and is projected to grow by 34.7 percent annually through 2035. More than 125 million LoRaWAN devices have already been deployed across 70 countries.

Because this network can last for years on a single battery while carrying readings over distances of several kilometers, sensors can now be planted even in a corner of a plastic greenhouse or open field where installing new electrical wiring would be difficult.

A conspicuous trend in 2026 is the shift of computation away from the cloud and into the communications relay device at the site, namely the gateway. Instead of sending every sensor reading to a distant server and waiting for an answer, a lightweight artificial intelligence model is embedded in the gateway so that it can make an immediate decision. This reduces both the time spent communicating with the server and the cost of that communication. In an experiment by researchers at the University of Nebraska in the United States, real-time soil-moisture monitoring using this method reportedly reduced water use by as much as 35 percent.

The sensor does not stop at taking a reading; it turns that reading directly into a decision.

It is nevertheless impractical to plant expensive sensors in every corner of a greenhouse. One alternative is to install only a small number of sensors and computationally draw a three-dimensional climate map of the entire greenhouse. Computational fluid dynamics (CFD) is a simulation technique that reproduces in a computer how air and gases actually flow and mix inside the greenhouse, much like bringing a weather map into the confined space of the greenhouse.

When it is combined with machine-learning models such as DeepCFD-OptNet, which links a convolutional neural network (CNN) that reads images, a temporal convolutional network (TCN) that follows changes over time, and particle swarm optimization (PSO), which searches among candidate values for the optimum, or with random forests, support vector regression (SVR), and deep neural networks (DNNs), readings from just a handful of sensors can reconstruct plausible distributions of temperature and carbon dioxide throughout the greenhouse.

This reconstruction is usually recalculated every few minutes to update the map, so values close to the most recent conditions can be inferred even in zones without sensors. Computation fills the gap in the tug-of-war between sensor deployment costs and measurement precision.

Redundant design, in which the same value is deliberately measured two or three times, is another way to protect measurement precision. If the system depends on a single sensor to measure temperature, the entire greenhouse operates with its eyes closed when that sensor fails. If two or three sensors with slightly different performance characteristics overlap in the same zone, however, their readings can be compared to determine which one is closest to failure.

Multiple inexpensive sensors are often more resilient to failure than a single expensive sensor, leaving the designer to decide which values should be measured in how many overlapping layers according to the greenhouse's size and budget. Only with such overlapping readings can the adaptive Kalman filter discussed below do its job.

Readings do not always arrive on time. A brief communications outage or a malfunctioning sensor creates a gap in the data. If the gap is short, no more than 10 steps, forward-backward linear interpolation fills it by drawing a straight line between the final reading before the break and the first reading after the connection resumes. If the gap extends beyond ten steps, that method alone becomes forced, so dynamic time warping (DTW) is used to find a segment in the historical record that followed a pattern similar to the current one and import that pattern to fill the blank.

The method resembles filling a missing page in a diary by consulting an entry from a day in the same season the previous year when the weather was similar.

Apart from filling gaps, there is also the problem of readings that spike or fluctuate. In the humid, dusty environment of a greenhouse, sensors often produce a momentary nonsensical reading. A self-adaptive Kalman filter is used to smooth these values. A Kalman filter assigns weights to the somewhat erratic value just read by the sensor and to the expected value inferred from the preceding trend, then quietly converges on a point somewhere between the two. Its principle is the same as the navigation calculation that stabilizes a vehicle's position by considering its previous speed and direction even when the satellite signal wavers.

It is called “adaptive” because the filter adjusts for itself how much each sensor should be trusted. The greater the difference between a new reading and the expected value, the more the weight of that sensor is reduced and the more weight is given to other sensors, preventing the entire decision from being shaken merely because moisture or dust has temporarily disrupted one sensor.

A typical double-film solar greenhouse at the Hengsheng Seedling Base in Ningxia, northwestern China, served as a field site where this filter was tested in practice. Four sets of WQX-TH, WQX-L, and WQX-R spectral sensors were distributed throughout the greenhouse, with a separate weather station installed outside. Once readings arriving simultaneously from the different sensor locations had first been smoothed with the adaptive Kalman filter, the subsequent temperature and humidity forecasts and the automatic decisions to open or close heating and ventilation equipment operated much more reliably.

Unless noise is removed at the measurement stage, even the most sophisticated decision logic placed on top of the data will be unstable at its foundation.

Even these smoothed readings cannot be fed directly into a calculation because their units differ. Temperature is measured in degrees Celsius, humidity in percent, light in lux, carbon dioxide in ppm, and acidity in pH, each on a different scale. Min-max normalization rescales them between 0 and 1, converting them into values that can be placed side by side and compared. The logic is the same as converting the currencies of several countries into a common reference currency before comparing the prices of goods. Only after this process can the values serve as input for a machine-learning pipeline.

Once the values have been smoothed and their scales aligned, it is time to turn them into physical movement. This work is divided among a NodeMCU ESP8266 or Raspberry Pi at the site, a control terminal with a local Qt display, and sometimes a cloud server. From the arrival of a reading to the resulting decision takes no more than a few seconds and can be even faster. Once the decision is made, the command travels through a relay module directly to irrigation pumps, ventilation fans, shade screens, carbon dioxide supplementation systems, misting nozzles, and supplemental LED lighting.

A number that begins with the green light on a sensor box culminates, a few seconds later, in a fan turning at the far end of the greenhouse.

Measured values acquire another use when they are stored and displayed. The values passing through the gateway accumulate in a time-series database, and the grower can scan graphs of temperature and humidity over the past day or week on the web or a mobile phone screen. If a value crosses a predefined danger threshold, a warning appears immediately by text message or app notification. The grower gains time to respond to an unexpected event such as a cold snap or power failure without remaining in the greenhouse all night.

Long-term records later become data that can be compared with yield or disease occurrence, and for produce requiring certification or traceability, they can also serve as evidence of the cultivation process.

Two commercial hydroponic lettuce greenhouses at Sandibuana Farm and Purwosari in Indonesia demonstrate this entire sequence. DHT11 temperature and humidity sensors, DS18B20 water-temperature sensors, FC-28 soil-moisture sensors, and pH and EC nutrient-solution sensors were installed, and their readings were collected through a NodeMCU ESP8266 gateway and carried to a cloud database. Over four months, 15,492 readings accumulated. They subsequently became the raw material for a decision system that predicts yield in advance and adjusts the environment automatically.

The case shows that even a few sensors and an inexpensive Wi-Fi module can build a data foundation at the scale of a commercial greenhouse.

A 300-square-meter winter greenhouse at Kangwon National University in Korea has a similar configuration. Equipped with a pellet boiler, heat-storage tank, and tube-rail heating system, the greenhouse uses a JNGW100 outdoor weather device and a GL840 logger to record, without missing a single reading, data from thermocouple sensors installed throughout the interior every minute. Only with this densely accumulated record does predictive heating become possible, allowing the water in the heat-storage tank to be preheated several hours before a bitter cold snap arrives.

Reducing the measurement interval to one minute is itself a foundation that must be laid before any sophisticated decision logic can be built on top of it.

A smart mist-control trial conducted through an agricultural innovation project in Aichi Prefecture, Japan, shows what changes when one more sensor is added. Instead of looking only at temperature and humidity inside the greenhouse, the system also reads outdoor solar radiation, meaning the intensity of the sunlight, and uses the values together to calculate in real time the vapor pressure deficit (VPD), a single number expressing how thirsty the air is.

If the sunlight suddenly intensifies and this number surges, the system sprays mist in short bursts, just enough to ease the plant's stress without wetting its leaves, while the grower watches the trend directly in a smartphone app. Combining separate values such as temperature and humidity to create a new number closely captures what the term “intelligent” measurement looks like in practice.

Large multitier cultivation racks and vertical farms in Kaohsiung, Taiwan, show what else becomes visible when a sensor network is stacked layer by layer. These farms divide orchids and hydroponic vegetables among shelves on several levels, install separate temperature, humidity, pH, and moisture sensors in each compartment, and gather the readings in a central field control box and cloud monitoring display. Because light and ventilation reach the upper and lower compartments differently, the measurement scheme that once treated the entire greenhouse as a single space has been subdivided by level.

Electricity and water consumption are measured as well, and carbon emissions are drawn on the display, enabling the sensors to serve as a dashboard that reveals the efficiency of the entire farm operation along with crop conditions.

Even a well-designed system has weak links. When relative humidity inside the greenhouse frequently exceeds 90 percent, corrosive gases spread with every pesticide application, and dust and droplets cling together, sensor sensitivity gradually dulls over time. This phenomenon is called drift: the actual temperature inside the greenhouse remains unchanged, but only the number shown by the sensor slowly moves out of alignment.

To prevent it, researchers are developing self-calibrating software that compares multiple sensor readings and corrects the scale without manual adjustment, along with sensors coated in durable nanomaterials that repel water and dust.

Regular calibration is indispensable as well. When a new sensor is first installed, initial calibration aligns the manufacturer's reference reading with the actual field value. Every few months thereafter, the scale is adjusted again by comparison with a standard solution or standard light source whose value has already been verified. A common method for a humidity sensor is to place it briefly in a saturated solution, such as a salt solution that maintains constant humidity, and use that reading as the reference.

Instead of having a person repeatedly walk through the greenhouse to perform this laborious procedure, software calibration is gradually taking over some of the work by comparing readings across sensors and identifying the one that has drifted out of alignment.

How to divide the space and place the sensors remains another unresolved challenge. The larger the greenhouse, the wider the climatic differences between the space in front of a ventilation fan and the center of the greenhouse, and between the upper and lower portions of the crop canopy. A regular grid with sensors placed at equal intervals can easily miss the areas with the largest temperature differences. Alternatives under consideration include stratified sampling, which divides the greenhouse into zones and deploys sensors at different densities, combined with CFD-based reconstruction.

Deciding where to place how many sensors is becoming as important as choosing which sensors to use.

Attempts to combine sensors from different manufacturers with programmable logic controllers (PLCs) and other controllers often cause readings to become misaligned or connections to fail altogether because the devices speak different communications languages. This also creates the burden of checking compatibility with existing equipment every time new machinery is brought onto the farm. Common standards are taking hold to address this problem, including OPC UA, a standard communications protocol widely used in industry; WAGRI, Japan's agricultural data-sharing platform; and the IoP cloud, which adapts the Internet of Things to agriculture.

Along with standardizing protocols, designers must create easy interfaces and controls that growers without engineering training can learn to use within a few days.

The real-world difficulties faced by artificial intelligence calculations deployed in the field are not fully revealed by test accuracy alone. There is a noticeable delay between an actuator receiving a command and physically moving, while physical factors such as wind and vibration intrude at unexpected moments. If communication is interrupted even briefly, the gateway cannot wait indefinitely for an answer from the cloud, so at least a minimal decision logic must remain on site. Running such decisions reliably even on an inexpensive microcontroller is the next objective.

As the shift of computation into the gateway becomes established, the artificial intelligence model installed there must itself become smaller and lighter. Techniques such as quantization, which reduces numbers to 8 bits to lower the computational load, and knowledge distillation, which transfers what a large model has learned into a smaller model, are being refined as answers. Selecting sensors, supplying power, laying a communications network, filling missing readings, filtering noise, and aligning scales may each seem minor in isolation, but if even one is omitted, every decision built on top of them becomes unstable.

Only after the reading from the sensor box attached to a greenhouse post has had its gaps filled, its noise removed, and its scale aligned does it become, within seconds, a decision that moves fans, pumps, and lights.

Whether the sensor inside the white box is actually reading the correct value is another question. A grower holds small bottles containing clear liquids adjusted precisely to pH 4, 7, and 10, dips the sensor tip into them one after another, and adjusts the number on the display to the standard value. This work is verification and calibration. An electrical-conductivity sensor is likewise immersed in a standard solution of known conductivity to reset its scale. The task is not finished after a single calibration; it must be repeated whenever the season changes, and at least every few months.

If six months or a year passes without verification and calibration, the sensor continues to produce numbers, but as time goes by those numbers gradually diverge from the values they purport to represent. Farms therefore maintain a separate calibration log recording when each sensor was calibrated and against which standard. Only with this record can they later distinguish which periods of data should be trusted and which should be viewed with suspicion.

Rules are also needed to decide which sensor to trust when several overlap. If three temperature sensors are placed side by side in the same zone, two of them may read close to 24 degrees Celsius. If the remaining sensor alone reads 31 degrees, the two readings supported by the majority are accepted and the outlier is placed on a watch list. A change that is unlikely to occur in reality, such as the temperature jumping by more than 5 degrees in ten seconds, is filtered out in the same way.

Instead of immediately removing a sensor flagged in this manner, the system observes it for several more days and confirms a fault only if the same discrepancy recurs. A single spike may be a chance event caused by a speck of dust passing by, but a recurring divergence must be treated as a problem with the sensor itself.

Values filled in by interpolation must be flagged. Whether they are connected by a straight line or imported from a similar historical pattern, these numbers are not readings actually taken by a sensor but values inferred through calculation. Without a flag, inferred and measured numbers can become mixed when environmental records are compared with yield several months later, leading the conclusion in the wrong direction.

More difficult is a case in which the interruption is not a communications outage but a genuine failure, such as a frozen pipe or a stopped pump, while interpolation fills the blank with smooth numbers and hides the problem from view. When a gap continues beyond a set period, an on-site inspection alert must therefore be issued in addition to the interpolation, preventing the calculation from quietly concealing the problem.

Designing a communications network requires deciding not only how much wiring to use but also how the system will endure a signal loss. If a wireless sensor node loses its connection to the gateway, it temporarily stores the readings in built-in memory rather than letting them disappear, then uploads the backlog all at once when the connection is restored. In a mesh arrangement in which nodes relay signals from one to another, a route through a neighboring node remains available even if a steel post blocks the radio waves. Wired sections do not need such a detour, but if the cable itself breaks, there is no alternative.

Critical trunk cables are therefore often duplicated by laying a spare line alongside them.

A screen viewed by the grower is of little use if it merely lines up numbers. If the normal range is colored green, conditions requiring attention yellow, and conditions demanding immediate action red, the state of the greenhouse can be judged with a glance. The problem is that if alarms sound too often, the grower may ignore a notification at the very moment it matters. Thresholds are therefore divided into levels: small fluctuations are displayed quietly on the screen, while a phone notification is sent only when several values enter the danger range together or the same warning persists beyond a set time.

Some sites escalate the response and place a phone call after a delay if the first notification receives no response. In one corner of the screen, predicted values for one hour and one day ahead are drawn beside the current reading, helping the grower act before a number rises instead of responding only afterward.

If all these devices are used only on large farms with ample budgets, the achievement is only half complete. The combined cost of sensors, gateways, and software subscriptions can easily reach a level that small farms hesitate to pay, and even after equipment is installed, it can be difficult in rural areas to find a technician who will repair it when it fails. A system that remotely operates pumps and vents is convenient, but it also carries the risk of unauthorized outside access and manipulation. If someone infiltrates the system in midwinter and opens the vents, an entire greenhouse can freeze solid within hours.

Another unresolved problem arises when years of environmental and growth records do not remain with the growers who created their value but are locked inside the equipment vendor's server, causing the record to be severed wholesale the moment the farm switches to another company's equipment. Equipment acquisition costs, maintenance personnel, remote-control security, and data portability remain challenges this technology must solve.

Sources and References

Academic Papers

1. Arfiyani, S. A., Wardihani, E. D., & Helmy. (2026). Smart Greenhouse Data Integration Using Machine Learning Approaches. Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 15(2), 291-305. DOI: 10.22146/jnteti.v15i2.21951.
2. Cai, Y., Liu, D., Xu, Q., Li, B., & Liu, B. (2026). Greenhouse Temperature and Humidity Prediction Method Based on Adaptive Kalman Filter and GWO-LSTM-Attention. Smart Agriculture, 8(1), 148-155. DOI: 10.12133/j.smartag.SA202506033.
3. Huang, X., Zhu, Q., & Huang, M. (2026). Online Detection System for Freshness of Fruits and Vegetables Based on Temporal Multi-source Information Fusion. Smart Agriculture, 8(1), 203-212. DOI: 10.12133/j.smartag.SA202505037.
4. Jamjachi Yauri, R. I., Herbozo Ramirez, J. R., & Ovalle, C. (2026). Sustainable greenhouse using IoT and machine learning to optimize the microclimate for lettuce cultivation. Bulletin of Electrical Engineering and Informatics, 15(3), 2138-2149. DOI: 10.11591/eei.v15i3.9877.
5. Park, S., Kim, D. H., & Oh, K. C. (2026). Field-based implementation of machine learning forecasting for energy-efficient temperature control in a greenhouse system. Journal of Agricultural Engineering, LVII:11eb730c.
6. Qiu, J., Liu, Y., Gao, X., Huang, Y., Zhang, H., Tian, F., Li, W., & Feng, Z. (2026). AgriAgent: End-to-End Large Model Agent System Architecture for Agricultural Environment Control. Smart Agriculture, 8(2), 220-236. DOI: 10.12133/j.smartag.SA202507042.
7. Wang, J., et al. (2026). Research Progress and Prospects of Intelligent Management and Control Technologies for Facility Vegetables. Smart Agriculture, 8(1), 105-120.
8. (2026). Systematic Review of Autonomous Weed Detection and Identification Using Deep Learning. Precision Agriculture.
9. 高山弘太郎. (2026). 施設園芸のメガスケールとストロングミニマム：環境制御高度化と作業自動化に向けて. 農業食料工学会誌, 87(4), 502-503.
10. (2026). 植物環境工学の研究展望 第二十七回 植物工場の環境評価・ 予測における数値解析研究の現状と展望. 植物環境工学, 38(2), 55-60 / 506-510.

Reports and Web Materials

1. Semtech. (2026). LoRaWAN reaches critical mass: 125 million devices and accelerating growth. Semtech Blog. https://blog.semtech.com/lorawan-reaches-critical-mass-125-million-devices-and-accelerating-growth
2. Atom Senses. (2026). 2026 LoRaWAN IoT trends: What's new in low-power wide-area networks. Atom Senses. https://atomsenses.com/company/news/2026-lorawan-iot-trends-whats-new-in-low-power-wide-area-networks
3. HashStudioz. (2026). LoRaWAN for IoT ecosystem. HashStudioz Blog. https://www.hashstudioz.com/blog/lorawan-for-iot-ecosystem/

Lectures and Video Materials

1. 高雄農來訊. (2025). 114年高雄智慧農業年度成果發表暨研討會. YouTube. https://www.youtube.com/watch?v=1Edc6pP4HnU
2. KYOTO SMART CITY EXPO. (2025). スマート農業の最前線～植物との対話からつくる未来の農業. YouTube. https://www.youtube.com/watch?v=PzHccCvmsRs
3. 愛知県農業水産局. (2026).【あいち農業イノベーションプロジェクト】プロジェクト2027 技術提案募集説明会（2026.6.10開催）. YouTube. https://www.youtube.com/watch?v=jpkoNqscGu8
4. 農業農村情報通信環境整備 準備会. (2026). 第5回オンラインセミナー「地域でデータを共有しよう！農業データの利活用最前線～データ駆動型農業WAGRIとIoPクラウド（SAWACHI）でのNext次世代型施設園芸システム～

」. YouTube. https://www.youtube.com/watch?v=k8Qvv6CKB2A

5. Supreme Knowledge Foundation. (2026). IoT and GNSS technology based Smart Precision Farming and Community Development. [Part I]. YouTube. https://www.youtube.com/watch?v=8VssYT1onkA
6. Kanal Pengetahuan Fakultas Teknologi Pertanian UGM. (2025). AI and IoT in Smart Agriculture: How Young Innovators Can Shape the Future of Farming. YouTube. https://www.youtube.com/watch?v=77_Aa0bdchQ
7. 農業農村情報通信環境整備 準備会. (2026). 第6回オンラインセミナー「自治体・ 農業DXを支える次世代無線:Wi Fi HaLowの最新動向と活用事例」. YouTube. https://www.youtube.com/watch?v=1Gn_rbR6_Tc
8. Life at PTPN. (2025). Webinar GO DIGI SERIES #5 IOT And AI Implementation In Agriculture. YouTube. https://www.youtube.com/watch?v=TvmtvwyHwUQ

3.2

Digital Twins and AI-Driven Environmental and Growth Optimization

A Digital Twin combines a virtual model that reproduces the structure and equipment of a real greenhouse inside a computer with real-time sensor data such as air temperature, humidity, carbon dioxide concentration, and soil moisture. It uses this information to calculate in advance how the greenhouse environment will change. An action such as opening a roof vent or adjusting an irrigation valve is first applied to the on-screen model, and its result is checked computationally before the same command is sent to the physical greenhouse equipment. It is called a twin because it moves through time together with the state of the greenhouse.

A digital twin is more than a three-dimensional image modeled on a greenhouse. If outward resemblance were all that mattered, a single photograph would be enough and there would be nothing to calculate. Its real substance lies in the living engine that calculates how the greenhouse air and soil, and the plants growing within them, change over time. At the center of this engine is version 4.8 of DSSAT, the Decision Support System for Agrotechnology Transfer, a crop growth model used widely around the world.

On a daily basis, the model calculates how nutrients produced through photosynthesis in the leaves are distributed among the roots, fruit, and other parts of the plant; how quickly the roots absorb water and nutrients from the soil; and how strongly the plant responds to stresses such as heat or drought. When a grower opens or closes a window on the screen, the engine recalculates to reflect that change and displays in advance how much thicker the greenhouse tomato stems will become and how much farther the leaves will droop by the next day.

Inside a greenhouse, air temperature, humidity, carbon dioxide concentration, sunlight intensity, and the amounts of water and fertilizer in the soil or growing medium do not move independently. They move as one interconnected system.

Opening a window lowers temperature and humidity together and even changes how quickly carbon dioxide escapes outside. The interconnection is no different from a shower in which turning one handle changes both water pressure and temperature. Conventional on-off control, which starts a fan at a preset temperature and stops it when the temperature falls, cannot keep pace with this interaction.

Its response always comes one step late. Heating fuel and water are wasted by that delay, while plants must endure days of conditions that are hotter, colder, or drier than necessary. This is why digital twins and artificial-intelligence control have moved to the center of greenhouse technology. The intertwined reactions can be caught in time only by calculating them on screen before moving the physical windows and valves.

A digital twin shows its greatest value when a grower tests an intervention never attempted before. If the grower tests in a real greenhouse what will happen when supplemental lighting is set much higher than usual or how fruit size will change when irrigation begins two hours earlier, the result will take at least a season and perhaps a full year to become known. If the judgment was wrong, that year's harvest is lost. In the twin greenhouse on the screen, however, the same intervention can be simulated in minutes and simply reversed if the result is poor.

The principle is the same as a novice pilot making mistakes repeatedly in a flight simulator before boarding a real aircraft. The very existence of a practice ground where failure is safe allows growers to test measures they once hesitated to try with greater confidence.

No person can manage all these calculations and decisions around the clock. Researchers therefore created artificial-intelligence agents that monitor the greenhouse and make judgments on the grower's behalf. AgriAgent, developed by researchers in China, is one example. The system is not merely a chatbot that answers questions. It is closer to a novice grower who lives in the greenhouse. It imitates the discipline of reading the instrument panel, remembering what was done the previous week, consulting a thick cultivation manual when necessary, and only then turning a valve.

AgriAgent's core decision-making function is performed by a Large Language Model (LLM) called Qwen2.5. Three differently sized models with 1.5 billion, 3 billion, and 7 billion parameters were tested side by side. Parameters are the numerical values set through learning inside a model. The more parameters a model has, the more situations it can distinguish and remember and the finer its judgments can become.

Once every seven days, AgriAgent moves in order through five stages: perception, memory, retrieval, reasoning, and action. Sensor data first create a state value organized into twelve numbers. These include the amount of water in the soil; the amounts of soil nitrogen and organic matter; the day's average and maximum temperatures; daily rainfall and wind speed; atmospheric carbon dioxide concentration; the intensity of light actually available for photosynthesis; the number of days since sowing; and two stress indices that express water shortage and delayed growth as single scores.

A device called a format converter turns these numbers into sentences people can read, attaches past records and the retrieval results described below, and sends them to the artificial-intelligence model. Spelling out the twelve numbers sentence by sentence amounts to translating into human language how thirsty the crop is now and how well it has grown. The procedure turns a table filled with numbers into a short report that a person can read and understand at once.

Agriculture contains a great deal of detailed knowledge found only in textbooks, and even the largest language model cannot memorize all of it. Instead of committing this knowledge wholesale to memory, AgriAgent opens a thick book to the exact page it needs whenever a question arises. To implement this approach, called Retrieval-Augmented Generation or RAG, the researchers used optical character recognition (OCR) to convert 5.1 million characters from agronomy textbooks and technical handbooks into digital documents. They selected 4.4 million useful characters from that material and divided them into smaller passages.

Each passage received a label identifying the chapter and section from which it came. A 768-dimensional embedding model called bge-large-zh, which converts sentences into numerical vectors and places similar meanings near one another, finds semantically relevant passages. A lexical search that checks for exact word matches is added, and the system selects five documents that fit the current situation and passes them to the artificial intelligence.

With the documents in hand, the artificial intelligence does not produce an answer immediately. It proceeds in stages through a Chain-of-Thought process. It first determines whether the crop is germinating or flowering and setting fruit, then checks whether water is insufficient or growth is lagging, and uses those judgments to decide on commands. The method resembles a person writing out the solution line by line instead of blurting out an answer from mental arithmetic.

The conclusion is organized as a standardized data object in JSON comprising six values: irrigation rate, supplemental-lighting intensity, ventilation status, heating status, fertilizer rate, and carbon dioxide release rate. Irrigation is set between 0 and 30 millimeters per hour, supplemental lighting between 0 and 200 percent of its normal level, and fertilization between 0 and 50 kilograms per hectare. As a final verification step, a tool executor checks once more that the values remain within safe bounds. It clips any value outside the range and only then sends the commands to the DSSAT engine or the physical greenhouse equipment.

Consider a Monday morning when sensor readings show less water in the soil than usual and a daytime maximum temperature 3 degrees higher than the seasonal norm. The water-stress index has risen close to its danger signal. AgriAgent reads this combination together with the fact that the crop is now in the fruit-sizing stage and a passage in its knowledge base explaining that dry soil at this stage readily causes physiological disorders. It concludes that irrigation should be raised above its usual level while midday supplemental lighting should be lowered slightly to reduce the burden on the leaves.

A human grower might make this judgment through experience and intuition. AgriAgent organizes it in seconds from numerical and documentary evidence.

Equipment can fail at any time. Suppose the irrigation system stops and can no longer deliver water to the soil. A person might panic, but AgriAgent uses causal reasoning to find another route. It can slow the loss of water through leaf transpiration by adjusting ventilation and heating, then raise carbon dioxide release to support photosynthesis and partly offset damage from the water shortage. The response resembles a cook discovering that one ingredient is missing from the refrigerator and adjusting the proportions of the remaining ingredients to complete the same dish.

Until a person arrives to repair the valve, this temporary adjustment helps the crop avoid the worst outcome.

To determine whether this judgment can be trusted, the system must be tested in advance against extreme weather. Using 30 years of weather data from 1994 through 2023, the researchers created three levels of virtual weather, normal, fluctuating, and extreme, and prepared a total of 90 annual scenarios. The extreme scenarios include compound hazards such as 20 consecutive days of hot, dry weather; more than 80 millimeters of torrential rain concentrated into five days; and a sudden reversal in which 20 dry days are followed immediately by 10 days of heavy rain.

AgriAgent is put through these situations so researchers can observe whether the crops survive and how well they endure. The worst combinations, which would rarely be encountered in a real greenhouse, can be experienced repeatedly and without risk on the screen.

Resilience refers to a concrete scene. In a scenario with 20 consecutive hot, dry days, a less capable control method allowed the stress index to enter the danger zone, after which it rarely recovered and the crop simply wilted. AgriAgent, by contrast, increased irrigation and supplemental lighting little by little at the beginning of the water shortage, holding the stress index at a low level. When the abrupt shift brought heavy rain, it quickly increased ventilation and reduced damage from excess moisture. Even in the worst weather, it did not give up. It kept making a slightly better choice for the conditions at every moment.

Language-based artificial intelligence such as AgriAgent is not the only kind of AI to enter the greenhouse. Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs), used for much longer, monitor the microclimate as well. Both are statistics-based programs that find patterns in numerical data. Working through a data-analysis tool called Orange, they examine greenhouse temperature, humidity, and ultraviolet intensity together and classify the current state into one of three categories: adequate, excessive, or deficient. The classification accuracy reached 99.99 percent.

That is scarcely one error in ten thousand decisions, a much finer level of surveillance than a person repeatedly checking the instrument panel by eye.

Temperature and humidity rise and fall sharply within a single day and affect each other, making them difficult to predict. Greenhouse readings fluctuate even when the sun is hidden briefly by a cloud and then reappears. The researchers first apply an adaptive Kalman filter to readings from sensors in several locations to settle this fluctuating noise. The procedure resembles compensating for camera shake to reduce blur in a photograph. Only after the noise has been removed can the next prediction model learn without being pulled in the wrong direction.

After removing the noise, Grey Wolf Optimizer (GWO) tunes the detailed settings of the prediction neural network. Like a pack of gray wolves gradually surrounding prey from several directions, the method repeatedly narrows multiple candidate settings until it finds a suitable combination. Attention is then added to a Long Short-Term Memory (LSTM) neural network refined in this way, allowing it to select and assign weight to the moments in the previous several hours that truly matter to the current prediction. This combination predicts in one pass how greenhouse temperature and humidity will move over the next several hours.

No matter how accurate a forecast is, it is useless unless it can be converted into a decision to switch physical heating equipment on and off. Data-Driven Predictive Control (DDPC) anticipates a drop in temperature and heats water before the cold wave arrives. A field trial was conducted in a 300-square-meter greenhouse in Chuncheon, Gangwon Province, equipped with a pellet boiler, a thermal-storage tank, and tube-rail heating. Data were collected at one-minute intervals.

A Gated Recurrent Unit (GRU), a neural network lighter than LSTM, repeatedly learned from 10,080 measurements covering the most recent seven days and was tested on the following day. The model forecast a fall in greenhouse temperature one to three hours before a cold wave and operated the boiler and hot-water valves in advance. It prevented the temperature from falling below the minimum the crops could tolerate while reducing unnecessary heating and saving fuel costs.

Heating and yield often pull in opposite directions. When the temperature is kept generously high, crops grow faster and set more fruit, but the boiler runs continuously and consumes fuel. Conserving heat cuts fuel costs, but crops suffering in low temperatures grow more slowly or drop flower buds, leading to a loss of yield. A digital twin does not calculate these two objectives separately. It overlays them on the same screen. Running the same winter scenario at several heating levels reveals the temperature at which the combined loss from fuel expense and reduced yield is smallest.

The GRU-based DDPC in the Chuncheon greenhouse ran the boiler briefly just before a cold wave and barely maintained the minimum temperature. This, too, was the result of recalculating that intersection in real time.

The performance of these prediction models can be checked through several numbers. The coefficient of determination, written as R-squared, expresses on a scale from 0 to 1 how closely predicted and observed values overlap. A value closer to 1 indicates a better fit. The microclimate prediction models produced values between 0.77 and 0.93. The GRU-based heating-control model in the Chuncheon greenhouse had an error of approximately 1.43 degrees Celsius by root mean square error, or RMSE. A difference of about one and a half divisions on a thermometer is sufficiently precise for heating control.

The model that classified greenhouse conditions as adequate, excessive, or deficient recorded AUC-ROC values between 0.988 and 0.991, meaning that it almost never confused the three states.

AgriAgent's capabilities were tested across several crops. On a DSSAT-based digital twin, the researchers selected five different crops: maize, millet, sugar beet, tomato, and Chinese cabbage. Maize represents a crop with an exceptionally efficient photosynthetic pathway; millet represents a drought-tolerant crop; sugar beet is grown for its root; tomato for its fruit; and Chinese cabbage for its leaves. Together they represent crops with widely differing modes of growth. The 90 weather scenarios were multiplied by three levels of severity, producing a total of 450 full-season simulations.

Compared with natural growth without human intervention, the model with 7 billion parameters clearly increased harvested dry weight in maize, millet, and sugar beet. When a low-temperature alert appeared during germination, it raised the intensity of supplemental-lighting intervention to 120 percent on its own and reduced cold injury. Control performance improved in order as parameter counts increased from 1.5 billion to 3 billion and 7 billion.

Looking at one maize simulation, the agent first counts the days since sowing to determine whether the crop is in the emergence stage or the ear-formation stage. The same low temperature does different damage on the tenth day after sowing than when ears are filling. If the growth stage is misidentified, even an accurate temperature prediction leads to the wrong prescription. In the actual test, the agent's decision to raise supplemental lighting in response to a low-temperature alert was possible because this growth-stage judgment came first. It goes beyond counting dates to read what each date means inside the crop.

Digital twins have moved beyond the laboratory into real commercial farms. At Sandi Buana Farm in Indonesia, 15,492 samples were gathered from high-resolution sensors measuring temperature, humidity, EC, and acidity while hydroponic lettuce was grown. EC measures how much fertilizer is dissolved in the water through its electrical conductivity, while acidity indicates how acidic or alkaline that water is. The data were fed into a machine-learning prediction algorithm to operate a system that automatically coordinates lettuce yield with variables in the growing environment.

Appropriate water and fertilizer concentrations were maintained without a person adjusting them one by one. The farm demonstrated that a decision system can be built from numbers accumulated in the field without a single laboratory simulation.

Language-model agents operating on digital twins and statistics-based prediction models used in greenhouses for much longer are not rivals. They are more like colleagues with different roles. The agent reads the growth stage and multiple variables together and balances irrigation, fertilization, and supplemental lighting within the larger picture. ANN, SVM, and LSTM-family models capture temperature and humidity changes occurring by the second and minute. The two levels often overlap in the field. A natural division of labor develops in which the agent makes major decisions and the statistical model fine-tunes them second by second.

One level decides whether to open the roof vent; the other monitors each second after it opens to see whether temperature behaves differently from the forecast. With only one level, the system would have remained a partial decision-maker, either understanding the larger picture but missing momentary fluctuations or catching the moment while failing to read the flow of an entire season.

Scaling an agent verified in one greenhouse to ten or one hundred greenhouses is not as easy as arithmetic. Every greenhouse has a different orientation and building material, and surrounding buildings cast different shadows, so the same command produces different results. When an agent is introduced into a new greenhouse, an adaptation period is therefore often provided during the first few weeks so it can learn the habits of that greenhouse. The knowledge base and decision patterns accumulated in an existing greenhouse provide the foundation, but the system is retuned little by little as numbers from the new site accumulate.

This trend extends beyond a handful of local experiments. Recent research has proposed placing a digital twin online as a simulator and connecting it with a large-language-model agent in a closed loop. The structure validates the safety of a command in real time before it is sent to physical equipment. The result shows that a language model can perform its role even within a loop that requires decisions timed precisely to an unfolding process.

Smart-greenhouse implementations have combined communication networks of wireless sensors and actuators with twin software to monitor microclimate and irrigation remotely, control them automatically, and predict the likelihood of disease. Without crossing the greenhouse threshold, a grower can check current conditions and test an intervention on screen before applying it in reality. In 2026, a data-driven greenhouse-management method integrating a crop growth model with autonomous climate control was presented, showing that digital twins are moving from display models to tools that both decide and act.

The screen a grower actually sees is a colorful dashboard. The twelve state values and six control commands appear side by side in graphs and bars, accompanied by one or two sentences summarizing why the artificial intelligence issued those commands. The grower can read the explanation and leave the command unchanged or edit the numbers by hand and return them if dissatisfied. The interface directly embodies the principle that the agent's judgment remains a draft until a person gives final approval.

Before this system was installed, growers had to walk through the greenhouse with a flashlight every dawn and check the thermometers one by one. On nights with cold-wave advisories, waking every two or three hours to inspect the boiler was not unusual. As the digital twin and agent take over this repetitive labor, the grower's day changes. The grower is called to the screen only when a problem arises. If the night is calm, morning arrives without a single alert. If something goes wrong, the grower's phone sounds at that moment. Patrols once made by hand have disappeared, but computation has taken over the work.

This does not mean the technology is complete. When a control method verified inside a DSSAT-based digital twin is transferred unchanged to a real greenhouse, the calculations and reality begin to diverge. The time between an actuator receiving a command and moving, the gradual drift in aging sensor readings, temperature differences between the corners and center of a greenhouse, and unforeseen external shocks all create a gap between the on-screen calculation and the physical result. Researchers call this the Sim-to-Real Gap.

They seek to close it by training the twin on harsher situations in advance and allowing it to continue learning after deployment in the field. The fine-tuning method called LoRA is well suited to this continual learning because it adjusts only a few controls to fit the new site rather than retraining the entire large model.

No matter how sophisticated the technology becomes, without people who can operate it, it ends up as an ornament in the corner of a greenhouse. The grower ultimately decides whether to accept AgriAgent's judgment or manually revise the numbers based on years of experience. Such trust is not built in a day. It thickens only after the agent continues making decisions for several seasons without a major mistake. A grower who has witnessed the on-screen twin behave differently from the real greenhouse several times will not readily trust even a highly accurate prediction afterward.

Running a model with 7 billion parameters requires expensive GPU servers and an uninterrupted communication network. These requirements create a barrier for farms without abundant capital and introduce delays every time information travels across the network. To address the problem, researchers use quantization (INT8), which stores each numerical value in a model with fewer bits, together with Knowledge Distillation, which teaches a smaller model the decision method learned by a larger one.

Just as an accomplished senior worker organizes practical knowledge and passes it to a junior, the decision patterns of the 7-billion-parameter model are transferred into models with 1.5 billion or 3 billion parameters so they can run on inexpensive equipment at the greenhouse. A model reduced in this way can complete its judgment on a small computer beside the greenhouse without sending data to the cloud. The agent must be able to operate in the greenhouse of an individual smallholder if the technology is not to remain a tool reserved for a few large farms.

The goal of increasing yield frequently conflicts with the goals of saving energy and reducing water use and carbon emissions. Stronger supplemental lighting increases yield but raises the electricity bill at the same time. Researchers continue to design multilayer reward calculations so that the agent does not favor only one side among these competing objectives. Growers must be able to see directly on the screen why the artificial intelligence decided to open a window at a particular time and what evidence supported that judgment.

However accurate a model may be, it is difficult to entrust it with field operations if its decision process is hidden like a black box. More research to improve explainability is needed before digital twins and intelligent control can leave the laboratory and become everyday tools in ordinary greenhouses.

Sources and References

Academic Papers

1. Qiu, J., Liu, Y., Gao, X., Huang, Y., Zhang, H., Tian, F., Li, W., & Feng, Z. (2026). AgriAgent: End-to-End Large Model Agent System Architecture for Agricultural Environment Control. Smart Agriculture, 8(2), 220-236. DOI: 10.12133/j.smartag.SA202507042.
2. Park, S., Kim, D. H., & Oh, K. C. (2026). Field-based implementation of machine learning forecasting for energy-efficient temperature control in a greenhouse system. Journal of Agricultural Engineering, LVII:11eb730c.
3. Cai, Y., Liu, D., Xu, Q., Li, B., & Liu, B. (2026). Greenhouse Temperature and Humidity Prediction Method Based on Adaptive Kalman Filter and GWO-LSTM-Attention. Smart Agriculture, 8(1), 148-155. DOI: 10.12133/j.smartag.SA202506033.
4. Arfiyani, S. A., Wardihani, E. D., & Helmy. (2026). Smart Greenhouse Data Integration Using Machine Learning Approaches. Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 15(2), 291-305. DOI: 10.22146/jnteti.v15i2.21951.

5. Xu, Q., Zheng, B., & Zhong, S. (2026). Artificial Intelligence Empowering Modern Agricultural Biological Breeding. Smart Agriculture, 8(3), 69-85.

6. Wang, J., et al. (2026). Research Progress and Prospects of Intelligent Management and Control Technologies for Facility Vegetables. Smart Agriculture, 8(1), 105-120.

7. Jamjachi Yauri, R. I., Herbozo Ramirez, J. R., & Ovalle, C. (2026). Sustainable greenhouse using IoT and machine learning to optimize the microclimate for lettuce cultivation. Bulletin of Electrical Engineering and Informatics, 15(3), 2138-2149. DOI: 10.11591/eei.v15i3.9877.

Reports and Web Materials

1. arXiv. (2025). A study validating closed-loop greenhouse control using a digital twin based on an online simulator and a large-language-model agent. https://arxiv.org/pdf/2507.07115

2. Springer. (n.d.). Implementing a smart-greenhouse digital twin: Remote microclimate monitoring and disease prediction through a wireless sensor and actuator network. https://link.springer.com/chapter/10.1007/978-981-99-0252-1_5

3. PLOS ONE. (n.d.). Remote condition monitoring and virtual intervention validation in a digital-twin greenhouse. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0344946

4. ScienceDirect. (2026). Data-driven adaptive greenhouse management integrating crop growth modeling and autonomous climate control. https://www.sciencedirect.com/science/article/pii/S2590005626000445

Lectures and Video Materials

1. 高雄農來訊. (2025). 114年高雄智慧農業年度成果發表暨研討會. YouTube. https://www.youtube.com/watch?v=1Edc6pP4HnU

2. Kanal Pengetahuan Fakultas Teknologi Pertanian UGM. (2025). Artificial Intelligence in Smart Greenhouses: Enhancing Crop Monitoring and Environmental Control. YouTube. https://www.youtube.com/watch?v=zFP6Nm4i4Ec

3. WiseFood Project. (2025). WiseFood and STELAR Webinar: Empowering Agrifood Apps With LLM-Based Dataset Search and Integration. YouTube. https://www.youtube.com/watch?v=Keqo17qBr_g

4. Regace Project. (2026). Agrivoltaics in Greenhouses - Turning Sunlight into Energy & Crops. YouTube. https://www.youtube.com/watch?v=rSJvjAxnY8w

5. KYOTO SMART CITY EXPO. (2025). スマート農業の最前線~植物との対話からつくる未来の農業. YouTube. https://www.youtube.com/watch?v=PzHccCvmsRs

Chapter 4

Data-Driven Precision Agriculture and Decision Support

4.1

Intelligent Irrigation and Fertilizer Supply Optimization

In an intelligent irrigation system, it is not a person who opens the valve, but sensors inserted into the soil and a predictive model running on a cloud server. The system calculates how dry the soil is, how high the daytime temperature has risen, and the percentage chance of rain tomorrow, then determines when to irrigate and how much water to apply. In field trials in Nebraska, United States, cornfields using such irrigation scheduling tools saved 29 percent to 43 percent of their water, while soybean fields saved 10 percent to 37 percent, with yields in both fields remaining similar to those achieved with conventional methods.

In traditional farming, when and how much water and fertilizer to apply depended on the grower's practiced touch and a fixed timetable. The water might be turned on at six in the morning, and fertilizer mixed in once a month. The problem is that soil and sky do not move according to a timetable. On some days, more water is poured onto already wet soil, leaving roots no room to breathe. On others, parched soil is left untouched and crops wilt. Excess water flows deep underground, carrying soil with it, while the nitrogen and phosphorus dissolved in that water run into streams.

As climate change intensifies alternating spells of drought and heavy rain, irrigation and fertilization based on intuition and timetables are causing ever greater losses.

Intelligent optimization of irrigation and fertilizer supply is a technology designed to reduce these losses. It integrates soil moisture sensors, weather observation equipment, the Internet of Things (IoT, a communications network that connects multiple devices through the internet so they can exchange information), and machine-learning prediction models into a single system.

The system measures from multiple angles the movement of water among soil, air, and crops; calculates water and nutrient needs that change with each stage of crop growth; feeds sensor data and weather forecasts into a model to predict requirements over the next several days; and sends those predictions to a cloud server that physically operates valves and pumps. A meta-analysis conducted in northeastern China found that combining drip irrigation and fertigation in place of conventional irrigation and fertilization increased crop yield, water productivity, and nitrogen use efficiency together (Li et al., 2021).

The finding suggests that delivering a precise amount directly in front of the roots is better than spraying water across an entire field.

The Nebraska field trials compared the methods growers normally used with commercial irrigation scheduling tools operated alongside them. They used far less water while producing the same yields. Cases have also been reported in which AI scheduling, continuously monitoring sensor readings and changing its decisions at each growth stage, saved as much as 87.97 percent of water. This figure, however, represents an upper bound observed for a particular crop under particular conditions, so it should not be assumed to apply unchanged to every farm.

Scientists call the relationship through which land, sky, and crops exchange water the Soil-Plant-Atmosphere Continuum (SPAC). The term describes water flowing continuously from soil to roots, from roots to stems and leaves, and from leaves back into the air. The amount of water in the soil is not the only factor that determines the rate of this flow. When the temperature is high, the wind is strong, and sunlight is intense, water evaporates rapidly from leaves, and roots draw water from the soil correspondingly faster.

The water evaporating from soil and the water transpiring from leaves are collectively called evapotranspiration (ET), and irrigation systems estimate it by continuously measuring temperature, humidity, solar radiation, wind speed, and precipitation. Capacitive or resistive moisture sensors are buried at several depths in the soil to monitor moisture in both the topsoil and the depths actually reached by roots. The system layers these data to create a picture of how much water remains in the field at that moment.

Crops do not require the same amount of water and fertilizer from birth to death. A small amount of water is enough during germination, when a seed sprouts, but demand rises steeply during the vegetative stage, when the plant grows vigorously. If water is scarce during flowering, flowers fall. During fruit set and enlargement, even a slight shortage can make fruit small or cause it to split. For some crops, water must instead be reduced during ripening to increase sugar content. Researchers express these changes with the Development Volume Index (DVI, an index that represents the crop's degree of development as a single number).

The crop's current level of development is expressed as a number between 0 and 1, then compared with a water-stress threshold set for each stage. The same logic is used to calculate different concentrations and amounts of fertilizer components, such as nitrogen, phosphorus, and potassium, for each growth stage. This approach differs at the roots from a single timetable that supplies the same amount throughout the year.

It is difficult for a person to calculate all these physical and growth data by hand. When the number of variables rises to ten or twenty, it becomes hard to judge which variable is decisive in the current situation. Over the past several years, machine-learning models have begun taking over these calculations in greenhouses and open fields. Using the numbers delivered by sensors and weather agency forecasts as inputs, a model determines whether irrigation is needed at that moment and, if so, how many liters to apply.

The choice of model depends on the farm's crop and scale, as well as the kind of data collected and the length of the collection period.

Random Forest is one of the most widely used models. As its name suggests, it plants many decision trees to form a forest, puts the judgment of each tree to a vote, and reaches a conclusion by majority rule. It produces stable results even when given intertwined variables such as soil moisture, temperature, humidity, rainfall, soil type, and crop type. When a research team in India tested the model on a web-based platform, users entered field conditions and received a yes-or-no answer on whether irrigation was needed. When it was needed, the system estimated a specific quantity of water, approximately 25 liters (Ajay Kumar et al., 2026).

In Indonesia, researchers equipped vegetable plots growing mustard greens and lettuce with a NodeMCU microcontroller, a temperature sensor, and a real-time clock module, collected 1,200 records, and used the same Random Forest algorithm to open and close irrigation valves automatically (Pradasari et al., 2026). The model appears to be adopted frequently in the field because it is resistant to overfitting, the phenomenon in which a model matches its training data but fails in an actual field.

Decision Tree models show strength in determining fertilizer quantities. One example is EcoGrow, a cloud platform built from conventional guidelines issued by the Horticultural Crop Research and Development Institute (HORDI) under Sri Lanka's Ministry of Agriculture. When a user enters the crop type, growing region, soil characteristics, irrigation interval, climatic zone, and fertilizer application timing, the platform states how many kilograms per hectare of three fertilizers, Urea, triple superphosphate (TSP), and muriate of potash (MOP), should be mixed.

Trained on 1,535 records covering 32 horticultural vegetable crops, the model achieved an accuracy of 97.39 percent and was selected for actual deployment on the strength of that result. EcoGrow goes further by combining on one screen a feature that locates nearby fertilizer retailers and an AI chatbot (a program that automatically responds in a conversational manner) that answers questions about fertilization methods. Farmers can mix fertilizer according to the formula shown on the screen without searching through numerical tables themselves.

Another approach combines Linear Regression with a Deep Neural Network (DNN, an artificial neural network that learns complex patterns through multiple stacked layers). Linear Regression draws something close to a single straight line showing the rate at which soil moisture declines over time. The deep neural network adds a photograph. It combines the judgment of a Convolutional Neural Network (CNN), which classifies soil photographs into types such as alluvial soil, black soil, clay soil, and red soil, with weather data and measurements of nitrogen, phosphorus, and potassium concentrations to issue a fertilizer prescription.

The ability to distinguish the visual appearance of soil and the ability to calculate numerical data are brought together in one model.

Reinforcement Learning is used in rice farming. Reinforcement Learning continuously observes the reward returned by an action and discovers better actions on its own through trial and error. Given weather forecast data, a model follows a Markov Decision Process (a mathematical framework that determines the next action by looking only at the current state) and learns when irrigation will be most advantageous. The model combines DVI tracking with Alternate Wetting and Drying (AWD, an irrigation method that repeatedly floods a rice paddy and then allows it to dry until the soil is exposed).

During water-sensitive periods, such as heading, it conserves water while applying precisely the amount required (Chen et al., 2021). In contrast to conventional rice farming, which keeps paddies continuously flooded, the results suggest that finely controlled cycles of drying and flooding are actually more beneficial to rice.

These models follow similar procedures when they operate in an actual field. Missing values in sensor data are filled in, numerical values are normalized to a range between 0 and 1, and text values such as crop type or soil type are converted into numbers. A model is trained in advance outside the server using the prepared data, and only the trained model is then uploaded to the cloud server. Field sensor nodes send new values to the server at intervals ranging from 1 minute to 15 minutes. As soon as the server receives those values, it runs inference to decide whether to irrigate and fertilize.

The decision passes through an IoT gateway and activates a relay switch. The amount of water actually delivered is measured again by a pulse water meter (a meter that counts one brief electrical signal each time water passes) and recorded in a database. An intelligent irrigation system exists only when this cycle of prediction, execution, and renewed measurement continues without interruption.

Looking closely at each sensor buried in the field shows where this judgment begins. Several capacitive or resistive moisture sensors, models such as the FC-28 and YL-69, are buried in the soil at different depths. Because moisture in the topsoil and moisture at root depth dry at different rates, measurements from only one location can easily miss actual conditions. A DS18B20 temperature and humidity sensor and a DS3231 real-time clock module are added, leaving a record of the exact hour and minute at which the soil reached a given temperature.

Nutrient-solution cultivation, meaning hydroponic farms that dissolve fertilizer in water and deliver it directly to roots without soil, requires a different set of sensors. An Electrical Conductivity (EC) sensor measures the total concentration of dissolved fertilizer salts; a pH sensor measures whether the water is acidic or alkaline; sensors measure nitrogen, phosphorus, and potassium concentrations; and a Total Dissolved Solids (TDS) sensor completes the set. If electrical conductivity is too high, reverse osmosis draws water out of root cells, causing damage similar to roots drying out. If it is too low, nutrients are insufficient.

If acidity moves outside the narrow range prescribed for each crop, the pathways through which roots absorb nutrients are themselves blocked. The control program at a nutrient-solution cultivation site uses electrical conductivity and acidity as reference points to calculate how much concentrated fertilizer and acidic solution to add to the water tank, then automatically activates pumps to correct a value the moment it moves outside the permitted range.

Sensors inside the field are not enough. The system draws at once on historical weather databases accumulated over more than 40 years, soil survey data, information on more than 30,000 crop varieties, and 7-day weather forecasts. It receives real-time temperature and precipitation probabilities from external services such as WeatherAPI, NASA POWER from the United States' National Aeronautics and Space Administration, and FAO CLIMWAT from the Food and Agriculture Organization of the United Nations. If rain is forecast for tomorrow, the system can reduce tonight's irrigation in advance.

In the sky, an unmanned aerial vehicle flies low over the field carrying a multispectral camera and visible-near infrared (Vis-NIR) spectroscopy equipment, recording the color and reflected light of crop leaves. Even when a leaf still appears healthy to the human eye, a developing shortage of nitrogen, potassium, or magnesium first changes the degree to which it reflects particular wavelengths of light. Analyzing these reflectance data with the Diagnosis and Recommendation Integrated System (DRIS) can identify which nutrient is deficient before visible symptoms appear.

In a trial that linked unmanned-aerial-vehicle data from an apple orchard to a fertigation model, nitrogen deficiency could be detected from the sky without removing leaves and sending them to a laboratory (Sun et al., 2023).

The path from numbers measured in a field to a server also has multiple layers. Low-cost microcontrollers such as the ESP32, NodeMCU ESP8266, Arduino, and PIC sit at field nodes and read sensor values. These values travel to the cloud through Wi-Fi; an RS485 wired bus commonly used in industrial settings; LPWAN, which covers a wide area with low power consumption; a fourth-generation mobile communications module; and MQTT, a communications protocol that exchanges short messages with little power. On the server side are backends such as a Firebase real-time database, FastAPI, and Flask.

Above them runs machine-learning code written with libraries such as Scikit-learn, Pandas, and NumPy. Growers do not need to understand this entire process. They only need to look at the dashboard displayed on a website or mobile phone. One platform in Maharashtra, India, displayed its interface in Marathi, Hindi, and English so that irrigation and fertilization advice could reach farms where English explanations were difficult to read (Shriwas et al., 2026).

A decision reached in the cloud must return to the field to produce a real effect. When the server decides to irrigate, its signal passes through an IoT gateway to a digital relay installed in the field. The relay switches a pump on or off and opens or closes valves for each zone. Variable nozzles that can alter the thickness of the water stream allow different quantities to be delivered to different zones through the same pipes. An electronic signal now opens and closes a faucet that a person once turned by hand. On every cycle, the signal begins at a sensor in the soil, travels around the cloud, and returns to the soil.

Some systems add a camera. A self-adaptive irrigation system integrating an ESP32 sensor terminal, an ESP32-CAM imaging sensor, and a FastAPI server does more than measure temperature, humidity, and soil moisture. It also monitors crop condition with a camera (Ramya et al., 2026). Once irrigation begins, a pulse water meter counts the actual volume that has passed through the pipe and sends that number back to the server. Instead of ending after issuing a command, the system confirms how fully the command was carried out, completing a closed-loop control structure.

The procedure is similar to assembling parts according to an instruction manual, then checking by hand one last time that the final screw has actually been tightened.

In Japan, the agricultural machinery company Kubota commercialized this system under the name ZeroAgri. Soil sensor information from inside a greenhouse and weather forecast data are sent to the cloud, where the server calculates and automatically dispenses optimal quantities of irrigation water and fertilizer. In mulched citrus orchards, the same soil-moisture visualization technology is used in the opposite direction. The purpose is not to conserve water, but to deliberately dry the soil to an appropriate degree, place the trees under mild stress, and thereby increase the fruit's sugar content.

The technology for supplying more water and the technology for supplying less operate for different purposes on the same sensors and the same cloud.

Several numbers show whether the system is working properly. Classification models that predict whether irrigation is needed or which type of fertilizer should be used are graded with metrics such as accuracy, precision, and recall. EcoGrow's 97.39 percent is one example of an accuracy measurement. Models that predict continuous values such as soil moisture measure the difference between observed and predicted values with Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). No matter how good those figures are, whether the technology actually helped farming must be measured again by other standards.

These are Water Use Efficiency (WUE), which indicates how much crop is produced with 1 cubic meter of water, and Nitrogen Use Efficiency (NUE), which indicates how much crop biomass is added with 1 kilogram of nitrogen fertilizer. The actual amount of water and fertilizer saved comes closer than the accuracy of an AI model to measuring the value of this technology.

The practice of delivering only a specified amount directly in front of the roots is called fertigation (the delivery of fertilizer dissolved in irrigation water to the root zone). The word combines irrigation with fertilization. Unlike conventional methods that spread fertilizer over the entire soil surface and then apply water, fertigation sends fertilizer dissolved in water through drip lines only to the narrow zone actually occupied by roots. In a trial of soilless strawberry cultivation in Italy, sensor-based fertigation management improved both resource use and crop performance (Bonelli et al., 2024).

Field trials in northeastern China also reported increased efficiency in the use of both nutrients and water when drip irrigation and fertilization technologies were combined (Fan et al., 2020). When water is applied across an entire field, fertilizer is dispersed just as widely, and nitrogen that leaks deep into soil beyond the reach of roots is simply lost. Precise delivery to a narrow zone prevents this loss at its source.

Efforts to determine nutrient status without removing leaves extend beyond photographs taken by unmanned aerial vehicles. A trial of water-saving tomato cultivation under protection, meaning roofed facilities such as plastic greenhouses or glasshouses, explored a method of estimating leaf nitrogen content without damage using hyperspectral imaging equipment that combines visible and near-infrared light (Hu et al., 2025). Monitoring water-saving irrigation together with leaf nitrogen status makes it possible to detect sooner when water and nutrient deficiencies begin to overlap.

The conventional method of picking a leaf and placing it in a laboratory analyzer takes time, and during that interval the crop has already moved to its next stage. If nitrogen status can be estimated merely by illuminating a leaf with a camera, irrigation and fertilization decisions can be made at once rather than being delayed by a day or two.

Another trend serves small farms that cannot afford expensive servers. Research is examining the installation of low-cost, low-power edge AI and TinyML (AI models made small and light enough to run even on very small microcontrollers) on farms with limited resources. Instead of sending data to a cloud server every time and waiting for a response, decisions are completed immediately inside a small device placed in the field. This kind of locally self-contained AI is most effective in remote areas where communications fail easily and on farms where internet charges are burdensome.

Research on smart irrigation systems combining IoT devices and machine learning has repeatedly confirmed that inexpensive sensors and microcontrollers alone can automate irrigation (Kathilkar et al., 2026). Equipment prices have already fallen substantially. The remaining task appears to be increasing the number of people who know how to operate that equipment.

Sensors left in the field for long periods become less reliable over time. Soil that receives frequent fertilizer applications accumulates more salts, while moisture and microbial films adhere to sensor surfaces, gradually reducing the sensitivity of moisture and electrical-conductivity sensors. A reading that was accurate when the sensor was first installed may become lower or higher than the true value several months later. This is drift, the phenomenon in which measurements gradually deviate from their original reference point over time.

If drift goes undetected for several weeks, the system may continue ordering water to be applied to soil that is already wet enough. To prevent this, research is under way on self-calibrating virtual sensor technology, which uses software to realign sensor values automatically, and on sensors made from corrosion-resistant nanomaterials. The effort resembles replacing a window that clouds after long exposure to soil with a material that does not cloud at all, rather than wiping it periodically.

The area of land that a single sensor can represent is smaller than one might expect. In a broad open field or a large greenhouse, soil properties, slope, and the density of crop foliage differ by zone. At the same moment, one zone may be wet while another is dry. If an entire field receives the same amount of water based on a single point sensor, wet areas become wetter and dry areas remain dry. Even within one greenhouse, the rows near the entrance receive different sunlight and wind from the rows farther inside. A grower cannot safely rely on one sensor merely because the crop is under the same roof.

Multiscale fusion technology, which integrates multispectral images from unmanned aerial vehicles and spatial interpolation with ground-sensor data in real time, is therefore regarded as a next-stage task.

Lower equipment prices do not make adoption effortless. Components such as the ESP32 have become much cheaper, but combining them and retuning a model for a particular farm still requires considerable work. The burden of learning screens and numbers wholly unlike familiar farming practices forms a real barrier to adoption. Attempts to lower this barrier add intuitive mobile interfaces, AI chatbots, and multilingual voice guidance. The cost of becoming comfortable with the machine remains far greater than the cost of buying it.

A model trained only on previously collected data becomes unstable when it encounters a situation it has never experienced. During an unprecedented heat wave or sudden drought, a model fixed by historical data may produce an absurd answer. To correct this weakness, researchers are studying online adaptive learning, in which the model continuously absorbs incoming field data and updates itself.

Combining this approach with a biophysical crop model such as DSSAT or AquaCrop, which simulates the crop's complete physiological response, makes multiobjective optimization possible, satisfying both the goal of conserving water and the goal of protecting yield. No matter how sophisticated sensors and AI become, their opponent is weather, which arrives each time with a different face.

Calculations that satisfy multiple objectives are not confined to theory. They are first validated in test plots at actual growing sites. In repeated controlled trials, one section of a greenhouse receives water in the conventional way while an adjacent section receives only the amount calculated by AI. The yield and water use of the two sections are then compared. Only after such trials accumulate can researchers distinguish an upper-bound figure observed under specific conditions from the average reduction that most farms can expect.

Without a procedure that verifies the numbers candidly, as in the Nebraska case, no new technology can easily earn farmers' trust.

Measurements of physical quantities across the soil continuum, growth-stage modeling, inference by multiple machine-learning models, and decisions in the cloud accumulate into a single command to open or close a valve. Decisions that conserve water and fertilizer are entrusted to the numbers sent by each sensor in the soil and to the code that reads them. The next morning, the grower checks on a mobile phone how many liters were delivered and at what hour the night before, and puts on boots to walk into the field only when a graph shows an abnormal spike.

The grower's role has narrowed from making every irrigation decision to reviewing the system's decisions after the fact.

Sources and References

Academic Papers

1. Ajay Kumar, C. V. R., Supriya, L., Navya Sree, P., & Santosh Kumar, M. (2026). Adaptive machine learning framework for precision driven smart irrigation optimization. World Journal of Advanced Research and Reviews, 30(1), 421-428. https://doi.org/10.30574/wjarr.2026.30.1.0841

2. Bonelli, L., Montesano, F. F., D'Imperio, M., Gonnella, M., Boari, A., Leoni, B., & Serio, F. (2024). Sensor-based fertigation management enhances resource utilization and crop performance in soilless strawberry cultivation. Agronomy, 14(3), 465. https://doi.org/10.3390/agronomy14030465

3. Chen, M., Cui, Y., Wang, X., Xie, H., Liu, F., Luo, T., Zheng, S., & Luo, Y. (2021). A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. Agricultural Water Management, 250, 106838. https://doi.org/10.1016/j.agwat.2021.106838

4. EcoGrow Development Team, & Sri Lanka HORDI. (2026). EcoGrow: Cloud-enabled machine learning driven fertilizer recommendation platform for horticultural vegetable crops in Sri Lanka. International Journal of Advanced Networking and Applications, 17(6), 7236-7243. https://doi.org/10.35444/IJANA.2026.17607

5. Fan, J., Lu, X., Gu, S., & Guo, X. (2020). Improving nutrient and water use efficiencies using water-drip irrigation and fertilization technology in Northeast China. Agricultural Water Management, 241, 106352. https://doi.org/10.1016/j.agwat.2020.106352

6. Hu, C., Zhao, T., Duan, Y., Zhang, Y., Wang, X., Li, J., & Zhang, G. (2025). Visible-near infrared hyperspectral imaging for non-destructive estimation of leaf nitrogen content under water-saving irrigation in protected tomato cultivation. Frontiers in Plant Science, 16, 1676457. https://doi.org/10.3389/fpls.2025.1676457

7. Kathilkar, S., Chavan, A., Sahu, K., Phalke, D., & Shedbale, S. (2026). Smart irrigation system using IoT device and machine learning. International Journal for Research in Applied Science & Engineering Technology, 14(5), 83111. https://doi.org/10.22214/ijraset.2026.83111

8. Li, H., Mei, X., Wang, J., Huang, F., Hao, W., & Li, B. (2021). Drip fertigation significantly increased crop yield, water productivity, and nitrogen use efficiency compared with traditional irrigation and fertilization practices: A meta-analysis in China. Agricultural Water Management, 244, 106534. https://doi.org/10.1016/j.agwat.2020.106534

9. Pradasari, N. I., Wahyudi, E., Atimi, R. L., Muhammad, A., Darmanto, & Pratiwi, I. (2026). IoT and machine learning-based smart watering model for water optimization in vegetable gardens. Zetroem, 8(2), 140-147.

10. Ramya, S. Y., Aravinda, T. V., Krishnareddy, K. R., & Ramesh, B. E. (2026). AI-driven self-adaptive smart irrigation system using IoT, computer vision and machine learning. International Research Journal on Advanced Engineering Hub, 4(6), 4074-4080. https://doi.org/10.47392/IRJAEH.2026.0526

11. Shriwas, M. K., Sindhi, K., & Thakre, P. (2026). Deep learning-based multilingual smart farming system for crop recommendation and nutrient monitoring. African Journal of Applied Research, 12(5), 260-280.

12. Sun, G., Hu, T., Chen, S., Sun, J., Zhang, J., Ye, R., & Liu, J. (2023). Using UAV-based multispectral remote sensing imagery combined with DRIS method to diagnose leaf nitrogen nutrition status in a fertigated apple orchard. Precision Agriculture, 24(6), 2522-2548. https://doi.org/10.1007/s11119-023-10051-7

Lectures and Video Materials

1. World AgriFood Innovation Conference (WAFI). (2026). WAFI2026 Global AgriFood Innovation Roadshow: China Agricultural University Special Session [Video]. YouTube. https://www.youtube.com/watch?v=SS2qVDa11nw

2. 農業農村情報通信環境整備準備会. (2025). Climate Change and the Future of Agriculture: The Importance of Data-Driven Agriculture Using ICT and IoT [Online Seminar Video]. YouTube. https://www.youtube.com/watch?v=iqLmEFT1wZo

3. 農林水産省(MAFF). (2025). Chugoku-Shikoku Regional Smart Agriculture Promotion Forum 2025 [Video]. YouTube. https://www.youtube.com/watch?v=16rooPtFOMM

4. Agricultural, Life & Environmental Sciences. (2026). Bentley Lecture in Sustainable Agriculture (2025) | AI, Machine Learning and the future of soils. YouTube. https://www.youtube.com/watch?v=ct0VrqKTmVU

5. Daugherty Water for Food Global Institute. (2025). Artificial Intelligence for Climate Smart Solution and Resilient Agriculture. (Archive). YouTube. https://www.youtube.com/watch?v=n1MYEcAeMRU

6. Global Soil Biodiversity Initiative. (2025). Harnessing AI to measure soil biodiversity. YouTube. https://www.youtube.com/watch?v=AT29crIuxow

7. 気象ビジネス推進コンソーシアムWXBC. (2026). FY2025 NARO Mesh Agricultural Weather Data Analysis Challenge! YouTube. https://www.youtube.com/watch?v=w1SrS3iwA4E

Reports and News

1. The Fence Post. (2026). 2026 Spring Homeland: Where artificial intelligence meets water stewardship in agriculture. https://www.thefencepost.com/news/2026-spring-homeland-where-artificial-intelligence-meets-water-stewardship-in-agriculture/
2. ScienceDirect. (2025). [Materials related to a comparative study of AI irrigation scheduling and field trials in Nebraska]. https://www.sciencedirect.com/science/article/pii/S2772375525002151
3. arXiv. (2026). [Materials related to trends in deploying low-cost, low-power edge AI and TinyML in agricultural fields]. https://arxiv.org/pdf/2603.15085

4.2
Spatiotemporal Data Fusion and Distribution and Quality Management

At four in the morning, a refrigerated truck pulls into the unloading bay of an agricultural produce logistics center in the southern part of the country. Each palletized box of strawberries carries an electronic data logger about the size of a fingernail, and the moment the truck door opens, the number on its display jumps from a value near freezing to nearly ten degrees Celsius, then slowly falls again after the door closes a few minutes later. Every one of these brief temperature fluctuations is a real factor that shortens the strawberries' remaining shelf life.

From the moment the strawberries leave the field until they enter the consumer's refrigerator, sensors measuring temperature, humidity, and gas concentrations exchange data as they change from one stage to the next.

The problem is that satellites, drones, and ground sensors collect data on different cycles. A satellite returns to the same orbit once every four to ten days and scans hundreds of hectares at a glance. A drone operates only on days when someone decides to fly it, but it photographs areas as small as the palm of a hand in fine detail. Ground sensors installed in fields and greenhouses produce readings without pause at intervals of a few seconds. When the records from the three devices are overlaid, their time and spatial scales differ, so they cannot be read as one coherent picture.

These differences in scale are precisely why yield forecasts and disaster early warnings in precision agriculture often miss the mark.

The industry calls the work of aligning these scales a climate-smart agriculture, or CSA, data pipeline. After the optical and multispectral imagery from satellites, high-resolution imagery from drones, and real-time readings from ground sensors are spatially and temporally aligned on the same coordinates and time scale, hybrid deep learning and graph neural networks, or GNNs, which learn from a mesh structure that represents each field as a node and relationships with adjacent fields as lines, combine them into a single flow.

The fused data identifies areas likely to suffer drought stress or pest and disease spread in advance, then calculates suitable planting and harvest times and expected yields under that year's weather scenarios.

Looking at several models now in use makes the structure of this network easier to grasp. A hybrid model that combines AMSTF, the Adaptive Multispectral Spatiotemporal Fusion model, and a GNN with a long short-term memory neural network, or LSTM, brings together multispectral images captured by satellites and drones at different times with ground-based weather and soil time series to estimate spatiotemporal changes in crop growth and final yield with high precision.

A fusion model that connects the three algorithms XGBoost, CatBoost, and LSTM uses time series of weather variables such as temperature, rainfall, and humidity, pesticide input, and the normalized difference vegetation index, or NDVI, to predict yield through regression. Another pipeline using a vision transformer, or ViT, converts images from satellites, drones, ground sensors, and even workers' smartphone cameras into small patches, learns through self-attention by comparing how closely they relate to one another, and produces field-level crop diagnosis maps and climate-response prescriptions.

In tests of these models, coefficients of determination showing how closely measured and predicted values agree commonly range from 0.85 to 0.95, and one report found that a spatial misclassification rate of 9.8 to 12.3 percent, caused by incorrectly dividing crop boundaries into separate zones, fell to 5.4 percent. Behind these numbers lies the demanding alignment work required to put readings from three kinds of sensors with different scales onto one common basis.

The term self-attention may sound unfamiliar, but a classroom group project provides an accessible analogy. The process resembles each student deciding how closely to listen to the others and gradually revising his or her own opinion. The model determines how strongly one area in a satellite image relates to another area in a drone image and to readings from ground sensors, assigns weights to those relationships, and uses the weights in its final decision. It differs from conventional statistical models because the data teaches the model what deserves closer attention instead of following rules set in advance by people.

One screen offers an easy way to understand how the overlaid map is used in practice. The farm manager's tablet displays the entire field as a mottled map of green, yellow, and red, and a red zone marks an area where water stress detected by satellite has been confirmed by signs of curling leaves in close-up drone images. If ground sensors support that finding by showing that soil moisture in the zone has actually fallen below the threshold, the model warns that pests or disease are likely to begin spreading from that area within three days.

The practical substance of a precision agriculture data pipeline is that three layers of evidence from satellites, drones, and ground sensors must come together to produce this one warning.

This pipeline runs on the physical equipment of three systems: satellites, drones, and ground sensors. On the satellite side, multispectral satellites such as Sentinel-2 and Landsat scan more than 1,000 hectares of farmland at once and regularly derive vegetation indices such as NDVI, GNDVI, and SAVI as well as soil moisture distributions. The strength of satellite sensors is that they can monitor an area of this size for roughly 20 percent of the cost of drones. Drones, by contrast, inspect small areas in much finer detail.

At a ground sampling distance of less than 5 centimeters per pixel, they capture near-infrared and red-edge bands and sometimes hyperspectral imagery, placing zones that look unusually wilted, zones overrun by weeds, and differences in leaf morphology within a field onto a single map. The resulting map is called a prescription map, and unlike the conventional practice of applying the same quantity across an entire field, it tells the farmer to intervene only where needed and only as much as needed.

On the ground, sensors placed throughout greenhouses and open fields measure temperature, humidity, carbon dioxide concentration, soil moisture, electrical conductivity, or EC, acidity, or pH, and light in real time. A total of 40 years of historical weather records, weekly forecasts, finely gridded agricultural weather data, soil survey data, and a database containing more than 30,000 crop varieties are then brought together through a cloud pipeline.

Compared with the period when farmers looked only at readings from one field, they can now open a nationwide weather map and a table of varietal traits at the same time and locate that field within them. Farm machinery has changed as well. Japanese cases have been reported in which specialized farm equipment, such as machines that trim tea plants, is fitted with a PLC, a small computer for machine control, GPS, and automatic steering sensors to upload operating time, travel distance, speed, working height, and working width automatically to the cloud.

In a tea plantation management project involving Kawasaki Kiko, Ochiai Cutlery Manufacturing, and Terada Seisakusho, accumulated operating records made it possible to visualize differences in work characteristics and field conditions among farms. Satellites and drones look down from the sky, while these machinery records preserve what amounts to the tracks left on the ground, making them useful later when tracing distribution histories.

Once data with such different characteristics has been collected in one place, it must be processed. Gaps caused by cloud cover in satellite images and missing values left when sensors briefly stop working are filled and noise is filtered out using a statistical technique called an adaptive Kalman filter. The data is then normalized so that values fall between 0 and 1, and augmentation is performed by rotating or flipping images to expand the sample and keep the training data from being skewed to one side, completing the preparation required before it enters the model. That covers the story while crops are growing in the field.

From the moment a crop is harvested and placed in a box, the same logic is applied to a different target. The difference is that the system now measures decay instead of growth and tracks temperature and ethylene gas instead of vegetation indices, while the principle of combining scales from different devices to reach a decision remains the same.

A data logger attached to a strawberry box measures only temperature. Commercial cold-chain management boxes, however, increasingly carry not only thermometers but also hygrometers, carbon dioxide sensors, ethylene gas sensors, and hydrophobic indicator labels that change color to show freshness. Near-infrared spectrometers that determine sugar content and acidity without cutting the flesh and 3D vision cameras that detect surface damage have joined this equipment. What these devices share is that they do not touch or damage the surface at all.

They can read the sugar content, damage, and freshness of a tray of strawberries inside a truck without opening the box.

A representative model that combines the signals produced by these devices into one decision is the cooperative-attention convolutional recurrent network, or Co-ACRN. Reported in 2026 by researchers at Jiangnan University in China, the model simultaneously accepts time-series data for temperature, humidity, carbon dioxide, and ethylene accumulated during cold-chain transportation and images of hydrophobic indicator labels.

The cooperative-attention module at the front aligns the temporal axes of the images and sensor streams, while the self-attention module at the back reviews the aligned information, reflects the overall context, and produces a real-time freshness score for fruits and vegetables without damaging their surfaces. The formula for assigning a freshness score is straightforward. It subtracts from one hundred the percentage of the total surface area occupied by defects. A product with no defects scores close to one hundred, while one that is half spoiled falls to around fifty.

This type of multisensor fusion is not limited to the case at Jiangnan University. Methods that combine temperature, relative humidity, ethylene, carbon dioxide, vibration received by a box, location data, image features, and spectroscopic signals have become a common approach in shelf-life prediction models. Industry assessments often identify three applications as suitable for immediate early adoption: computer-vision grading, Internet of Things-based cold-chain monitoring, and shelf-life prediction.

Cameras and sensors can be added on top of existing logistics facilities without replacing them, but changing the cultivation method itself takes much longer. Even when the theory appears impressive, the technology must be placed on aging hardware such as pipes, refrigerators, and trucks, so adoption tends to proceed first where intervention is easier.

The US cold-chain equipment company DeltaTrak has described predictive shelf-life analytics as changing the way cold-chain freshness is managed. Such predictive shelf-life analysis commonly gathers sensor readings from refrigerated containers in real time and informs shippers and retailers in advance how much time remains before each box reaches a store shelf. A different approach is used for fruits such as dates, whose interiors are difficult to assess from appearance alone.

Research has reported reading the combination of odor molecules emitted as the flesh ripens through multichannel gas sensors and feeding the values into a machine-learning model to estimate shelf life. Ripe and spoiled fruit release different combinations of odor molecules, and the machine reads these differences as patterns in the electrical signals from gas sensors.

Grading brings another combination of methods into view. For Le Lectier, a Western pear brand from Niigata Prefecture in Japan, an image-segmentation model called Mask R-CNN combined with GCBlock cuts out defects and contaminated areas on the pear's surface at the pixel level. The proportion of the extracted contaminated area is not converted directly into a grade but first passes through a method called fuzzy logic. Fuzzy logic is a computational method for handling boundaries that cannot be divided cleanly, such as hot and cold.

It is easy to classify three percent contamination as clearly high grade and 30 percent as clearly low grade, but for ambiguous values in between, probabilities are distributed across several grades in a way similar to the visual judgment of a human inspector. The resulting value enters a fuzzy inference system containing rules derived from expert experience, which assigns the final grade automatically.

The final grades are divided into four levels, from Aka-shu, the top grade marked with a red stamp, through Ao-shu one level below it, Ryo at an acceptable level, and Mu-in, which is difficult to sell as a commercial product. Cameras and fuzzy rules have taken over work once done by people who spent all day turning pears one by one and stamping them. Because each fruit has different characteristics, the industry does not use one uniform model.

DenseNet and EfficientNet-family models are applied to fruits such as mangoes, tomatoes, and kiwifruit, which differ in skin thickness, color, and softening rate, to grade both external appearance and internal ripeness. Surface defects determine a pear's grade, whereas color matters more for tomatoes and firmness matters more for kiwifruit, so the criteria themselves differ by fruit.

Results from tests of these assessment models tend to produce fairly high numbers. Classification performance indicators such as accuracy, precision, recall, and F1 score often range from 95 to 99 percent, and mAP and IoU, which measure how precisely defect areas are segmented, are reported as well. The mAP value averages, across several thresholds, how closely the defect location found by the model overlaps the actual defect location, while IoU divides the overlapping area of two regions by their combined area, so a value closer to one indicates a more precise segmentation.

By numerical measures, the scores assigned by machines have reached a level that compares favorably with the visual sorting once performed by people.

This change appears to extend beyond reducing labor. Human inspectors concentrate differently in the morning and afternoon, and habits acquired through repetition can cause them to grade the same pear slightly differently from day to day. Cameras and fuzzy rules do not become tired and maintain the same standard throughout the day and throughout the season. It remains to be seen, however, whether machines can fully match the subtle sense that people acquire through years of experience, such as noticing that a fruit looks sound on the outside but may be soft within.

Once graded, agricultural products move directly into inventory and traceability management. Produce harvested from fields or greenhouses is grouped by basket or steel rack, and each unit receives a barcode, QR code, or smart tag. A smart agriculture project in Kaohsiung, Taiwan, is a representative case. The project uses tags attached to blue plastic baskets or steel racks to upload field-batch inventory to a database in real time and automatically links greenhouse environment alerts to the government's agricultural production and distribution traceability system.

Improved space utilization for multilevel cultivation racks and logistics warehouses has been reported as well. Each tag provides a way to verify the source and travel route of a box.

This tracking function becomes useful when an incident occurs. In a case introduced by the TaiwanICDF in 2025, production data from farms and data from food ingredient suppliers were linked through a single commercial-area map platform. If a food safety problem arises in a school meal, the system immediately traces which farmer's field supplied the produce in that meal. It then identifies all other schools that received products from the same farm, allowing preventive action to be taken at once. The method has shifted from finding a cause after an incident to identifying the related scope before the incident spreads.

Scanning a produce batch number brings up on screen the field and harvest date of the box and the trucks and logistics warehouses through which it passed. Misdistributed books or manufactured goods can be recalled, but spoiled produce cannot be taken back once it has reached the table. The value of a traceability system therefore depends less on cleanup after an incident than on how quickly it can contain the incident within a narrow scope before it spreads.

A shelf-life prediction value by itself remains only partial information. Modular systems are now emerging that place shelf-life prediction, the traceability required by the US Food Safety Modernization Act, and carbon-emissions accounting on one data foundation so that producers, shippers, and retailers can view the same screen. The practice of logging into several separate systems has shifted to accessing one data foundation.

Postharvest management technology centered on predictive analytics around 2018 has since shifted its emphasis toward blockchain and digital twins, while TinyML, a lightweight form of artificial intelligence, supports traceability and information disclosure at sites with limited power or communications infrastructure. As shelf life, history, and carbon emissions are displayed together on one screen, the risk of freshness deterioration and gaps in traceability increasingly appear for the same box at the same moment.

In logistics warehouses, deep-learning object-detection models such as YOLOv5 and YOLOv8 count boxes entering and leaving in real time and connect the results to a digital twin, a virtual model that replicates the entire warehouse. Even while seated at an office monitor, a manager can check what is stacked on each rack and in what quantity, as well as which zones have temperature or humidity readings outside the thresholds. If a problem appears, the manager can first test an intervention on the virtual model and take the corresponding action in the physical warehouse when the result looks acceptable.

A digital twin is a virtual model that reproduces the actual state of a logistics warehouse. Every new box entering the warehouse and every rack whose temperature rises slightly are reflected on the virtual screen in real time, and the manager can zoom in on any desired area. If the numbers on a physical thermometer and the virtual screen begin to diverge, that discrepancy can signal a sensor failure or communications delay. This makes it possible to view the physical warehouse and the virtual model at the same time.

The temperature value that jumped at the predawn logistics center ultimately appears as one point somewhere on this virtual screen. A brief spike each time the truck door opens may look trivial to the human eye, but on the digital twin it remains as a record that allows the manager to compare how many times the truck opened its door that day and whether its temperature fluctuated more often than other trucks. If one truck shows an unusually poor pattern, the record can prompt an inspection of its refrigeration equipment. A pattern invisible when the numbers were scattered as individual readings becomes apparent only after they accumulate.

The Internet of Plants, or IoP, SAWACHI cloud in Kochi Prefecture, Japan, extends this idea across an entire region. It connects environmental data from more than 1,600 greenhouses and shipping data from more than 3,000 farms to WAGRI, Japan's national agricultural data platform, and runs sixteen analytical models to forecast the region's total shipments and schedule distribution. Shipping information once recorded farm by farm has expanded into data used to forecast total shipments across Kochi Prefecture and adjust distribution schedules.

Compared with the previous practice in which each farm estimated its own shipments for the following week, this approach is closer to viewing supply and demand across the entire region at once and setting the order accordingly.

Several recurring metrics in these cases can be summarized. They include the space utilization rate, which measures how fully multilevel cultivation racks and logistics warehouses are filled; the production and distribution traceability rate, which measures the share of cases successfully traced back to the producing farm and delivery destination when a problem occurs; and the reduction rate, which measures lower pesticide and energy use achieved through precise prescriptions and management.

Results from cultivation racks in Kaohsiung and Taiwan's school-meal traceability case are often reported by converting them into one of these three figures. Each number may look small by itself, but repeatedly using a little more warehouse space and detecting incidents several days earlier has a material effect on a farm's overall profit.

The need to train a separate model for pears, mangoes, and strawberries is a hidden cost in this field. A production area growing more than ten products must build more than ten models and retrain each one with new data every season. The burden of maintaining and updating as many models as there are crops is substantial, and companies other than large distributors may find the cost difficult to bear. Research therefore appears to be moving toward building one foundation model that works across many fruits and adjusting only the final decision stage for each fruit.

No matter how precise prediction, grading, and traceability become, a person makes the final decision. Even if a model assigns the Mu-in grade or warns that pests or disease will spread in three days, the field manager still decides whether to discard the box or when to apply a control agent. Data pipelines, grading models, and distribution traceability systems prepare the information needed for that decision rapidly and accurately, while people determine what to do and how to do it at the next stage. Technology has changed the quality and speed of the material available for judgment, but judgment itself remains a human responsibility.

The cases discussed so far span Taiwan, China, Japan, and the United States. Faced with the same problems in refrigerated trucks, cold storage facilities, and school cafeterias, each country has arrived at an answer by combining satellite data, cameras, and gas sensors in somewhat different proportions. Their languages and government organizations differ, but they share the goal of detecting the moment temperature rises and finding one spoiled box several days earlier. It is reasonable to say that the technology is converging in a similar direction across borders.

These systems, however, do not communicate smoothly with one another. Taiwan's production and distribution traceability system, Japan's WAGRI, and the US Food Safety Modernization Act record histories through different fields and code systems. When produce grown in one country is sold across a border, considerable manual work is still needed to transfer records between these different ledgers. Until a common standard allows one batch number to carry the same meaning in every country, a final step performed by people is likely to remain in cross-border traceability.

This picture is not yet complete. Differences in scale among satellite observations made every several days, occasional drone flights, and second-by-second ground-sensor records remain unresolved even by research on spatiotemporal interpolation and high-resolution reconstruction using generative artificial intelligence. High humidity inside cold-chain boxes, condensation on lenses, and fluctuating gas concentrations gradually distort readings from optical and gas sensors. Just as fogged eyeglasses blur a person's view, moisture prevents sensors from reading their original target values accurately.

Correcting the problem appears to require both calibration algorithms that adjust values automatically through software and surface materials resistant to moisture.

Questions about who owns data collected by farms, how it should be valued and traded, and how rules should protect farmers' rights while blockchain prevents falsification appear to require institutional solutions rather than technological ones. Another future task is to add explainable artificial intelligence to digital twins of logistics warehouses so that a machine can tell the field manager in words why it assigned a box the Mu-in grade. Farmers who receive only a grade without a reason may be frustrated because they cannot see the basis for the judgment.

As grading moves from people to machines, systems must explain what the machine examined and how it reached its decision.

Data from a refrigerated truck's temperature logger is linked in one pipeline to satellite imagery, a logistics warehouse's digital twin, and a cafeteria's traceability lookup screen. In this structure, a missing or incorrect record at one stage makes decisions at the remaining stages less accurate as well. The outstanding tasks are to unify the different traceability code systems used by each country and reduce measurement errors in moisture-sensitive equipment such as temperature, humidity, and gas sensors.

Sources and References

Academic Papers

1. Choudhari, A., Katkar, R., Gupta, M., Wathe, N., Kakde, S., & Poyam, P. (2025). Design of an integrated spatiotemporal deep learning framework for autonomous precision weed detection, treatment, and recurrence prediction in drone-based smart farming sets. International Journal of Applied Mathematics, 38(2s), 1339-1366.
2. Hoque, M. D. J., et al. (2024). Incorporating meteorological data and pesticide usage for crop yield prediction using machine learning. IEEE Access, 12, 40947-40961.
3. Huang, X., Zhu, Q., & Huang, M. (2026). Online detection system for freshness of fruits and vegetables based on temporal multi-source information fusion. Smart Agriculture, 8(1), 203-212. DOI: 10.12133/j.smartag.SA202505037.
4. Pasupuleti, M. K. (2026). Vision transformers for crop intelligence and sustainable precision agriculture. International Journal of Academic and Industrial Research Innovations (IJAIRI), 6(4), 1-45.
5. Wang, J., et al. (2026). Research progress and prospects of intelligent management and control technologies for facility vegetables. Smart Agriculture, 8(1), 105-120.
6. Nakazawa, Kensuke, & Yamazaki, Tatsuya. (2026). Study of a pear grading method using deep learning incorporating global contextual information and fuzzy inference. Journal of Agricultural Information Research, 35(2), 65-78.

Reports and Online Materials

1. CABI Digital Library. (2026). Review of trends in artificial intelligence-based postharvest management and shelf-life prediction. CABI Reviews. https://www.cabidigitallibrary.org/doi/10.1079/cabireviews.2026.0011
2. DeltaTrak. (2026). How predictive shelf-life analytics redefines freshness in the cold chain. Food Logistics. https://www.foodlogistics.com/software-technology/supply-chain-visibility/article/22954945/deltatrak-how-predictive-shelflife-analytics-redefines-freshness-in-the-cold-chain
3. National Center for Biotechnology Information. (2026). Material on a machine-learning model for estimating the shelf life of dates using multichannel gas sensors. PubMed Central. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12252097/
4. Preliminary paper on a modular data-based system integrating shelf-life prediction, traceability, and carbon accounting. (2026). arXiv. https://arxiv.org/pdf/2511.05920

Lectures and Video Materials

1. Kaohsiung Agricultural News. (2025). Year 114 Kaohsiung Smart Agriculture Annual Achievement Presentation and Seminar. YouTube. https://www.youtube.com/watch?v=1Edc6pP4HnU
2. Weather Business Consortium WXBC. (2025). 2025 NARO Mesh Agricultural Weather Data Analysis Challenge. YouTube. https://www.youtube.com/watch?v=w1SrS3iwA4E
3. TaiwanICDF. (2025). Third International Development Assistance Field Forum 2025: Exploring Global Smart Agriculture Development. YouTube. https://www.youtube.com/watch?v=0qUeP8A4Pro
4. Chabashira. (2026). Smart Agriculture in Tea Plantations: Kawasaki Kiko, Ochiai Cutlery Manufacturing, and Terada Seisakusho. YouTube. https://www.youtube.com/watch?v=zmFbOC5RcJU
5. Li Tsai-Jung. (2025). A88 Institute for Information Industry Satellite Precision Agriculture AI Introduction Workshop: Taoyuan AI Satellite Multispectral Imagery and Drone Applications. YouTube. https://www.youtube.com/watch?v=dtT3bafdKgw
6. World AgriFood Innovation Conference (WAFI). (2026). WAFI2026 Global AgriFood Innovation Roadshow: China Agricultural University Special Session. YouTube. https://www.youtube.com/watch?v=SS2qVDa11nw
7. World AgriFood Innovation Conference (WAFI). (2026). WAFI2026 Global Agrifood Innovation Roadshow: Session 2. YouTube. https://www.youtube.com/watch?v=DQrCqqyNAqY

8. Preparatory Committee for Agricultural and Rural Information and Communications Environment Development. (2026). Sharing Data Across the Region: The Front Line of Agricultural Data Use, Data-Driven Agriculture with WAGRI and the IoP Cloud (SAWACHI), and Next-Generation Facility Horticulture Systems. YouTube. https://www.youtube.com/watch?v=k8Qvv6CKB2A

9. Centre for Networked Intelligence, IISc. (2025). Webinar on Track-1: Crop identification using satellite data [IndiaAI Impact Gen-AI Hackathon]. YouTube. https://www.youtube.com/watch?v=9xi5_1rrJ9o

10. Sampurna Challenge. (2026). Post-Harvest Innovation, Startups & Agri Value Chain Solutions | Webinar | Sampurna 2026. YouTube. https://www.youtube.com/watch?v=0HTys4rPdwE

11. AgriBusiness Forum. (2026). AI, Data, Digital Transformation of the agrifood value chain. YouTube. https://www.youtube.com/watch?v=OMFBQz27YyU

12. Facebook Li Tsai-Jung Assistance for Technology Companies Entering Science Parks, Group Meal Robots, Electric Tricycle Logistics Food Trucks, Global Aiyu Jelly, Global Robots, and Global AI Genealogy Planning. (2025). A89 Private Institute for Information Industry Satellite Precision Agriculture AI Introduction Workshop, Taoyuan AI Satellite Multispectral Imagery and Drone Applications. Organizer: Institute for Information Industry; Co-organizer: King Young Technology Co. [FarmiSpace x Dahe Rice x King Young Drones x Institute for Information Industry] Joint Precision Agriculture Project. YouTube. https://www.youtube.com/watch?v=L27xkkSrSa0

13. Washington State Academy of Sciences (WSAS). (2026). AI for Post Harvest. YouTube. https://www.youtube.com/watch?v=hJlOrgNlQG4

Chapter 5

AI-Based Digital Phenotyping and Breeding Biotechnology

5.1

High-Throughput Crop Phenotyping

Suspended from rails in a greenhouse ceiling, a gantry system takes photographs from several angles with multiple cameras each time it passes a pot, then automatically combines those images to build a three-dimensional model of every plant. Work that once kept a person walking through the greenhouse with a tape measure all day is completed by the machine for thousands of pots overnight. The computational speed of reconstructing three-dimensional models from photographs has risen sharply in recent years.

A three-dimensional foundation model published in 2026 cut the average reconstruction time from 6 minutes 31 seconds to 1.58 seconds, while a three-dimensional Gaussian splatting method achieved a processing speed of 6.39 frames per second.

This use of machines to measure the outward characteristics of plants is called phenotyping. A phenotype is the way genetic information encoded in a seed appears in a mature plant as height, leaf size, stem thickness, signs of disease, and other observable characteristics. Phenotyping encompasses the entire process of measuring and recording those characteristics. For many years, people performed all of this work. Breeders placed tape measures against plants to measure height, judged leaf color by eye, assigned disease severity grades, and sometimes pulled up entire plants to weigh them.

Human measurement introduces errors that vary with the observer's judgment and condition on a given day. A single variety trial can involve thousands or tens of thousands of lines, putting continuous observation throughout the growing period beyond human reach from the outset. Breeders had long recognized this bottleneck, but they lacked an effective tool to break through it.

Sensor engineering, robotics, and artificial intelligence supplied that tool. Lasers and cameras measure plant bodies, unmanned aerial vehicles and wheeled robots travel through fields and greenhouses collecting photographs and video, and deep learning distinguishes leaves from fruit in the images. The technology that integrates these three elements is called high-throughput crop phenotyping, often shortened to HTP. If a person can measure one hundred pots in a day, HTP equipment can measure thousands in the same period without touching a pot or removing a leaf.

This technology invariably accompanies discussions of smart breeding because accurate measurement is the first step in selecting a superior variety, and HTP places that step in the hands of machines.

The central principle of HTP is to capture the three-dimensional structure and light-reflectance characteristics of plant populations growing in greenhouses and fields quickly and nondestructively, then convert those data into numbers within a framework in which genes, the environment, and outward characteristics interact. The same seed grows differently in a well-watered field and a dry one. Appearance is not determined by genes alone; it is shaped jointly by genes and the environment. Breeders call this relationship G×E×P, from genotype, environment, and phenotype.

HTP fills in phenotype, the final and exceptionally difficult term to measure in that relationship. Without phenotype data, even the best genetic map cannot be tested against what actually happens in a field.

Laser scanning is one of the most widely used methods for capturing three-dimensional structure. LiDAR emits laser light and measures, in microseconds, how long it takes for that light to strike an object and return, then calculates the distance. When millions of these distance measurements are plotted as coordinates, the individual points combine into a three-dimensional image that reproduces the plant's outward form. This collection of points is called a point cloud. Its principle resembles a mosaic that builds the outline of a face from countless sand-grain-sized pieces.

The PlantEye F600 from PhenoSpex is a leading example of a laser scanner, producing point clouds so dense that adjacent points are only 0.1 millimeter apart. The instrument includes sensors that measure reflectance in both the near-infrared and red bands, allowing it to capture three-dimensional form and calculate the Normalized Difference Vegetation Index (NDVI), an indicator of leaf health, at the same moment.

RGB-D depth sensors, which combine color images and distance information in a single frame; stereo vision, which calculates distance from the displacement between overlapping images taken by two or more cameras; and Multi-View Stereo (MVS) reconstruction, which assembles photographs taken from many angles, serve the same purpose. These methods resemble the way two human eyes perceive distance.

From the resulting point cloud, a computer automatically extracts measurements that people once took with a ruler. Plant height, leaf-covered area, leaf angle, stem thickness, and even the volume formed by overlapping leaves can now be calculated together by analyzing one point cloud, whereas each item once had to be measured by hand. If thousands of pots are measured at the same time every day, the system automatically plots growth curves showing how much each plant has grown from day to day. Ground scanners alone are not enough for fields that are vastly larger than greenhouses. Another eye looking down from the sky is needed.

That eye is the unmanned aerial vehicle, or UAV. A drone flying low over a field carries an ultra-high-resolution color camera with more than 20 million pixels together with a multispectral sensor that separately records five wavelength bands: blue, green, red, red-edge, and near-infrared. A hyperspectral sensor that divides light between 400 and 1000 nanometers into closely spaced bands and a thermal sensor that displays surface temperature in color may be added as well. The multispectral sensors made by MicaSense and commonly mounted on drones are representative instruments for capturing these five wavelength bands.

The collected images are converted into measures of vegetation cover across the field and vegetation indices such as NDVI, GNDVI, and SAVI. Diseased or dry leaves reflect less near-infrared light and fail to absorb red light, whereas chlorophyll in healthy green leaves absorbs red light and strongly reflects near-infrared light. The difference in reflectance between the two wavelengths therefore reveals leaf health as a single number. After flying at low altitude for several hours, one drone can map vegetation indices across tens of hectares and use color to mark the plots where growth has stalled.

Deep learning separates leaves, flowers, and fruit in photographs and point clouds. Models such as convolutional neural networks (CNNs), Vision

Transformers (ViTs), and graph neural networks (GNNs) perform this task. Tracing and cutting out the boundary of an object down to each pixel in a photograph is called instance segmentation. In practical terms, it means labeling one ear of corn, the ear next to it, and then the neighboring leaf as separate objects. When these segmented organs are photographed repeatedly from the same position over days or weeks, the rates at which leaves unfold and fruit expands can be tracked through time.

Those trajectories can then be used to estimate total mass, biomass, and final yield in advance.

In a maize field in Henan Province, China, instance segmentation is already moving through real rows of crops. A wheeled unmanned ground vehicle, or UGV, travels slowly between the maize rows and scans both sides with a large panoramic camera mounted at the front. Midday sunlight changes from moment to moment, and upper leaves conceal the ears below, making it difficult even for a person to count how many ears are present. The CornYOLO model deployed there uses the lightweight YOLO11n neural network as its backbone and adds a frame reconstruction module (FRM) to find ears that are only partly visible behind leaves.

Its mean average precision, or mAP, reached 91.5 percent, a level comparable to counting by eye. During a single circuit of the field, the robot automatically builds a table recording the number and location of ears on every plant. That table becomes source data for selecting yield-related traits.

A similar method is used to monitor flowering in maize and rice. When YOLOv5 and a Vision Transformer are applied to canopy images captured from a drone, they automatically distinguish the emergence and flowering status of pollen-producing tassels in three stages: not yet flowering, just beginning to flower, and fully flowering. Aggregating these results automatically produces a growth index showing the current developmental stage of the entire field. Flowering time, once judged by people walking the field every few days, can be checked again at much finer intervals. For breeders, flowering time determines which lines should be crossed.

Improving that criterion to day-level precision can change the success rate of a crossing design.

Not every farm can afford to develop a dedicated model such as CornYOLO, so open-source tools that anyone can download and use are developing alongside it. PlantCV v2 is an open-source program that scripts procedures for analyzing leaf color and shape in images. A user supplies photographs, and the program automatically extracts leaf area, canopy volume, and the proportion of discoloration. kmSeg divides colors into clusters using the statistical technique K-means to distinguish plants from the background, while PhenoImage processes large volumes of fluorescence images showing how chlorophyll receives light.

These tools allow even university laboratories with modest research budgets to perform phenotyping comparable to that of large companies.

Inside greenhouses, fixed systems running on gantries or rails have become standard. Equipment from companies such as LemnaTec, PhenoSpex, and Field Scanalyzer runs on rails installed in greenhouses or roofed experimental fields, automatically carrying thousands of pots or plots beneath scanners at scheduled times. Because the machines follow a predetermined order and timetable around the clock, they can measure the same pot at three in the morning under exactly the same conditions as at noon.

Human hands cannot replicate this repeatability, and it is precisely what phenotyping requires when the same plant must be photographed every day from the same angle and under the same lighting.

Conditions differ outside the greenhouse in broad, unroofed fields. Because the area is too large for rails, drones and ground robots equipped with real-time kinematic positioning, or RTK-GPS navigation, which fixes location to within 2 centimeters, take over. The drones scan broad areas from the air while ground robots examine ears up close from below the canopy. By dividing the work this way, they can survey an entire breeding trial containing thousands of lines within a few days.

Because the positional information exchanged by both machines lies in the same coordinate system, data seen from above by the drone can later be overlaid with data seen from the side by the robot.

Low-cost alternatives are available for sites with limited equipment budgets. Devices such as MVS-Pheno and LITERAL, as well as compact systems assembled from an Arduino microcontroller and a few cameras, can perform three-dimensional reconstruction. They cost only a fraction, perhaps a few percent, of the price of a commercial gantry system, yet they are adequate for a small laboratory or an individual farm to scan one of its own plots. Phenotyping can now begin without expensive equipment.

All the equipment discussed so far measures visible parts above the ground. Roots present a completely different problem because they are buried in soil. Devices such as MultipleX-Lab and RhizoPot address it by growing plants inside transparent tubes and passing light through the sides to photograph root shadows. Without pulling up the roots, these systems repeatedly record over time how many lateral roots extend in each direction and how deeply they descend. This makes them valuable for breeding root structures that withstand drought. They supply from below ground information that aboveground photographs cannot provide.

No matter how precise a camera is, its images contain soil, gravel, weeds, shadows, and other noise. In the first stage of removing this background, formulas that transform color space use indices such as ExG and CVI to separate green plants from everything else, together with topological filtering that follows the terrain. Point clouds contain too many points for a computer to handle efficiently in their original form. They are reduced to a manageable size through Plant-denoising-net (PDN), a neural network trained to remove plant-specific noise, and self-organizing-map down-sampling (SOM Down-sampling), which thins points evenly.

Every grain of soil and faint shadow must be filtered out so that only a clean representation of the plant remains for the next stage. Without this preparation, noise mixed into the raw point cloud disrupts the calculations that follow.

Multi-View Stereo reconstruction algorithms and Edge_MVS-Former, a Vision Transformer that detects boundaries more sensitively, are used together to reconstruct the three-dimensional form of an individual plant from several cleaned images. The resulting three-dimensional model passes through another instance-segmentation neural network, such as TrackPlant3D, which separates stems and leaves and organizes them by organ. Speed was the problem. Earlier methods took several minutes to turn dozens of input photographs into one well-formed three-dimensional model.

In a breeding facility that must image thousands of greenhouse pots every day, those minutes accumulated until they consumed the entire day's schedule.

Neural Radiance Fields, or NeRF, offered a clue to solving this speed problem. NeRF takes several photographs captured from different angles and trains a neural network to infer the color and light intensity at any point in three-dimensional space viewed from any direction. The network plausibly fills the gaps between images, including angles the cameras did not capture, and produces a much more natural three-dimensional scene than earlier methods. Yet each point on the screen had to be drawn by querying the neural network individually, so the method long carried a reputation for being accurate but slow.

A three-dimensional foundation model published in 2026 directly challenged that reputation. After pretraining on large volumes of photographs from multiple plant species, it reused that learning whenever photographs of a new plant arrived. It cut the average reconstruction time from 6 minutes 31 seconds to 1.58 seconds while preserving shape accuracy and phenotypic measurements.

Three-dimensional Gaussian splatting, or 3DGS, is another method that raised the speed. NeRF repeatedly asks a neural network for color values. By contrast, 3DGS models a plant surface in advance as hundreds of thousands of small, blurred, translucent blobs, then spreads them across the screen at once, like scattering tufts of dyed cotton, to complete the image. Once the position, color, and degree of blur of each blob have been calculated, a graphics card can render them directly without a query-and-response process through a neural network. In experiments against Instant-NGP and Nerfacto, 3DGS achieved 6.39 frames per second.

This result helped shift the center of gravity in three-dimensional representation from neural networks toward explicit structures of this kind.

Separating individual plants as distinct objects in images taken by several cameras remains difficult. PlantSegNeRF combines Neural Radiance Fields with a procedure that matches the same organs across images taken from multiple viewpoints, separating three-dimensional point clouds by individual plant even when there are few samples or the crop species has never been seen before. Researchers have produced methods that allow a robot to decide for itself which next camera angle will yield the most new information.

One such method, SSL-NBV, uses self-supervised learning, in which the robot learns without being given each correct answer, to choose the next camera position without a person specifying it every time. A five-stage process has been proposed for constructing field-scale point clouds at a physically accurate scale. It proceeds from data acquisition through semantic segmentation of plants and background, sparse reconstruction that first establishes a limited set of points, dense reconstruction that fills in the points, and finally the extraction of phenotypic traits.

Comparisons among this flood of methods soon followed. In 2026, Frontiers in Plant Science published a review comparing studies of three-dimensional crop phenotyping under field conditions with an emphasis on point clouds. Around the same time, Sensors published a review of three-dimensional reconstruction technologies as a whole. Both reviews identify the same trend: the center of gravity is moving away from reliance on a single expensive instrument and toward combinations of several inexpensive cameras and fast software.

Several metrics determine whether these methods are useful in practice. Mean average precision, or mAP, measures how successfully the system identifies targets such as maize ears or leaves. Intersection over Union, or IoU, measures how closely the boundary of a segmented leaf overlaps the real boundary. Root mean square error, or RMSE, expresses down to the millimeter how far a three-dimensional model constructed with a laser or camera departs from measurements taken with a ruler.

The time required for a small field computer to process a single image and the number of images processed per second determine whether the system can operate in real time. The coefficient of determination measures how closely weights or heights estimated from images match values obtained through direct destructive sampling. In several studies, this coefficient has exceeded 0.75, adding evidence that image-based measurements can replace destructive surveys. Traits classified by grade rather than number, such as flowering stage or disease severity, are evaluated separately with the F1-Score and a confusion matrix.

The photographs, point clouds, and numbers produced each day by cameras and robots vary in type and scale. Researchers need an orderly set of agreements to share these data. The basic framework is the FAIR principles. Data must be

Findable, Accessible, Interoperable among different programs, and Reusable without difficulty later. The practical way to follow FAIR is to attach metadata stating what was recorded, when, and where to every item, from raw sensor images and processed three-dimensional point clouds to shooting time and location and the final phenotypic measurements, then organize it all in a central database.

PHIS, the Phenotyping Hybrid Information System, is one example of these principles translated into a concrete system. PHIS adopts the standardized vocabulary of the Plant Ontology (PO) and assigns a unique label to every experimental subject, sensor, and event in a greenhouse or field. Moving one pot from one side of a greenhouse to the other, the appearance of a spot on a leaf, and a reading from a temperature sensor are all linked through unique persistent identifiers, or URIs. Data from an experiment conducted several years ago can therefore be compared with data from the current month by the same criteria.

This identifier system implements the interoperability and reusability required by FAIR at the level of experimental data.

OpenSILEX, which works in tandem with PHIS, is a software architecture that connects phenotypic data scattered across many locations and the web services used to manage them through one standard gateway. The international research community has separately defined the minimum information that must be recorded for a phenotyping experiment. The standard, called MIAPPE, or Minimum Information About a Plant Phenotyping Experiment, requires the experimental location and date, cultivation conditions, and measurement method to be documented without omission. Systems such as PHIS and OpenSILEX use it as part of their underlying architecture.

The ICASA Version 2.0 standard, which covers field-trial environments, farm-operation histories, and weather data, serves a similar purpose. These standards translate the different names used by laboratories in different countries and universities into a common language.

Phenotypic data organized in this way become valuable when they meet genomic data. Databases such as CropGS-Hub and PHENOPSIS DB bring genomic information and high-throughput phenotyping data together so that Genomic Selection models can connect and use both directly instead of searching for them separately. If phenotypic databases had not followed common standards, technologies for measuring appearance and technologies for reading genes would have used incompatible formats, making it impossible to connect phenotype data with genomic data in the first place.

In a greenhouse where the PhenoSpex PlantEye is in active use, one scanner passes the same pots each day and accumulates daily measurements of canopy height, leaf area, and leaf angle. These values are copied directly into a virtual greenhouse, a Digital Twin inside a computer, creating a second greenhouse on the screen that grows alongside the real one. The grower works backward from how much water and nutrients the virtual plants require and decides when and how far to open or close the nutrient-solution valves in the physical greenhouse. Instead of plants saying they are thirsty, changes in leaf angle and canopy volume speak for them.

In breeding trial fields planted with soybeans and winter wheat, a single drone carries a multispectral sensor and LiDAR together. When canopy-height information and leaf-color information are combined with appropriate weights, the aboveground biomass accumulated in the field and the amount of nitrogen it contains can be estimated with considerable precision. At harvest, a clustering method automatically counts soybean pods that overlap in the images. Breeders can therefore use a single drone photograph to screen lines that appear likely to yield well before they plow up the plots and weigh the crop.

Large research institutions working with flowers and fruit trees deploy OpenSILEX and PHIS together to manage the entire mass of data. Tens of thousands of color photographs and three-dimensional images accumulated each day, real-time sensor readings, and genotype information for each variety are gathered in one system. Labels based on the Plant Ontology standard eliminate confusion in which different research teams use different names for the same pot. With less time spent finding and sharing data, researchers can move directly to the next step of connecting AI breeding models to those data.

Standardization may not produce a dazzling result by itself, but it visibly shortens the time required for such a result to appear.

Despite these achievements, several problems remain unresolved. As crops grow and the canopy becomes dense, upper leaves conceal lower flowers, fruit, and even signs of disease. A single camera or laser scanner positioned above the canopy often cannot see lower traits at all. The most promising current solution appears to combine images taken from different angles by ground robots and drones with point-cloud reconstruction calculations that fill in hidden, unrecorded leaves from surrounding information.

Perfectly restoring the shape of a leaf missing entirely from an image remains difficult, but field practitioners judge that even an estimate is far better than nothing.

The sheer size of the data is another obstacle. A point cloud at 0.1-millimeter resolution or a hyperspectral image containing hundreds of wavelengths can expand to gigabytes per image, too large for a small computer standing at the edge of a field to process in real time. Down-sampling that thins the points, Knowledge Distillation that transfers the knowledge of a large neural network to a smaller one, and hardware acceleration that embeds a neural processing unit (NPU) directly inside the equipment are being studied side by side to reduce data volume or speed computation.

Lighter equipment can deliver answers immediately in the middle of a field, so this race to reduce computational weight is likely to continue for some time.

Outdoor weather frequently confuses artificial intelligence as well. Under strong direct sunlight, the surfaces of leaves reflect white and lose color information. Wind makes the same leaf appear in a different shape in every photograph. Passing clouds change brightness from moment to moment, causing a model that just learned from one image to lose its way in the next. A model trained under controlled light and wind in a greenhouse often loses accuracy sharply when moved directly into an open field. This phenomenon is called Domain Shift.

The principal methods for addressing it are to expand training data by artificially mixing diverse weather and lighting conditions or to retune the model to a new environment through Domain Adaptation. A model whose accuracy has not been tested outside the laboratory ultimately risks remaining a partial solution that works only inside a greenhouse.

Nearly all the technologies discussed so far are directed at leaves and fruit above the ground, making integrated phenotyping that tracks underground roots in the same pipeline the next major task. Explainable AI (XAI) is needed as well to show people which images caused an artificial-intelligence system to classify a plant as diseased or likely to yield well, and why. Techniques such as Grad-CAM and SHAP display a color-coded heat map of the regions in an image that a neural network examined when making a decision. Breeders need such visual evidence before they can trust a machine's judgment and plan the next crosses around it.

After a gantry has passed, what remains is not a handful of photographs but millions of coordinate values and the trait measurements extracted from them. Once these measurements are organized according to common standards, they can connect data from several greenhouses with data from laboratories in other countries. Breeders can then verify in numbers how a seed carrying a particular gene grows in a particular field.

Sources and References

Academic Papers

1. Xu, Q., Zheng, B., & Zhong, S. (2026). Artificial Intelligence Empowering Modern Agricultural Biological Breeding. Smart Agriculture, 8(3), 69-85.
2. Gao, G., Wang, Q., Song, L., Feng, H., Shi, L., Yang, H., Liu, Y., & Yue, J. (2026). Object Detection Method of Maize Ears Within Canopy Based on CornYOLO. Smart Agriculture, 8(1), 167-177. DOI: 10.12133/j.smartag.SA202509005.
3. Moldvai, L., & Nyéki, A. (2025). Innovative computer vision methods for tomato (Solanum Lycopersicon) detection and cultivation: a review. Discover Applied Sciences, 7, 975. https://doi.org/10.1007/s42452-025-07613-x
4. Harnessing artificial intelligence in plant breeding: Innovations in digital phenotyping and breeding methods. (2026). Plant Methods. DOI: 10.1007/s13007-026-01286-0.
5. Neveu, P., Tireau, A., Hilgert, N., Negre, V., Mineau-Cesari, J., et al. (2019). Dealing with multi-source and multi-scale information in plant phenomics: The ontology-driven Phenotyping Hybrid Information System. New Phytologist, 221(1), 588-601.
6. OpenSILEX. (2026). Plan de gestion logiciel.
7. 田中佑, 郭威, 山口友亮, 桂圭佑, 中嶌洸太, & 戸田陽介. (2025). "AI時代"の作物学研究. 日本作物学会紀事, 94(3), 329-330.
8. 大澤央, 齋藤正博, & 八木岡敦. (2026). 物体検出AIを活用したスイートコーン収穫適期予測技術の開発. 農業食料工学会誌, 88(2), 85-91.
9. arXiv. (2026). Ultra-fast crop reconstruction based on a three-dimensional foundation model with preserved phenotypic accuracy. https://arxiv.org/abs/2607.01753
10. arXiv. (2025). PlantSegNeRF: Three-dimensional instance point-cloud reconstruction of plants through channel-coupled Neural Radiance Fields based on multiview image instance correspondence. https://arxiv.org/pdf/2507.00371

11. arXiv. (2024). SSL-NBV: Robotic three-dimensional plant reconstruction through self-supervised next-best-view selection. https://arxiv.org/pdf/2410.14790

12. Frontiers in Plant Science. (2026). Comparative review of point-cloud-based three-dimensional crop phenotyping under field conditions. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2026.1731852/full

13. Sensors. (2026). Comprehensive review of three-dimensional reconstruction technologies for crop phenotyping. https://doi.org/10.3390/s26092730

Lectures and Video Materials

1. IPPN - International Plant Phenotyping Network. (2025). Dr. Michael Selveraj: AI-Powered Phenomics: Transforming Crop Science for Climate-Resilient Future. YouTube. https://www.youtube.com/watch?v=eoTvCz3YTzo

2. Greg Bronevetsky. (2025). Predictive Pattern Recognition of Plant Growth Traits in Simulated and Controlled Environments. YouTube. https://www.youtube.com/watch?v=uKepfZAnKbk

3. Bioingene. (2026). International Webinar- Biotechnological strategies to improve plant traits for enhanced productivity. YouTube. https://www.youtube.com/watch?v=duUd8ajVo44

4. Solvi. (2026). Webinar: Validating AI-Based Plant Counts with Real Field Data. YouTube. https://www.youtube.com/watch?v=B0KuJdRtmaQ

5.2

Predictive Breeding and Smart Biological Breeding Technologies

Crop breeding began with phenotypic selection: growers walked into their fields and selected superior individuals solely by what they could see. Mendel's laws of inheritance and statistics were later brought in to calculate performance across parents and offspring. Marker-Assisted Selection (MAS), which uses particular locations in genes as markers to identify desired traits in advance, then became established. Genomic Selection (GS) followed, reading tens of thousands of markers scattered across the entire genome at once. Selection accuracy increased, while the time required to grow one generation and move to the next became shorter.

This method gives breeders a value calculated in advance for how much a seed not yet planted in soil is likely to yield when it grows. The value is called the Genomic Estimated Breeding Value (GEBV). It is a prediction calculated solely from the genetic information carried by each seed.

Traits governed by many genes, such as yield, presented the problem. Unlike plant height or disease resistance, which may be determined by one or two genes, yield emerges from hundreds or thousands of genes with small effects interacting with one another. Those interactions appear differently depending on the field in which the same seed is planted. The same variety may perform poorly in a drought-stricken field and well in a field with ample water. This reciprocal influence between genes and the environment is called Genotype by Environment Interaction (G×E). Genomic Selection alone has clear limitations in resolving it precisely.

Smart biological breeding emerged to place information on the cultivation environment into the calculation alongside the genome and phenotype.

Smart biological breeding feeds three kinds of data, Genomics, Phenomics, and Enviromics, together into a single artificial-intelligence model. Genomics describes the genetic information a seed inherits. Phenomics describes the height, leaf color, yield, and other characteristics the seed actually exhibits as it grows. Enviromics describes the temperature, rainfall, and soil properties of the field in which the crop grew. Viewed separately, the three are only scattered pieces of information.

Once passed through multiple layers of an artificial neural network and gathered in a single Latent Space, a compressed space where different kinds of data can be placed on common coordinates and compared, their interconnections become visible. Before seeds are physically planted, the computer sketches how a given genotype will perform in a particular field. A breeder can select twenty lines from an on-screen ranking because this calculation has already taken place behind it.

As climate change makes droughts and heat waves more frequent, identifying in advance varieties that remain stable in particular environments has become a central breeding task. The Integrated Genomic-Enviromic Prediction (iGEP) system developed for this purpose feeds historical weather data, future weather projections, soil properties, and location information into a multilayer neural network together with a genomic map of Single Nucleotide Polymorphisms (SNPs), variations in which a single letter of the genetic sequence differs among individuals.

The target cultivation regions and climate zones are grouped as a Target Population of Environments (TPE). The system estimates in advance how well a particular genotype will endure and how much it will yield when extreme conditions such as drought or high temperatures strike within that TPE. It is now possible to gauge to some degree which varieties will withstand next summer's heat wave without planting them in a real field and waiting several months.

Other systems add a growth model that simulates the crop's entire physiology. Crop growth models such as APSIM and DSSAT use equations from physics and chemistry to represent the processes by which sunlight, temperature, and water become leaves and fruit. A crop growth model linked to a genomic prediction model, known as a genomic-crop model (CGM-G2P), simulates the physiological processes through which a particular gene ultimately affects yield. Older statistical models showed only correlations between numbers.

This approach instead makes predictions while revealing what a gene actually does inside a plant, allowing a breeder who receives a prediction to follow its basis one step at a time.

All these calculations require data. At the bottom layer of the data supporting predictive breeding lies the genome itself. Next-generation sequencing (NGS) and third-generation sequencing technologies that read longer regions in a single pass sequence the entire genome again to identify Single Nucleotide Polymorphisms, insertions and deletions (InDels), and Structural Variants (SVs) in which a large structure changes as a whole. Recognizing that one genomic map from one variety is insufficient, breeders construct and use Pangenome maps that overlay the genomes of multiple lines.

Transcriptomics showing how strongly genes are expressed, Proteomics showing protein abundance, Metabolomics showing metabolites, and Epigenetics data showing changes in expression without changes to the gene sequence itself are layered on top over time.

Enviromic data arriving from fields are equally dense. Internet of Things sensors installed in fields, nationwide gridded agricultural weather data, NASA POWER satellite data from the U.S. National Aeronautics and Space Administration, and soil-survey databases recording soil properties are combined. Together they build a time series of daily minimum and maximum temperatures, accumulated temperature, Vapor Pressure Deficit (VPD), which indicates how dry the air is, soil moisture, and concentrations of nitrogen, phosphorus, and potassium.

Researchers in different institutions and countries need matching formats before they can share these data. The four FAIR principles, which require data to be Findable, Accessible, Interoperable with other programs, and Reusable, have therefore become the standard for data management. With the adoption of BrAPI, the Plant Breeding API and a common specification for exchanging breeding data, along with ICASA, a standard for crop cultivation data, different research institutes can now exchange their accumulated data in standardized formats.

The collected data are loaded onto integrated platforms such as Breedbase, OpenSILEX, and CropGS-Hub and organized into a knowledge base where genomes, phenotypes, and environments can be searched from one screen. Knowledge scattered through papers and old breeding notebooks is not left buried. Sentence-analysis systems such as PlantConnectome scan tens of thousands of papers, automatically find statements linking genes to traits, and organize the relationships into a Knowledge Graph made of points and lines. Even a single observation handwritten in a notebook by a breeder who retired decades ago can now be retrieved with one search.

Accumulating data is not the same as identifying useful patterns within them. Early genomic prediction calculations were handled by statistical models such as GBLUP, or Genomic Best Linear Unbiased Prediction, and RKHS, or Reproducing Kernel Hilbert Space. Both solve equations to estimate how much each of tens of thousands of markers contributes to a trait, then sum those contributions to produce an individual's expected performance.

Their reach expanded with machine-learning techniques such as Random Forest, which narrows answers by branching through multiple questions; Extreme Gradient Boosting (XGBoost), which refines predictions by reducing errors step by step; Support Vector Machine (SVM); and LASSO regression. These statistical and machine-learning models remain useful for identifying the few hundred truly important markers among tens of thousands and filtering the rest as noise.

Deep neural networks came after the statistical models. SoyDNGP, used in soybean breeding, arranges genomic information like a three-dimensional image and passes it through a three-dimensional convolutional neural network (3D CNN) to predict yield, harvest time, and oil and protein content. DNNGP accepts several kinds of omics data at once. CropARNet was tested on fifty-three agricultural traits across rice, maize, cotton, and millet. Cropformer combines a convolutional neural network with Self-Attention, a calculation that learns where to place more weight within the input.

DPCformer handles distorted nonlinear relationships between genotype and phenotype. As these models appeared in succession, predictive performance rose step by step. They capture Epistasis, in which genes suppress or amplify one another, an effect that cannot be obtained merely by adding the effects of two or three genes.

Models such as AutoGP and the multi-trait MtCro have been applied to crops with long histories of accumulated data, including maize, rice, and wheat. In a single calculation, they produce results for yield, disease resistance, and grain quality together. As network depth increases, a model is more likely to fall into

Overfitting, memorizing minute features of the training data so that it performs well on test data but fails in the real field. Dropout, which randomly disables neural-network nodes, and techniques such as L2 regularization, which restrain computational values from growing excessively large, are used to prevent it.

Newer models increasingly report not only their accuracy but how stable their performance remains in fields they have never seen.

Artificial intelligence for reading genomes has gone a step further by treating genetic sequences themselves as language, much as a model learns human language. AgroNT, the Agro-Nucleotide Transformer, and FloraBERT are pretrained through Self-Supervised Learning, without labeled answers, on the sequences of noncoding regions before and after genes that regulate expression, namely promoters and enhancers. Once trained, these models can receive the sequence of a plant species they have never encountered and immediately identify gene-expression levels and the activity of regulatory regions. PlantCaduceus goes further.

It trained simultaneously on the genomes of sixteen distantly related angiosperm species and has identified more accurately than supervised models why particular genetic sequences have been conserved through long evolutionary history.

Artificial intelligence has moved beyond prediction into direct genome editing. Introduced by its developers as the first system of its kind for plants, OpenCRISPR-1 learns from vast existing data on the three-dimensional structures and nucleotide sequences of CRISPR-Cas gene-editing systems, then designs novel cleavage enzymes and guide RNAs (gRNAs), short genetic sequences that carry the scissors to the desired location. This approach can reduce Off-target effects, in which an unintended location is cut, compared with using cleavage enzymes taken directly from nature.

Because gene editing must cut precisely at the desired position, this accuracy determines success or failure.

A vast Breeding Foundation Model that unifies all these predictive and design capabilities is under development as well. Like large language models that learn human speech, a breeding foundation model follows a Scaling Law: performance improves by a predictable amount as model size, training-data volume, and computation increase together. Once the model crosses a certain threshold, Emergent Abilities appear and it suddenly performs tasks it could not do before. The point at which that threshold will be crossed cannot be known in advance.

One example is Chain-of-Thought reasoning that examines step by step how a combination of edits to several genes will affect an entire crop. Larger models, however, are not free from Hallucination, in which they confidently produce plausible but incorrect answers. Retrieval-Augmented Generation (RAG), which makes the model search an authorized breeding database before answering, is therefore used alongside them. Low-Rank Adaptation (LoRA) retunes only a few computational pathways instead of training the whole model again from the beginning, allowing specialized versions for a particular crop and region to be made without an expensive server.

The process by which a trained model produces predictions is broadly divided into two stages. In the offline stage, a vast combined dataset of genomics, phenomics, and enviromics is distributed across a server cluster containing hundreds or thousands of GPUs. Three-dimensional parallelization and ZeRO, an optimization method that reduces memory waste so large models can be trained on limited hardware, are used to train the model over several weeks. Once training is complete, the model moves to the online stage.

At a working breeding site, it receives genomic maps of candidate female and male parental lines together with weather scenarios for the target growing region. It then completes In-silico Cross Screening, experiencing the performance of tens of thousands of cross combinations inside the computer, in only a few seconds. The rankings filling the screen are the product of that several-second calculation repeated again and again through the night. An answer that once required planting a crop and growing an entire generation can now be obtained inside a computer before a seed enters the soil.

Several numbers test whether these predictions can be trusted in practice. The target for predictive accuracy, which indicates how closely actual performance and model predictions move together, is between 0.70 and 0.85. The benchmark for the coefficient of determination, the share of total performance variation explained by the model, is between 0.75 and 0.90.

Traits expressed as actual numbers, such as yield, heading date, and stem height, are checked again with mean absolute error (MAE), the average difference between predicted and observed values, and root mean square error (RMSE), the square root of the mean of those squared differences. A model's ability to identify diseased leaves is measured by precision, the proportion of leaves predicted as diseased that are actually diseased; recall, the proportion of truly diseased leaves found by the model; and the F1 score that combines them.

The precision with which gene-editing scissors cut the intended location is measured separately by on-target efficiency, the share of cuts at the intended position, and the Off-target rate, the share made at unintended positions.

Breeders watch another metric more closely: how quickly a variety improves in practice because of that accuracy. This value, called the rate of genetic gain, is calculated by multiplying selection intensity, predictive accuracy, and the remaining genetic variation for the trait, then dividing the product by the time needed to move from one generation to the next. The shorter the generation interval and the more accurate the prediction, the greater the improvement achieved within the same period. Heritability, the proportion of total trait variation attributable to genes, is examined as well.

Higher heritability means genes determine more of the trait than the environment does, making selection more effective. Common benchmark suites such as SeedBench have been created to pit large models against one another and compare, by standardized criteria, how well they perform gene searches, interpret phenotypes, and design crosses.

Evidence that these calculations have changed real fields comes from the International Maize and Wheat Improvement Center (CIMMYT), part of the Consultative Group on International Agricultural Research (CGIAR). Predictive breeding systems combining whole-genome marker data, multispectral indices measured by unmanned aircraft, and records of weather stress were applied to maize and wheat breeding programs across East Africa and Latin America. Under severe drought and high-temperature conditions, their accuracy in predicting yield and drought tolerance rose by approximately 15 percent to 25 percent over that of conventional statistical models.

The number of plots in outdoor field trials declined, and the time from crossing to obtaining the next generation fell by roughly half. Before the next growing season in an increasingly drought-prone Africa, breeders could screen which lines would survive before planting them.

SoyDNGP is already operating as a publicly available web service in soybean breeding. When a breeder uploads genomic data for a line, a three-dimensional convolutional neural network calculates yield, harvest time, and oil and protein content, then automatically identifies promising parental combinations. CropGS-Hub and the AutoGP platform for maize, rice, and wheat similarly integrate genotype and phenotype data in web databases and allow any researcher to access them and run virtual-cross simulations.

Data accumulated by one laboratory over several years can immediately be picked up and used by another institute on the other side of the world.

Syngenta and InstaDeep applied their jointly developed AgroNT to regulatory sequences across diverse plant species, ranking in advance how sensitively genes would respond to environmental change without planting a single seed. By calculating weights for thousands of potential regulatory regions at once, they produced molecular evidence needed to design climate-resilient varieties. The growing frequency of partnerships between seed companies and artificial-intelligence firms signals that this technology has moved beyond laboratory walls and begun to serve as a tool in commercial breeding.

This collaboration carries a different weight because it has entered real breeding pipelines instead of remaining only in papers.

Gene-editing tools designed by OpenCRISPR-1 successfully performed precise edits to genes responsible for crop drought resistance and disease resistance. Results showed fewer unintended edits than when naturally occurring cleavage enzymes were used unchanged. Selecting and correcting exactly the desired line and letter within the genome of a single seed has advanced beyond a possibility to a demonstrated result. Whether that result will persist for several generations in fields outside the laboratory will become certain only after several more growing seasons.

In Kochi Prefecture, Japan, a cloud platform called SAWACHI connects environmental data from more than 1,600 greenhouses and production records from 3,000 farms with WAGRI, the national agricultural big-data network. It runs artificial-intelligence models that predict growth and yield and gauges how well candidate breeding lines fit real farming conditions. In Aichi Prefecture, another kind of artificial intelligence is used to improve carnation varieties.

Researchers investigate epigenetic phenomena in which environmental conditions or the propagation process alter gene expression and change petal color, then apply that knowledge to select new varieties with the desired colors precisely. Pooling data from thousands of greenhouses and probing the genetic mechanism behind a single petal color differ in scale, but both proceed side by side within the same predictive-breeding framework.

The achievements are clear, and the unresolved problems are equally clear. Models trained in greenhouses or simulated computer environments may lose much of their predictive power when they confront unprecedented extreme weather or compound environments in which several stresses strike at once. This divide between virtual space and a real field is called the Sim-to-Real Gap. Domain Shift, in which judgment becomes unstable under conditions absent from training, causes it.

Reducing the gap requires a steady addition of data gathered from more diverse environments and an online adaptive learning system that continually updates the model with incoming field data.

The inability of a model to explain why it produced an answer is no small problem either. Deep neural networks achieve high accuracy but cannot translate their reasons for recommending a particular cross into biological evidence a breeder can understand. Such a model is called a Black-box because it delivers a judgment while hiding the path that led to it. Efforts to open that box include explanatory techniques such as SHAP and Grad-CAM, as well as studies that connect Knowledge Graphs to attach visible biological causal relationships to predictions.

However high the accuracy, a model is unlikely to win the full trust of an experienced breeder until its evidence can be seen.

Computational cost creates another barrier. Training and running a breeding foundation model with billions of parameters requires expensive GPU servers. Small research laboratories and small breeding companies that cannot afford such equipment have difficulty even approaching the technology. Quantization, which lightens a model by reducing its computational values to a few integers, and Knowledge Distillation, which transfers knowledge learned by a large model to a small one, are being studied as ways to lower this threshold.

Renting servers in the cloud is an option, but many companies are reluctant to entrust breeding data to outside servers. The ability to pay for a single server is becoming another line separating those who can reach the latest breeding technology from those who cannot.

The wall that ultimately remains is human. People who can work with advanced artificial-intelligence algorithms while understanding both molecular genetics and field agronomy are still rare, slowing the transfer of laboratory results into variety improvement in real fields. Laws and institutions have not kept pace with the technology in deciding how crops edited by artificial intelligence should be reviewed and approved or who should own the data and models.

However powerful the ability to predict and design a seed becomes, the path to the field remains blocked unless people and rules are ready to examine that seed and take responsibility for it.

Sources and References

Academic Papers

1. Xu, Q., Zheng, B., & Zhong, S. (2026). Artificial intelligence empowering modern agricultural biological breeding. Smart Agriculture, 8(3), 69-85. https://doi.org/10.12133/j.smartag.SA202602016
2. Harnessing artificial intelligence in plant breeding: Innovations in digital phenotyping and breeding methods. (2026). Plant Methods. https://doi.org/10.1007/s13007-026-01286-0
3. Shokor, F. (2025). Deep learning models for genomic prediction accounting for heterosis in crossbreeding systems and multi-trait selection (Doctoral dissertation). Université Paris-Saclay, INRAE, AgroParisTech.
4. Qiu, J., Liu, Y., Gao, X., Huang, Y., Zhang, H., Tian, F., Li, W., & Feng, Z. (2026). AgriAgent: End-to-end large model agent system architecture for agricultural environment control. Smart Agriculture, 8(2), 220-236. https://doi.org/10.12133/j.smartag.SA202507042
5. Yang, J., Yang, W., Yang, S., He, L., & Zhang, D. (2026). Intelligent Q&A method for crop diseases and pests using LLM augmented by adaptive hybrid retrieval. Smart Agriculture, 8(1), 52-61. https://doi.org/10.12133/j.smartag.SA202506026
6. 田中佑, 郭威, 山口友亮, 桂圭佑, 中嶌洸太, 戸田陽介. (2025). AI時代の作物学研究. 日本作物学会紀事, 94(3), 329-330.
7. APSIM crop model innovations for breeding and deep learning. (2026). in silico Plants. https://doi.org/10.1093/insilicoplants/diaa016
8. CropARNet: A genomic selection model based on self-attention and a deep residual network. (n.d.). ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2772899425000242

9. Cropformer: A crop-phenotype prediction model combining a convolutional neural network and self-attention. (n.d.). ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2590346224006448

10. DPCformer: Predicting nonlinear genotype-phenotype relationships with convolutional neural networks and attention. (n.d.). arXiv. https://arxiv.org/html/2510.08662v1

11. PlantCaduceus: A multispecies plant DNA language model pretrained on the genomes of sixteen angiosperm species. (n.d.). Computomics. https://www.computomics.com/news-reader/foundation-models-in-plant-breeding-an-introduction-copy.html

12. AI breeder: Artificial-intelligence research on genomic prediction for crop breeding. (n.d.). ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2949952623000109

Lectures and Video Materials

1. Plant-OpenCRISPR1. (2025). Plant-OpenCRISPR1: The world's first AI-designed genome editing platform for plants [Video]. YouTube. https://www.youtube.com/watch?v=DEkNOyybGtE

2. World Food Prize, & CIMMYT. (2025). Predictive breeding: Bridging genetics, artificial intelligence and agriculture (B. M. Prasanna) [Video]. YouTube. https://www.youtube.com/watch?v=FWcxeEC90JA

3. WAF. (2025). WAF interactive webinar on biotechnology and AI contributions to plant breeding [Video]. YouTube. https://www.youtube.com/watch?v=xkVT9D7Nqy0

4. 農業農村情報通信環境整備準備会. (2026). 地域でデータを共有しよう、農業データの利活用最前線、データ駆動型農業WAGRIとIoPクラウド(SAWACHI)でのNext次世代型施設園芸システム(第5回オンラインセミナー) [Video]. YouTube. https://www.youtube.com/watch?v=k8Qvv6CKB2A

5. 愛知県農業水産局. (2026). あいち農業イノベーションプロジェクト2027技術提案募集説明会 [Video]. YouTube. https://www.youtube.com/watch?v=jpkoNqscGu8

Support This AI Library Book

If this book was useful, you can support future translations,
public editions, and open AI knowledge projects through PayPal.

PayPal

https://paypal.me/kimkjkj

Scan the QR code or open the link above.