Policy Research Journal
ISSN (E): 3006-7030 ISSN (P): 3006-7022 Volume 4, Issue 3, 2026

ARTIFICIAL INTELLIGENCE AND COMPUTER VISION FOR AUTOMATED IDENTIFICATION, POPULATION MONITORING, AND EARLY WARNING OF AGRICULTURAL INSECT PESTS

Muhammad Zaheer Ehsan[1], Ambreen Fatima[2], Rameesha Amjad Ali[*3], Sajid Ali Shah[4]

[1]*University of agriculture Faisalabad*
[2]*Graduated from University of Jhang*
[*3]*University of Agriculture Faisalabad*
Assistant Professor Entomology, Agriculture College Panjgur

[1]zaheerehsan594@gmail.com, [2]amber.u0005@gmail.com, [*3]dr.rameesha.sidhu@uaf.edu.pk
[4]syedsajid@ucdmj.luwams.edu.pk

DOI: https://doi.org/10.5281/zenodo.23208691

Keywords
artificial intelligence, computer vision, agricultural pest management, deep learning, object detection, edge computing, smart traps, population monitoring, early warning systems, Integrated Pest Management

Article History
Received: 30 January 2026
Accepted: 15 March 2026
Published: 31 March 2026

Copyright @Author
Corresponding Author: *
Rameesha Amjad Ali

Abstract

Global agricultural production loses over 40% of its potential yield annually to biotic stressors, primarily insect pests, threatening food security and supply chain stability. Traditional pest management, reliant on manual field scouting and calendar-based chemical spraying, is labor-intensive, costly, and prone to observer bias, sampling errors, and reporting delays. This review synthesizes recent advances in Artificial Intelligence (AI) and Computer Vision (CV) for automated pest identification, population monitoring, and early warning systems. We analyze four benchmark datasets–IP102, AgriPest, Pest24, and GrainPest–highlighting domain-specific challenges including fine-grained morphological distinctions, multi-instar variation, long-tailed class distributions, and micro-target detection below 15×15 pixels. We evaluate deep learning architectures spanning handcrafted feature extractors, Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and one-stage real-time detectors. Refined ResNet-50 pipelines achieve 76.0% Top-1 accuracy on IP102, while specialized YOLO variants such as EP-YOLO, AEP-PEST, and AgriPest-YOLO reach mAP50 scores of 70.5–75.2% on dense micro-target benchmarks. We examine cyber-physical smart traps integrating edge computing hardware, anti-occlusion mechanical clearing systems, and low-power LoRaWAN telemetry for autonomous field operation. Furthermore, we assess microclimate-driven predictive models combining Growing Degree-Day phenological calculations with XGBoost and LSTM architectures, where five-day rolling input windows elevate emergence prediction accuracy to 86.3% while reducing false-positive rates to 11%. Key challenges include visual distribution shifts under field conditions, thermal constraints on edge processors, and extreme class imbalances affecting rare species detection. We conclude by outlining strategic pathways toward scalable, proactive Integrated Pest Management deployment.

1. INTRODUCTION

Agricultural production globally loses over 40% of its total potential yield annually due to biotic stressors, primarily insect pests, plant pathogens, and invasive weeds. This massive economic loss severely threatens global food security and supply chain stability across both developing and industrialized agricultural economies (Li & Wang, 2024). Traditional pest management strategies rely predominantly on periodic manual field scouting and calendar-based chemical insecticide applications. Manual monitoring requires trained field technicians to visually inspect crop canopies or physically visit widely scattered light and pheromone traps on a weekly or bi-weekly basis (Aziz et al., 2025). This human-dependent process is labor-intensive, costly, logistically challenging, and inherently prone to observer bias, sampling errors, and reporting delays. Consequently, prophylactic chemical treatments are frequently applied on rigid calendar schedules rather than in response to verified ecological thresholds. This chemical over-reliance accelerates the development of pesticide resistance among target pests, elevates production costs, causes non-target toxicity to beneficial pollinators, and leads to persistent environmental degradation (Lima et al., 2020).

To address these vulnerabilities, modern digital agriculture has integrated Artificial Intelligence (AI) and Computer Vision (CV) as automated sensing and analytical engines. Modern vision systems facilitate non-destructive, continuous, and high-throughput monitoring of field crop canopies and trapped insect specimens (Kapetas et al., 2025). By coupling deep neural network architectures such as Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and real-time object detectors with cyber-physical Internet of Things (IoT) field sensors, digital farming platforms can automatically detect, localize, identify, and quantify target pest species at scale (Batz et al., 2023).

The ultimate objective of automated visual monitoring is to facilitate a paradigm shift in crop protection, transitioning from reactive, calendar-based chemical spraying to proactive Integrated Pest Management (IPM) (Dong et al., 2025). Achieving this transition requires linking vision-derived real-time population counts with microclimate environmental data such as canopy air temperature, relative humidity, and thermal heat unit accumulations to drive predictive machine learning forecasting models. These predictive systems issue localized early warnings regarding population surges or emerging infestation thresholds, allowing farmers to execute targeted biological or chemical interventions precisely when and where they are required (Cardim Ferreira Lima et al., 2020).

This review provides a comprehensive synthesis of AI and computer vision methodologies applied to agricultural pest management. Section 2 analyzes public benchmark pest datasets and their underlying taxonomic and ecological complexities. Section 3 evaluates state-of-the-art deep learning architectures for pest classification, localized object detection, and lightweight edge optimization (Partel et al., 2019). Section 4 explores edge computing hardware, mechanical trap anti-occlusion engineering, and low-power IoT telemetry. Section 5 evaluates microclimate-driven predictive modeling paradigms for early warning systems. Section 6 outlines key deployment challenges and future research frontiers, while Section 7 provides actionable strategic conclusions (Adetunji et al., 2023).

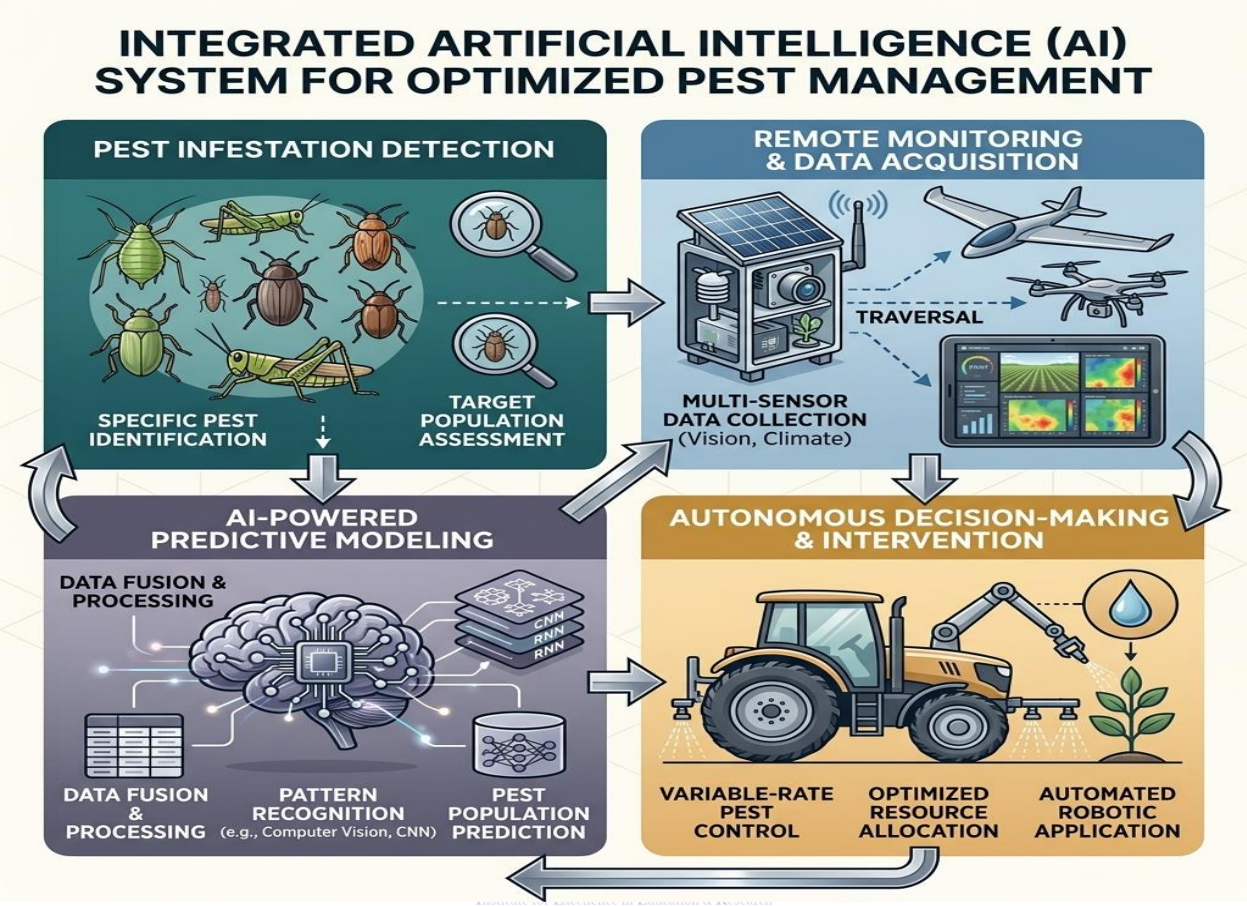


Figure 1: Conceptual Framework and Operational Workflow of an Integrated AI-Driven System for Automated Agricultural Pest Detection, Predictive Modeling, and Precision Intervention.

2. Benchmark Datasets and Taxonomic Complexity

The real-world efficacy and generalizability of deep learning models in agricultural computer vision are tied directly to the scale, taxonomic diversity, and visual fidelity of benchmark datasets. Unlike general-purpose computer vision datasets containing high-contrast, macro-scale everyday objects, agricultural pest benchmarks present domain-specific challenges (Sczyrba et al., 2017). These include subtle fine-grained morphological distinctions between species, multi-instar developmental variations, high intra-class variance, complex field crop backgrounds, and severe class imbalances. Benchmark datasets have evolved through three structural paradigms: broad web-scraped classification benchmarks, domain-specific in-the-wild detection datasets, and specialized automated trap datasets focused on micro-target localization (Kim et al., 2026) .

2.1 The IP102 Dataset Benchmark

Introduced to address the limitations of early datasets constrained by small sample sizes and artificial laboratory conditions, IP102 represents a benchmark for insect pest recognition. IP102 encompasses 75,222 images partitioned across 102 distinct pest species, maintaining an average of 737 samples per class with a standard train-validation-test split ratio of 6:1:3 (Wu et al., 2019). The dataset is organized hierarchically into eight super-classes divided between Field Crops (Rice, Corn, Wheat, Beet, Alfalfa) and Economic Crops (Vitis, Citrus, Mango). A localized subset of approximately 19,000 images is annotated with bounding box coordinates to support multi-target

object detection research (Hettiarachchi, C. N., & Meegama, 2026).

The core challenge of IP102 stems from its natural long-tailed distribution, which mirrors ecological field realities where a few dominant species generate the vast majority of visual instances while minor species remain underrepresented. Furthermore, IP102 incorporates diverse growth stages including eggs, larvae, pupae, and adults within single species categories, forcing deep learning networks to learn representations that tolerate high intra-class variance while simultaneously disentangling fine-grained inter-class similarities between closely related insect taxa (Nurfauzi et al., 2023).

2.2 AgriPest: Domain-Specific In-the-Wild Detection

While web-scraped datasets contain noise and variable framing, AgriPest was established specifically for wild, in-field pest detection under uncontrolled canopy environments. Over a seven-year collection period, AgriPest accumulated 49,700 images spanning four major agricultural field crops: wheat, rice, corn, and rape (brassica). The dataset encompasses 14 core pest species annotated by professional agricultural taxonomists with over 264,700 precise bounding boxes (Wang et al., 2021).

Images in AgriPest were acquired using field equipment comprising height-adjustable stands, Charge-Coupled Device (CCD) camera modules set to a 4 mm focal length and f/3.3 aperture, mobile control clients, and environmental sensors measuring local temperature and relative humidity. This standardized acquisition strategy ensured variations in camera-to-target distance, variable natural lighting, leaf occlusion, and realistic insect poses, yielding a benchmark for tiny object detection in complex field canopies (Christakakis et al., 2024).

2.3 Pest24 and GrainPest: Automated Traps and Post-Harvest Environments

For automated multi-target monitoring within physical trap devices, Pest24 serves as a standardized benchmark for dense tiny object detection. Pest24 contains 25,378 field images capturing 24 distinct pest categories that cause widespread agricultural damage. Captured via automated light and pheromone trap devices, Pest24 contains dense clusters of micro-targets where individual insects frequently occupy fewer than 15×15 pixels, or less than 0.25% of the total image canvas area (Athanassiou et al., 2025).

In post-harvest and storage protection, the GrainPest dataset fills a specialized niche by focusing on insects located on the bulk surfaces of stored grains in agricultural granaries. GrainPest contains 16,358 images and 66,372 annotated instances across six primary stored-grain pest species (Ma et al., 2024). A critical feature of GrainPest is its explicit inclusion of "hard negative samples" non-pest visual distractors such as cracked grain kernels, mold spots, shadows between kernels, and blackened ear bases which are essential for training detectors to minimize false-positive trigger rates in granary storage environments (Badgujar et al., 2026).

Table 1: Structural Comparison of Benchmark Agricultural Insect Pest Datasets

Dataset	Total Images	Target Classes	Primary Vision Task	Dataset Characteristics & Domain Focus
IP102	75,222	102	Image Classification / Bounding Box Object Detection (~19k subset)	Web-collected & field images; hierarchical taxonomy (8 super-classes); natural long-tailed distribution; high intra-class growth stage variance.
AgriPest	49,700	14	Tiny Object Detection & Multi-Species Localization	Field crop canopy imaging (wheat, rice, corn, rape); captured via custom CCD hardware; 264.7K expert bounding box annotations.

Pest24	25,378	24	Dense Multi-Target Tiny Object Detection	Automated physical trap imaging; extremely small target scale (<15x15 pixels); high spatial density; uniform background plate.
GrainPest	16,358	6	Granary Bulk Surface Pest Detection	Bulk grain surface (corn, paddy, wheat); 66,372 instances; includes hard negative distractors (mold spots, kernel shadows).

3. Computer Vision Architectures for Pest Detection and Classification

Computer vision models applied to agricultural pest monitoring have evolved from classical handcrafted feature extractors to deep convolutional networks, attention-guided transformers, and optimized real-time object detectors (Tetila et al., 2020).

3.1 Handcrafted Features versus Deep Convolutional Neural Networks

Early automated pest identification systems relied on manual feature engineering, combining localized spatial descriptors such as Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and Speeded-Up Robust Features (SURF) with shallow classifiers like Support Vector Machines (SVM) or Random Forests (Fooladgar & Kasaei, 2020). While combination descriptors (such as HOG paired with SURF) achieved acceptable performance (~90% accuracy) on small, isolated laboratory datasets, they failed when evaluated against large-scale, complex benchmarks like IP102 due to an inability to model non-linear variations in field lighting, pose, background clutter, and multi-instar morphology (Huang et al., 2022).

The adoption of deep Convolutional Neural Networks (CNNs) resolved these limitations by learning hierarchical visual representations directly from raw image pixels. Baseline comparative evaluations on the IP102 benchmark demonstrated a performance divergence between training deep models from scratch versus fine-tuning models pre-trained on large-scale datasets such as ImageNet. For example, a ResNet-50 architecture trained from scratch achieved an accuracy of only 35.7%, whereas fine-tuning the pre-trained counterpart elevated accuracy to 49.4% (Lin et al., 2020).

This performance gap underscores the role of transfer learning in agricultural domains, where pre-trained weight initialization provides foundational representations of edge, texture, and shape primitives. However, early fine-tuned CNN baselines exhibited a wide divergence between Top-1 Accuracy (49.4%) and Geometric Mean Accuracy (G-mean: 31.5%) (Shamsipour et al., 2024). Mathematically, for a dataset with K classes where Acc_i denotes the class-specific accuracy for class i, G-mean is defined as:

$$G\text{-mean} = \sqrt[K]{\prod_{i=1}^{K} \text{Acc}_i}$$

The lower G-mean value highlighted the vulnerability of standard cross-entropy loss functions to class imbalance, as the network biased its predictions toward majority categories at the expense of rare pest species (Nugroho, 2018).

To counter these limitations, advanced optimization pipelines have been introduced. Incorporating regularization and optimization techniques specifically Mixup data augmentation, label smoothing, Exponential Moving Average (EMA) weight tracking, and Stochastic Weight Averaging (SWA) elevated fine-tuned ResNet-50 performance on IP102 to 76.0% classification accuracy with a weighted F1-score of 0.75 (Attallah, 2025).

3.2 Vision Transformers and Hybrid Attention Frameworks

Vision Transformers (ViTs) introduce global self-attention mechanisms that capture non-local dependencies across image patches, contrasting with the localized receptive fields of standard

CNNs. In comparative benchmarks, ViT architectures (such as ViT-B/16) demonstrated strong capacity for modeling global spatial context and complex multi-scale reasoning (Moorthy et al., 2025). On the IP102 dataset, ViT-B/16 achieved a test classification accuracy of 69.17%, outperforming lightweight CNN baselines such as MobileNetV2 (53.8%). However, this accuracy gain incurs significant computational cost: ViT-B/16 requires 87.3 million parameters (349 MB memory footprint), compared to MobileNetV2's 4.2 million parameters (16 MB memory footprint), posing trade-offs for edge hardware deployment (Hatamizadeh et al., 2024).

Hybrid frameworks combine the computational efficiency of convolutional layers for local feature extraction with transformer blocks for long-range spatial context, optimizing multi-scale representations without incurring prohibitive parameter overheads (Ding et al., 2022).

3.3 One-Stage Real-Time Object Detectors and Micro-Target Innovations

In real-time multi-target tracking, the You Only Look Once (YOLO) framework series dominates practical field applications. One-stage detectors frame bounding box regression and class probability estimation as a single unified tensor mapping, enabling high-speed processing essential for edge deployment (Xie et al., 2025).

When applied to dense, tiny pest targets in benchmarks like Pest24 and AgriPest, standard YOLO backbones often suffer from feature loss during aggressive spatial downsampling. To preserve micro-target details, specialized architectural modifications have been developed. Spatial-to-Depth Convolution (SPD) modules replace traditional strided convolutions and max-pooling layers with depth-space transformations, preserving fine-grained spatial pixel features of sub-15-pixel insects prior to feature aggregation (An & Shi, 2024). Efficient Multi-Scale Attention (EMA) modules group channels into sub-features to retain spatial information across parallel sub-networks, suppressing background noise from trap plates or leaf surfaces without destroying tiny target signatures. Additionally, Global Spatial Pyramid Pooling Fast (GSPPF) blocks enhance receptive field fusion across extreme scale variations, capturing both isolated tiny insects and larger aggregated insect clusters (Ma et al., 2025).

To address bounding box regression inaccuracies on dense, overlapping micro-targets, researchers integrated the Wise-IoU v3 ($WIoUv3$) loss function. $WIoUv3$ replaces traditional Distance or Complete IoU loss functions with a dynamic, non-monotonic focusing mechanism, formulated as:

$$L_{WIoUv3} = r_W R_{WIoU} L_{IoU}, \quad r_W = \frac{\beta}{\delta \alpha^{\beta-\delta}}$$

where β represents the outlier degree of the prediction box, and α, δ are tuned hyper-parameters. This dynamic weighting reduces harmful gradient updates from extreme geometric outliers while preventing low-quality anchor alignments from dominating training on dense insect collections (Zhang et al., 2024).

Models incorporating these modifications such as EP-YOLO (70.5% mAP50), AEP-PEST (75.2% mAP50), and AgriPest-YOLO (71.3% mAP50) outperform standard object detection baselines (Faster R-CNN, baseline YOLOv4/v8) on dense agricultural benchmarks. Furthermore, adaptive multi-model detection ensembles combining modern YOLO backbones (v5, v8, v10) with Non-Maximum Suppression (NMS) fusion strategies have reached mAP50 scores up to 97.22% on target datasets (Jin et al., 2026).

Table 2: Comparative Performance Metrics of Deep Learning Frameworks on Benchmark Datasets

Benchmark	Architecture / Model	Methodological Novelty / Strategy	Primary Performance Metric
IP102	ResNet-50 (Baseline)	ImageNet Fine-Tuning Baseline	49.4% Top-1 Accuracy (31.5% G-Mean)
IP102	ResNet-50 (Refined)	Mixup, Label Smoothing, EMA, SWA	76.0% Top-1 Accuracy (0.75 Weighted F1)

IP102	MobileNetV2	Lightweight Convolutional Backbone	53.80% Test Accuracy
IP102	Vision Transformer (ViT-B/16)	Global Self-Attention Mechanism	69.17% Test Accuracy
IP102	Adaptive YOLO Ensemble	Fusion of YOLOv5, v8, v10 via Adaptive NMS	97.22% mAP50 (98.58% Accuracy)
Pest24	EP-YOLO	Spatial-to-Depth (SPD) + EMA Attention	70.50% mAP50 (47.30% mAP50-95)
Pest24	AEP-PEST	Compact Multi-Scale Feature Architecture	75.20% mAP50 (147.3 FPS inference)
Pest24	YOLO-LCE	Asymmetrical Channel Splitting + Ghost Head + WIoUv3	1.7% mAP50 improvement over YOLOv8 baseline
AgriPest	AgriPest-YOLO	Global Spatial Pyramid Pooling Fast (GSPPF) + CLA	71.30% mAP50

4. Cyber-Physical Smart Traps and Edge Computing Infrastructure

Transitioning computer vision models from static offline evaluation to continuous field deployment requires integrating physical traps, edge compute modules, multi-sensor payloads, and low-power telemetry into unified cyber-physical systems (Cao et al., 2021).

4.1 Edge Computing Hardware and Sensing Payloads

Smart traps operate as autonomous monitoring nodes deployed directly in crop fields, orchards, granaries, or livestock facilities. Edge computing hardware executes neural network inference locally on the device, removing the requirement to continuously stream raw, uncompressed video feeds to cloud servers. This local processing model significantly reduces cellular telemetry costs, lowers power consumption, and eliminates network latency barriers (Beckman et al., 2016).

Edge processing hardware architectures are typically selected based on operational power constraints and required computational throughput. Single-board microcomputers, such as the Raspberry Pi 3/4 and Raspberry Pi Zero series, serve as versatile edge units due to their low cost, low idle power consumption, and rich peripheral interface options (Bui et al.,2024). For high-throughput edge tasks requiring dense object detection at higher frame rates, AI-accelerated processors such as the NVIDIA Jetson Nano provide specialized Tensor Cores suitable for running quantized vision models. Ultra-low-power microcontrollers, such as the ESP32-based LilyGO-T-7670, are deployed to manage system duty-cycling, power state transitions, and environmental sensor polling (Capra et al., 2019). Sensing payloads combine high-resolution optical devices with environmental monitoring units. Optical capture relies on high-resolution RGB CMOS camera modules (e.g., Raspberry Pi Camera Module v3) or industrial CCD sensors configured with fixed focal lengths to capture fine morphological structures (Albanese et al., 2022). To minimize unnecessary image capture and power draw, optoelectronic infrared (IR) beam barriers are installed at trap entrances. Interruption of the IR beam triggers the camera and illumination source only when an insect enters the chamber. Environmental sensor suites, such as digital HTU21D modules, continuously record ambient temperature and relative humidity. In specialized environments such as livestock housing, secondary gas sensors (e.g., MQ-9) monitor ambient carbon monoxide and methane levels (Joglekar et al., 2021).

4.2 Mechanical Maintenance and Anti-Occlusion Engineering

A persistent mechanical bottleneck in autonomous field trap operation is the rapid accumulation, decay, and physical layering of trapped insect specimens. On sticky boards or

standard funnel trap bottoms, dead insects accumulate over extended operational periods. This stacking causes severe visual occlusion and body deformation, degrading the identification accuracy of computer vision models over time (Xue et al., 2024).

To overcome visual decay and occlusion limitations, modern cyber-physical smart traps incorporate active mechanical clearing systems. In automated funnel traps, an affordable servomotor mechanism is attached to the detachable bottom plate of the collection bucket (Zhu et al., 2026). Operated by the edge microcontroller, the servomotor periodically rotates the bottom plate, dumping accumulated, sun-dehydrated insect carcasses out of the camera's view field. This automated physical clearing completely resolves insect overlap, extending autonomous operational lifetimes to several months without human intervention (Chang et al., 2025).

4.3 Low-Power IoT Telemetry and Network Architectures

Because smart traps are frequently deployed in remote agricultural environments lacking grid electricity or high-speed broadband, network protocol selection determines field deployment feasibility (Karaagac et al., 2019).

LoRaWAN (Long Range Wide Area Network) operates on sub-GHz license-free radio bands, delivering long-range wireless communications up to 10–15 km in open agricultural landscapes with minimal power consumption. LoRaWAN is suited for transmitting lightweight numerical telemetry such as vision-derived insect count logs, species identification indices, local temperature, relative humidity, and battery state-of-charge (Dongare et al., 2017). Where image validation is required, cellular modules (GSM/LTE-M/NB-IoT) transmit compressed thumbnail images directly to centralized cloud platforms. In greenhouse, granary, or facility-level deployments, local Wi-Fi mesh networks provide high-bandwidth local connectivity (Zhalmagambetova et al., 2025).

Executing deep learning inference at the edge and transmitting only structured metadata via LoRaWAN minimizes power consumption and communication overhead, enabling long-term solar-powered field operation (Casalbore et al., 2026).

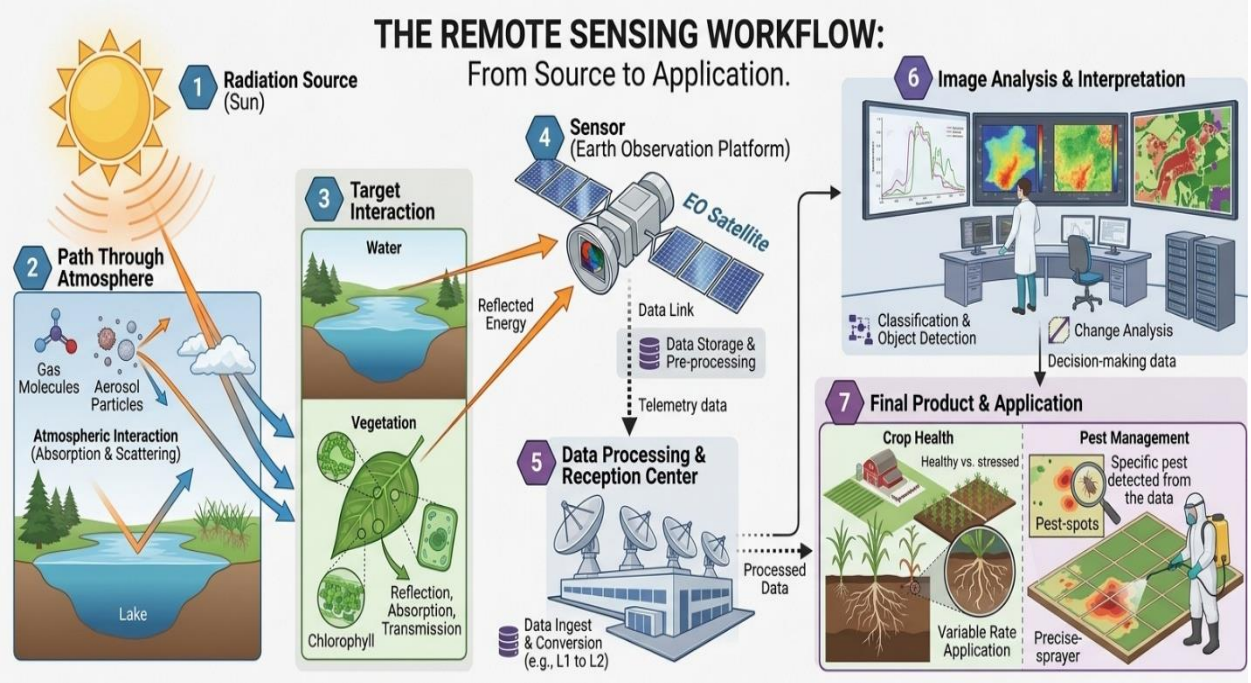


Figure 2: End-to-End Remote Sensing and Satellite Earth Observation Workflow for Crop Health and Precision Pest Management.

5. Microclimate Data Integration and Predictive Early Warning Models

While computer vision smart traps provide spatial and temporal counts of pest populations, vision data alone cannot predict future population spikes. Constructing true proactive early warning systems requires integrating vision-derived population metrics with microclimate environmental data within predictive machine learning forecasting frameworks (Majdoubi et al., 2025).

5.1 Thermal Time and Phenological Modeling

Insect development, physiological growth rates, and seasonal emergence dates are intrinsically governed by ambient thermal conditions. Phenological forecasting models rely heavily on the concept of Growing Degree-Days (GDD) or thermal time accumulation. GDD measures the accumulated heat units above a specific physiological base threshold temperature (T_{base}) required for a target insect species to progress through developmental stages (Hufkens et al., 2018).

The daily accumulation of Growing Degree-Days (GDD_d) is calculated as:

$$GDD_d = \max\left(\frac{T_{\max,d} + T_{\min,d}}{2} - T_{\text{base}},\ 0\right)$$

where $T_{\max,d}$ and $T_{\min,d}$ represent the daily maximum and minimum ambient temperatures, respectively. When cumulative thermal time reaches species-specific physiological thresholds, the early warning system forecasts key biological milestones, such as adult emergence, egg laying, or larval feeding surges (Chuine et al., 1999).

5.2 Machine Learning and Time-Series Forecasting Paradigms

Modern predictive early warning platforms integrate thermal calculations with advanced machine learning architectures to model complex non-linear relationships across microclimate variables (Bontempi ett al., 2012).

Decision-tree ensemble models, particularly XGBoost (eXtreme Gradient Boosting), excel at processing tabular microclimate features such as ambient air temperature, relative humidity, and solar radiation alongside historic vision trap counts. XGBoost models efficiently capture non-linear feature interactions with minimal computational training overhead. For complex temporal sequences, Long Short-Term Memory (LSTM) recurrent neural networks capture multi-day temporal dependencies in weather patterns and population dynamics (Kong et al., 2025). In comparative evaluations, LSTM models achieved strong precision in predicting temporal population trajectories, achieving Root Mean Square Error (RMSE) values as low as 0.119 in insect emergence forecasting. Furthermore, temporal convolutional networks and spatial transformer models link spatial trap networks with predictive weather grids to simulate regional pest density dynamics (Torres et al., 2021).

Empirical evaluations demonstrate that expanding predictive models from single-day inputs to multi-day chronological windows (such as 3-day or 5-day historical microclimate features) significantly improves forecasting accuracy. Shifting to a 5-day rolling input window elevated pest emergence prediction accuracy from 76.5% to 86.3%, while simultaneously reducing false-positive warning rates to 11%. This predictive window provides agricultural managers with actionable advance notice, enabling precise, targeted biological control applications or selective chemical spraying prior to catastrophic population surges (Hall & Rasheed, 2025).

6. Challenges, Open Research Directions, and Strategic Outlook

Despite progress in deep learning architectures and cyber-physical IoT hardware, several technical hurdles hinder wide-scale commercial deployment in field agriculture (Hossain et al., 2016).

Real-world field environments introduce visual distribution shifts that degrade the performance of models trained on static benchmarks. Dynamic canopy lighting, deep shadow foliage, water droplet reflections on camera lenses, dust accumulation, and variable weather conditions degrade visual input quality (Zhang et al., 2026). Furthermore, models trained exclusively on adult specimens struggle when encountering early-stage instars, damaged insect fragments, or non-target

arthropods. Developing domain adaptation frameworks, self-supervised representation learning, and synthetic data augmentation methods represents an active research direction (Rani & Mitra, 2026).

Hardware constraints pose another barrier to permanent edge deployment. High-performance edge processors (such as the Raspberry Pi 4 or Jetson Nano) experience significant thermal dissipation and power draw when executing continuous deep learning inference under direct solar exposure in tropical field environments. Overheating can trigger thermal throttling or hardware failure. Future hardware developments must emphasize hardware-aware neural architecture search (NAS), low-bit model quantization (INT8/INT4), channel pruning, and passive thermal enclosure engineering (Lichtenthaler, 2011).

Finally, natural ecosystems feature extreme long-tailed class imbalances, where minor or emerging invasive species are underrepresented in training datasets. Standard supervised learning biases neural networks toward majority classes, resulting in low detection rates for rare pest species. Addressing this requires zero-shot and few-shot learning algorithms, class-rebalancing loss functions, and point-guided prompt architectures capable of detecting novel target species with minimal human annotations (Cerulli & Reale, 2026).

Conclusion

This review demonstrates that AI and computer vision have fundamentally advanced agricultural pest management, enabling a paradigm shift from reactive calendar-based spraying toward proactive Integrated Pest Management. Benchmark datasets—IP102, AgriPest, Pest24, and GrainPest—have matured to reflect real-world complexities including long-tailed distributions, multi-instar morphology, and sub-15-pixel micro-targets. Deep learning architectures have correspondingly evolved: refined CNN pipelines achieve 76.0% accuracy on IP102, Vision Transformers capture global spatial context at higher computational cost, and specialized one-stage detectors (EP-YOLO, AEP-PEST, AgriPest-YOLO) attain 70.5–75.2% mAP50 on dense pest targets through innovations such as Spatial-to-Depth convolution, Efficient Multi-Scale Attention, and Wise-IoU v3 loss. Cyber-physical smart traps integrating edge processors, servomotor-based anti-occlusion mechanisms, and LoRaWAN telemetry now support months-long autonomous operation. Predictive early warning systems coupling Growing Degree-Day phenology with XGBoost and LSTM models achieve 86.3% emergence prediction accuracy using five-day rolling windows, reducing false-positive warnings to 11%. Despite these advances, translation to widespread commercial deployment faces persistent barriers: visual distribution shifts under dynamic field lighting, thermal throttling of edge processors under solar exposure, and severe class imbalances that suppress rare invasive species detection. Future research must prioritize domain adaptation, self-supervised and few-shot learning, hardware-aware neural architecture search, low-bit quantization, and passive thermal engineering. Interdisciplinary collaboration among entomologists, computer scientists, and agricultural engineers will be essential to deliver scalable, equitable, and sustainable automated pest surveillance systems that safeguard global food security.

REFERENCES

Adetunji, C. O., Olaniyan, O. T., Anani, O. A., Inobeme, A., Osemwegie, O. O., Hefft, D., & Akinbo, O. (2023). Artificial intelligence and automation for precision pest management. In *Sensing and artificial intelligence solutions for food manufacturing* (pp. 49-70). CRC Press.

Albanese, A., Nardello, M., & Brunelli, D. (2022). Low-power deep learning edge computing platform for resource constrained lightweight compact UAVs. *Sustainable Computing: Informatics and Systems*, *34*, 100725.

An, J., & Shi, Z. (2024). YOLOv8n-enhanced PCB defect detection: A lightweight method integrating spatial–channel reconstruction and adaptive feature selection. *Applied Sciences*, *14*(17), 7686.

Athanassiou, C. G., Morrison, W. R., Steiskal, V., & Riudavets, J. (2025). Monitoring of stored product insects and decision support systems at the post-harvest stages of durable agricultural commodities: a review. *Entomolgia Generalis*.

Attallah, O. (2025). Multi-domain feature incorporation of lightweight convolutional neural networks and handcrafted features for lung and colon cancer diagnosis. *Technologies*, *13*(5), 173.

Aziz, D., Rafiq, S., Saini, P., Ahad, I., Gonal, B., Rehman, S. A., ... & Nabila Iliya, M. (2025). Remote sensing and artificial intelligence: revolutionizing pest management in agriculture. *Frontiers in Sustainable Food Systems*, *9*, 1551460.

Badgujar, C. M., Swaminathan, S., & Gerken, A. R. (2026). Electronic nose for agricultural grain pest detection, identification, and monitoring: A review. *Journal of Stored Products Research*, *115*, 102840.

Batz, P., Will, T., Thiel, S., Ziesche, T. M., & Joachim, C. (2023). From identification to forecasting: the potential of image recognition and artificial intelligence for aphid pest monitoring. *Frontiers in Plant Science*, *14*, 1150748.

Beckman, P., Sankaran, R., Catlett, C., Ferrier, N., Jacob, R., & Papka, M. (2016, October). Waggle: An open sensor platform for edge computing. In *2016 IEEE sensors* (pp. 1-3). IEEE.

Bontempi, G., Ben Taieb, S., & Le Borgne, Y. A. (2012). Machine learning strategies for time series forecasting. In *European big data management and analytics summer school* (pp. 62-77). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bui, T. A., Lee, P. J., Liang, C. S., Hsu, P. H., Shiu, S. H., & Tsai, C. K. (2024). Edge-computing-enabled deep learning approach for low-light satellite image enhancement. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 4071-4083.

Cao, K., Hu, S., Shi, Y., Colombo, A. W., Karnouskos, S., & Li, X. (2021). A survey on edge and edge-cloud computing assisted cyber-physical systems. *IEEE Transactions on Industrial Informatics*, *17*(11), 7806-7819.

Capra, M., Peloso, R., Masera, G., Ruo Roch, M., & Martina, M. (2019). Edge computing: A survey on the hardware requirements in the internet of things world. *Future Internet*, *11*(4), 100.

Cardim Ferreira Lima, M., Damascena de Almeida Leandro, M. E., Valero, C., Pereira Coronel, L. C., & Gonçalves Bazzo, C. O. (2020). Automatic detection and monitoring of insect pests—A review. *Agriculture*, *10*(5), 161.

Casalbore, M., Spadaccino, P., Pisani, P., Cuomo, F., & Panella, M. (2026). From Telemetry to Network Diagnostics: Supporting Operators in Massive IoT over LoRaWAN. *Internet of Things*, 101979.

Cerulli, G., & Reale, E. (2026). Conclusions: Contributing to Global and National Trends in Open Science: Challenges and Future Directions. *Research Infrastructures for Open Science in Social Science Research*, 377.

Chang, Y., Xu, H., Zhou, J., Ma, Z., Nie, P., & Aziz, Z. (2025). Design and analysis of terrestrial laser scanner based on a 3-SPR parallel mechanism for improved anti-occlusion scanning. *Measurement Science and Technology*, *36*(1), 016032.

Christakakis, P., Papadopoulou, G., Mikos, G., Kalogiannidis, N., Ioannidis, D., Tzovaras, D., & Pechlivani, E. M. (2024). Smartphone-based citizen science tool for plant disease and insect pest detection using artificial intelligence. *Technologies*, *12*(7), 101.

IEEE Institute for Excellence in Education & Research

Chuine, I., Cour, P., & Rousseau, D. D. (1999). Selecting models to predict the timing of flowering of temperate trees: implications for tree phenology modelling. *Plant, Cell & Environment*, *22*(1), 1-13.

Ding, M., Xiao, B., Codella, N., Luo, P., Wang, J., & Yuan, L. (2022, October). Davit: Dual attention vision transformers. In *European conference on computer vision* (pp. 74-92). Cham: Springer Nature Switzerland.

Dong, Y., Liu, L., Zhai, X., & Li, W. (2025). Artificial intelligence in agricultural pest and disease management: Current applications and future prospects. *Advances in Resources Research*, *5*(2), 971-986.

Dongare, A., Hesling, C., Bhatia, K., Balanuta, A., Pereira, R. L., Iannucci, B., & Rowe, A. (2017, March). OpenChirp: A low-power wide-area networking architecture. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (pp. 569-574). IEEE.

Fooladgar, F., & Kasaei, S. (2020). A survey on indoor RGB-D semantic segmentation: from hand-crafted features to deep convolutional neural networks. *Multimedia Tools and Applications*, *79*(7), 4499-4524.

Hall, T., & Rasheed, K. (2025). A survey of machine learning methods for time series prediction. *Applied Sciences*, *15*(11), 5957.

Hatamizadeh, A., Heinrich, G., Yin, H., Tao, A., Alvarez, J. M., Kautz, J., & Molchanov, P. (2024, May). Fastervit: Fast vision transformers with hierarchical attention. In *International Conference on Learning Representations* (Vol. 2024, pp. 29368-29391).

Hettiarachchi, C. N., & Meegama, R. G. N. (2026). Advanced Machine Learning Techniques for Pest Classification: A Comprehensive Review. *Computational Intelligence*, *42*(2), e70229.

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of organizational computing and electronic commerce*, *26*(1-2), 14-40.

Huang, X., Li, Z., Zhang, M., & Gao, S. (2022). Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. *Frontiers in Oncology*, *12*, 994950.

Hufkens, K., Basler, D., Milliman, T., Melaas, E. K., & Richardson, A. D. (2018). An integrated phenology modelling framework in R. *Methods in Ecology and Evolution*, *9*(5), 1276-1285.

Jin, H., Huang, S., Gao, D., Geng, Y., Chen, S., Mei, X., & Zhu, J. (2026). HDS-YOLO: A high-performance real-time framework for wind turbine blade defect detection via synergistic feature preservation. *Array*, *30*, 100923.

Joglekar, A., Gurrala, G., Kumar, P., Joseph, F. C., Kiran, T. S., Sahasranand, K. R., & Tyagi, H. (2021). Open-source heterogeneous constrained edge-computing platform for smart grid measurements. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-12.

Kapetas, D., Christakakis, P., Faliagka, S., Katsoulas, N., & Pechlivani, E. M. (2025). AI-driven insect detection, real-time monitoring, and population forecasting in greenhouses. *AgriEngineering*, *7*(2), 29.

Karaagac, A., De Poorter, E., & Hoebeke, J. (2019). In-band network telemetry in industrial wireless sensor networks. *IEEE Transactions on Network and Service Management*, *17*(1), 517-531.

Kim, J., Kim, N., Cha, J. H., Ma, J., & Lee, I. (2026). Comprehensive benchmarking of metagenomic binning tools reveals key factors for improved genome recovery. *Nature Communications*, *17*(1), 3467.

Kong, X., Chen, Z., Liu, W., Ning, K., Zhang, L., Muhammad Marier, S., ... & Xia, F. (2025). Deep learning for time series forecasting: a survey. *International Journal of Machine Learning and Cybernetics*, *16*(7), 5079-5112.

Li, C., & Wang, M. (2024). Pest and disease management in agricultural production with artificial intelligence: Innovative applications and development trends. *Advances in Resources Research*, *4*(3), 381-401.

Lichtenthaler, U. (2011). Open innovation: Past research, current debates, and future directions. *Academy of management perspectives*, *25*(1), 75-93.

Lima, M. C. F., de Almeida Leandro, M. E. D., Valero, C., Coronel, L. C. P., & Bazzo, C. O. G. (2020). Automatic detection and monitoring of insect pests—A review. *Agriculture*, *10*(5), 161.

Lin, W., Hasenstab, K., Moura Cunha, G., & Schwartzman, A. (2020). Comparison of handcrafted features and convolutional neural networks for liver MR image adequacy assessment. *Scientific Reports*, *10*(1), 20336.

Ma, C., Li, J., Qi, X., Sun, M., & Zhou, H. (2024). Pest manager: a systematic framework for precise pest counting and identification in invisible grain pile storage environment. *arXiv preprint arXiv:2409.07483*.

Ma, Y., Xi, C., Ma, T., Sun, H., Lu, H., Xu, X., & Xu, C. (2025). I-YOLOv11n: a lightweight and efficient small target detection framework for UAV aerial images. *Sensors*, *25*(15), 4857.

Majdoubi, J., Jamel, L., Mansouri, W., Alattas, K. A., Al-Sharafi, A. M., & Alshammari, A. (2025). An AI-driven data approach to enhancing climate resilience through climate change temperature prediction and early warning systems. *Expert Systems with Applications*, 130891.

Moorthy, S., KS, S. S., Arthanari, S., Jeong, J. H., & Joo, Y. H. (2025). Hybrid multi-attention transformer for robust video object detection. *Engineering Applications of Artificial Intelligence*, *139*, 109606.

Nugroho, K. A. (2018, October). A comparison of handcrafted and deep neural network feature extraction for classifying optical coherence tomography (OCT) images. In *2018 2nd international conference on informatics and computational sciences (ICICoS)* (pp. 1-6). IEEE.

Nurfauzi, A. H., Azhar, Y., & Chandranegara, D. R. (2023). Penerapan Model EfficientNetV2-B0 pada Benchmark IP102 Dataset untuk Menyelesaikan Masalah Klasifikasi Hama Serangga. *Jurnal Repositor*, *5*(3).

Partel, V., Nunes, L., Stansly, P., & Ampatzidis, Y. (2019). Automated vision-based system for monitoring Asian citrus psyllid in orchards utilizing artificial intelligence. *Computers and Electronics in Agriculture*, *162*, 328-336.

Rani, J., & Mitra, B. (2026). Future Directions, Trends, Challenges, and Opportunities in Research. *Innovative Approaches to Social Research Challenges*, 189-212.

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., ... & McHardy, A. C. (2017). Critical assessment of metagenome interpretation—a benchmark of metagenomics software. *Nature methods*, *14*(11), 1063-1071.

Shamsipour, G., Fekri-Ershad, S., Sharifi, M., & Alaei, A. (2024). Improve the efficiency of handcrafted features in image retrieval by adding selected feature generating layers of deep convolutional neural networks. *Signal, image and video processing*, *18*(3), 2607-2620.

Tetila, E. C., Machado, B. B., Astolfi, G., de Souza Belete, N. A., Amorim, W. P., Roel, A. R., & Pistori, H. (2020). Detection and classification of soybean pests using deep learning with UAV images. *Computers and Electronics in Agriculture*, *179*, 105836.

Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: a survey. *Big data*, *9*(1), 3-21.

Wang, R., Liu, L., Xie, C., Yang, P., Li, R., & Zhou, M. (2021). Agripest: A large-scale domain-specific benchmark dataset for practical agricultural pest detection in the wild. *Sensors*, *21*(5), 1601.

Wu, X., Zhan, C., Lai, Y. K., Cheng, M. M., & Yang, J. (2019, June). Ip102: A large-scale benchmark dataset for insect pest recognition. In *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 8779-8788). IEEE.

Xie, J., Yuan, B., Guo, C., Li, H., Wang, F., Chu, P., ... & Guo, H. (2025). KL-YOLO: a lightweight adaptive global feature enhancement network for small-object detection in low-altitude remote sensing imagery. *IEEE Transactions on Instrumentation and Measurement*, *74*, 1-13.

Xue, Z., Yang, J., Chen, R., He, Q., Li, Q., & Mei, X. (2024). AR-assisted guidance for assembly and maintenance of avionics equipment. *Applied Sciences*, *14*(3), 1137.

Zhalmagambetova, U., Neftissov, A., Biloshchytskyi, A., Kazambayev, I., Shimpf, A., Kazhibekov, M., & Snopkov, D. (2025). A secure telemetry transmission architecture independent of GSM: an experimental LoRa-based system on Raspberry Pi for IIoT monitoring tasks. *Applied Sciences*, *15*(17), 9539.

Zhang, J., Chen, X., Ye, X., Yang, Y., & Ai, B. (2026). Large language model in materials science: Roles, challenges, and strategic outlook. *Advanced Intelligent Discovery*, *2*(4), e202500085.

Zhang, L., Xiong, N., Gao, W., & Wu, P. (2024). Improved detection method for micro-targets in remote sensing images. *Information*, *15*(2), 108.

Zhu, H., Zong, L., Liu, C., & Guo, J. (2026). A 3D-enhanced occlusion-aware correlation filter for assembly quality inspection in confined spaces. *Engineering Applications of Artificial Intelligence*, *165*, 113418.

IEER
Institute for Excellence in Education & Research