Agri Articles

(e-Magazine for Agricultural Articles)

Volume: 06, Issue: 04 (JULY-AUG, 2026)

Available online at http://www.agriarticles.com

Deep Learning-Based Identification and Classification of Insect Pests in Agricultural Crops

[*]Dr. C. Sankar

Department of Veterinary Parasitology, Veterinary College and Research Institute, Udumalpet-642205, Tiruppur, Tamil Nadu, India

[*]Corresponding Author's email: csankardr@gmail.com

Agricultural insect pests are responsible for substantial reductions in crop productivity and quality, while conventional scouting remains labour-intensive, subjective and difficult to scale. Deep learning has emerged as a powerful computer-vision approach for automated pest identification, classification and detection from images acquired using smartphones, digital cameras, smart traps, unmanned aerial vehicles and other sensing platforms. Convolutional neural networks can learn discriminative visual features directly from images, while object-detection architectures such as YOLO, SSD and Faster R-CNN can locate individual insects in complex scenes. More recent vision transformers and hybrid CNN–transformer models provide alternative mechanisms for learning fine-grained visual patterns and long-range relationships. Public datasets such as IP102 have helped establish benchmarks, although class imbalance, inter-species similarity, small-object detection and domain shift between curated datasets and field conditions remain major limitations. This review examines the workflow of deep learning-based insect pest recognition, including image acquisition, annotation, preprocessing, augmentation, model training, transfer learning, classification, object detection and field deployment. It also discusses mobile and edge AI, explainable artificial intelligence, integration with IoT and decision-support systems, current challenges and future research directions. Robust expert-verified datasets, field validation, lightweight models and uncertainty-aware systems will be essential for translating high laboratory performance into dependable agricultural pest-management tools.

Introduction

Insect pests are among the major biological constraints to agricultural production. Accurate identification is fundamental to integrated pest management because management decisions depend on knowing the pest species, developmental stage, abundance and distribution. Conventional identification relies heavily on field scouting, trapping and expert taxonomic knowledge. Although these approaches remain indispensable, they can be slow and difficult to implement continuously across large agricultural areas. Recent machine-vision research is shifting pest identification toward automated and potentially real-time systems that can support earlier intervention and more targeted crop protection (Islam et al., 2026). Deep learning is a branch of machine learning in which multilayer neural networks learn hierarchical representations from data. In image analysis, convolutional neural networks (CNNs) have been particularly influential because they can learn patterns related to shape, texture, colour and spatial structure without requiring manually designed features. The development of object-detection architectures has further enabled systems to locate pests within photographs containing leaves, stems, soil, traps and multiple insects. The availability of large datasets has accelerated this field. The IP102 benchmark contains more than 75,000 images across 102 insect-pest categories and includes bounding-box annotations for about 19,000 images, providing a major benchmark for classification and detection research (Wu et al., 2019). However, the dataset also illustrates the difficulty of agricultural pest recognition because insect categories can show large within-class variation, strong similarity between species and an imbalanced distribution of images. Recent reviews indicate that the field is moving from conventional CNN classifiers toward object detectors, attention mechanisms, vision transformers and hybrid architectures. Nevertheless, high accuracy on curated datasets does not necessarily guarantee reliable field performance. Differences in lighting, background, insect orientation, scale, occlusion, crop variety and camera type can produce substantial domain shifts. Therefore, agricultural deep learning should be evaluated not only by accuracy but also by robustness, generalization, computational efficiency and usefulness for real pest-management decisions (Islam et al., 2026).

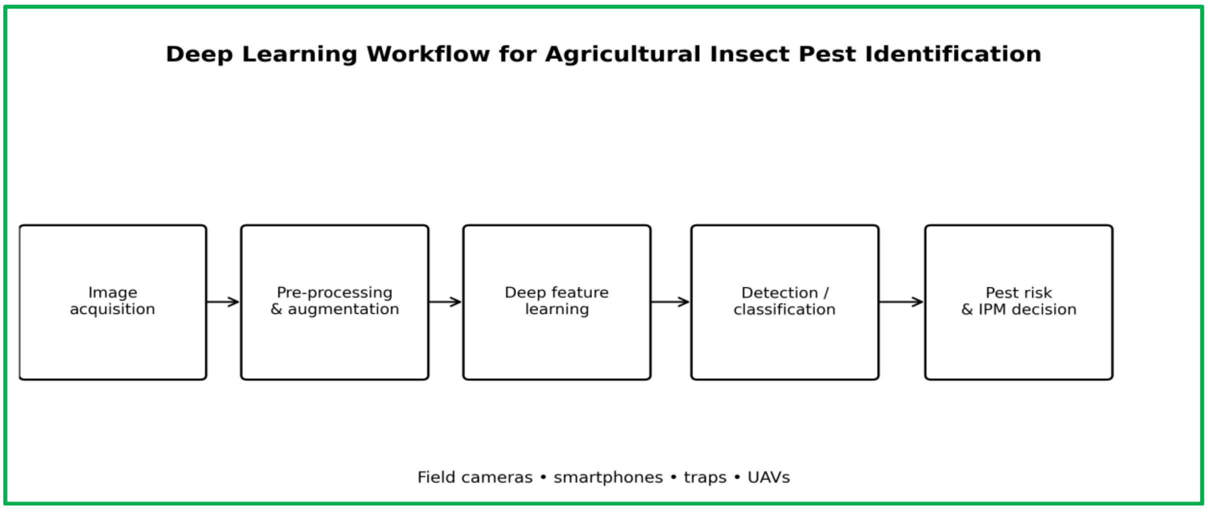

Figure. General workflow for deep learning-based insect pest identification from image acquisition to pest-management decision support.

Image Acquisition for Insect Pest Recognition

The first stage of an automated pest-recognition system is image acquisition. Images may be collected from smartphones, DSLR or digital cameras, sticky traps, pheromone traps, light traps, fixed field cameras, greenhouse cameras, UAVs and other sensing platforms. The choice of imaging platform determines the resolution, viewing angle, illumination and type of biological information available to the model. For field applications, images should represent the variability that the model will encounter after deployment. Important factors include morning and afternoon illumination, shadows, rain or dust, different crop growth stages, leaf orientation, pest density, multiple insects in one image and different distances between the camera and target. If training images are collected only under controlled laboratory conditions, the resulting model may learn background or imaging characteristics rather than biologically meaningful pest features. Trap-based imaging is especially suitable for automated monitoring because the insect can be concentrated on a defined surface. Field scouting images provide more realistic ecological conditions but introduce greater complexity. UAV imagery can provide broad spatial coverage, although small insect targets may be difficult to resolve. Combining image sources can therefore improve generalization when the training dataset is carefully designed.

Dataset Development, Annotation and Pre-processing

Deep learning models require representative and accurately labelled datasets. For classification, each image is assigned a pest category, whereas object detection requires bounding boxes around individual insects. Segmentation tasks require pixel-level masks and are useful when insect shape or area must be measured precisely. Annotation should ideally be verified by trained entomologists because incorrect species labels can propagate errors through the model. Dataset diversity is as important as dataset size. A useful dataset should include different pest life stages where possible, including eggs, larvae or nymphs, pupae and adults when these stages are visually distinguishable and relevant to management. Metadata such as crop, location, season, camera type, lighting, pest density and management conditions can improve subsequent analysis and help identify domain-shift problems. Pre-processing may include resizing, normalization, background correction and removal of unusable images. Data augmentation can introduce controlled changes in rotation, scale, brightness, contrast, cropping and noise. Augmentation can improve robustness, but excessive transformations that produce biologically unrealistic images should be avoided. Class imbalance is another major concern; rare pest classes can be overshadowed by abundant categories during model training. Balanced sampling, class-weighted losses, focal loss and targeted data collection can help address this problem.

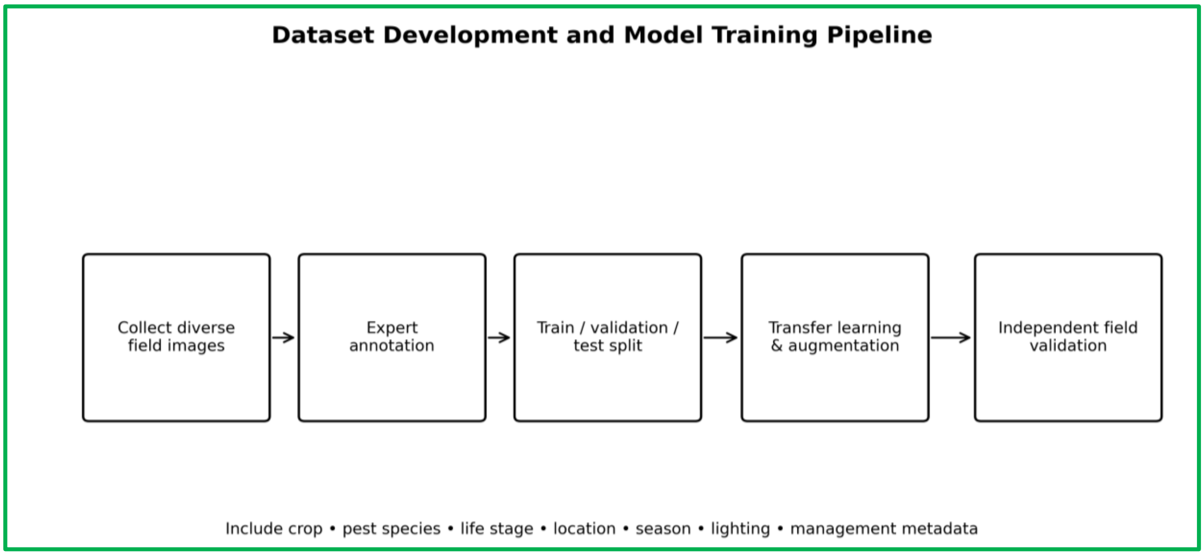

Figure. Recommended dataset and model-development pipeline emphasizing expert annotation, representative sampling and independent field validation.

Convolutional Neural Networks for Pest Classification

CNN-based classification models learn increasingly abstract representations through successive convolutional and pooling operations. Earlier layers generally learn low-level patterns such as edges and textures, while deeper layers learn combinations of features that can distinguish pest categories. Architectures such as ResNet, EfficientNet and MobileNet have been widely adapted to image-classification tasks and can be fine-tuned for agricultural datasets. Transfer learning is particularly valuable when agricultural datasets are smaller than general computer-vision datasets. A model pretrained on a large image corpus can be adapted to insect categories by replacing or fine-tuning the final classification layers. Fine-tuning can reduce training requirements and improve performance when the target images contain sufficient visual similarity to the pretraining domain. CNN classifiers are most appropriate when the image contains a single insect or when the pest has already been cropped from the background. Their limitation is that they do not inherently provide the location of multiple insects in a complex field image. This distinction makes it important to select the architecture according to the practical question: species classification, individual detection, counting or segmentation.

Object Detection Using YOLO, Faster R-CNN and SSD

Object detection combines localization and classification. Instead of assigning one label to an entire image, a detector predicts where insects occur and assigns a class to each detected object. YOLO-family detectors are widely used because they are designed for fast inference and can be adapted for real-time applications. Faster R-CNN provides a two-stage detection framework and can offer strong detection performance, while SSD provides another single-stage alternative suitable for relatively efficient deployment. Agricultural insect detection is challenging because insects can occupy only a small fraction of an image and may overlap with leaves, stems or other insects. Multi-scale feature extraction and feature-pyramid approaches can improve detection of small objects. Attention mechanisms can help a model focus on informative regions, while image tiling or higher-resolution inputs can increase the effective size of small pests. Detection performance should be evaluated using precision, recall, F1-score and mean average precision rather than accuracy alone. A model with high image-level accuracy may still miss many individual insects if they are small or partially occluded. For pest management, false negatives can be especially important because failure to detect a developing infestation may delay intervention.

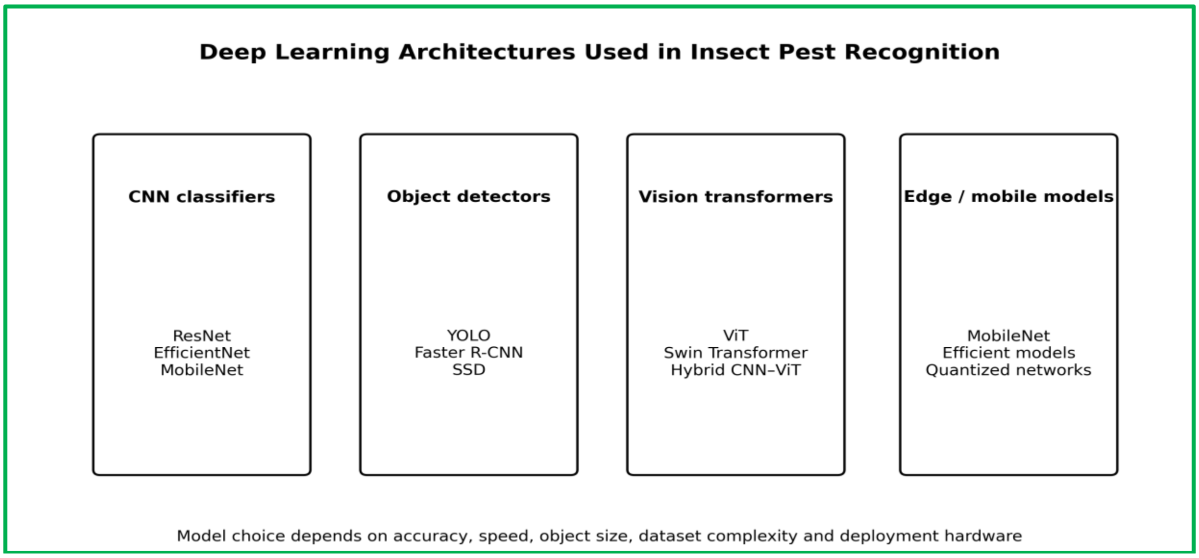

Figure. Major deep learning architecture families used for agricultural insect pest classification and detection

Vision Transformers and Hybrid Deep Learning Models

Vision transformers (ViTs) apply transformer-based attention mechanisms to image patches and can model relationships between distant regions of an image. Their growing use in agricultural image analysis reflects interest in learning broader contextual information beyond local convolutional features. Hybrid CNN–transformer systems combine convolutional feature extraction with attention-based representation learning and may be useful for fine-grained discrimination among visually similar insect species. A 2025 review of crop-pest classification reported a transition from conventional CNN-based studies toward hybrid and transformer-based models, while identifying imbalanced data, small-object detection, limited generalization and edge-device constraints as persistent challenges. These findings suggest that architectural innovation alone is unlikely to solve the agricultural pest-recognition problem; data quality and field validation remain equally important.

Fine-Grained Classification of Similar Pest Species

One of the most difficult problems in insect recognition is distinguishing closely related species with similar morphology. Two species may share colour, body shape and wing patterns while differing in small anatomical features. This is a fine-grained visual recognition problem in which the model must learn subtle differences rather than broad categories. Attention mechanisms, high-resolution feature extraction, multi-scale learning and hierarchical classification can improve fine-grained recognition. A hierarchical system may first identify the crop or pest family and then classify the species. Expert-guided feature annotation can also help researchers determine whether the model is focusing on biologically meaningful structures. Life-stage recognition is another important research gap. Many current datasets and models focus primarily on adult insects, although immature stages may cause substantial crop damage and are often the most appropriate targets for early intervention. Expanding datasets to include larvae, nymphs and eggs where feasible could improve the practical value of automated pest diagnosis (Islam et al., 2026).

Transfer Learning, Data Augmentation and Model Optimization

Transfer learning reduces the amount of labelled agricultural data required to train a model from the beginning. A pretrained network can be fine-tuned on insect-pest images, with deeper layers optionally frozen during initial training. This approach is useful when researchers have limited images for individual pest species. Data augmentation can simulate some of the variability encountered in the field. Rotation, flipping, cropping, scaling, brightness changes and controlled blur can increase training diversity. However, augmentation should preserve the biological identity of the insect. Researchers should also avoid leakage between training and testing sets; images of the same insect or nearly identical frames should not be distributed across both sets because this can produce overly optimistic performance. Model compression is important for field deployment. Quantization, pruning and knowledge distillation can reduce memory and computational requirements. Lightweight networks such as MobileNet-family architectures may be appropriate for smartphones, embedded cameras and edge devices where cloud connectivity is unreliable.

Performance Evaluation and Validation

A reliable pest-recognition study should report multiple complementary metrics. Accuracy is useful for balanced classification datasets, but precision, recall and F1-score are often more informative when classes are imbalanced. For object detection, mean average precision at defined intersection-over-union thresholds is commonly reported. Confusion matrices can identify pest pairs that are frequently misclassified. Validation should be conducted at multiple levels. A random image split can estimate performance under similar conditions, but an independent field test is more informative for deployment. Ideally, the test set should include locations, seasons, crop stages, camera conditions and pest densities that were not represented in training. Such external validation helps determine whether the model has learned general pest characteristics rather than dataset-specific backgrounds.

Explainable Artificial Intelligence for Pest Identification

Explainability is increasingly important because agricultural users need confidence that model predictions are based on insect characteristics rather than irrelevant background patterns. Methods such as class-activation mapping and related visualization approaches can indicate which image regions influenced a prediction. When a model consistently focuses on the insect body, wings, head or other relevant structures, the explanation can support biological interpretation. Explainability can also reveal failure modes. If a model focuses on leaf colour, trap background or image borders, its apparent accuracy may not transfer to a different field. Combining model explanations with entomological expertise can therefore improve dataset design, identify biases and increase confidence before deployment.

Mobile, Edge and IoT-Based Pest Recognition

Smartphones provide an accessible platform for image-based pest identification. A farmer or extension worker can capture an image and submit it to a mobile model or cloud service. Edge AI offers an alternative in which inference is performed locally on the device, reducing dependence on continuous internet connectivity and improving response time. IoT-enabled traps can combine automated imaging with environmental sensors. A smart trap can record pest images at regular intervals, classify insects and transmit counts to a dashboard. When integrated with temperature, humidity and crop-growth information, these systems can provide a more complete picture of pest risk. The 2026 machine-vision review highlights the potential of combining machine vision with IoT for real-time and proactive pest management (Islam et al., 2026).

Integration with Integrated Pest Management

Deep learning should be considered a decision-support technology rather than a replacement for entomological diagnosis. Automated identification can support scouting by rapidly screening large numbers of images, but management decisions should also consider pest abundance, crop stage, natural enemies, economic thresholds and local recommendations.

A practical system can follow a sequence of detection, counting, risk assessment and intervention. For example, automated trap images may indicate increasing pest abundance; the system can then combine the count with weather and crop-stage information to generate an alert. An extension worker can verify the diagnosis before recommending a control measure. This workflow reduces the risk of treating an image-classification result as an automatic pesticide prescription.

Challenges and Limitations

The main limitation is the gap between curated datasets and real agricultural environments. Insects can be partially hidden, very small, motion-blurred or visually similar to plant structures. Lighting changes, rain, dust, shadows and background clutter can reduce model performance. These conditions are substantially different from laboratory photographs. Dataset imbalance is another persistent problem. Some pest species are photographed frequently while economically important but less common species may have few examples. Long-tailed distributions can cause a model to optimize for common classes while performing poorly on rare classes. IP102 was specifically developed to expose these challenges, containing a natural long-tailed distribution and substantial variation within and between categories (Wu et al., 2019). Generalization across geographical regions is also difficult. A model trained on one country or crop system may encounter different pest morphologies, varieties, backgrounds and management conditions elsewhere. Domain adaptation, diverse training data and local fine-tuning can help, but independent validation remains essential. Computational cost, energy consumption and connectivity can restrict deployment. Large transformer or detection models may require hardware that is unavailable to smallholder farmers. Lightweight edge models, model compression and offline inference can improve accessibility. Data privacy, ownership and governance should also be considered when images and farm information are uploaded to cloud platforms.

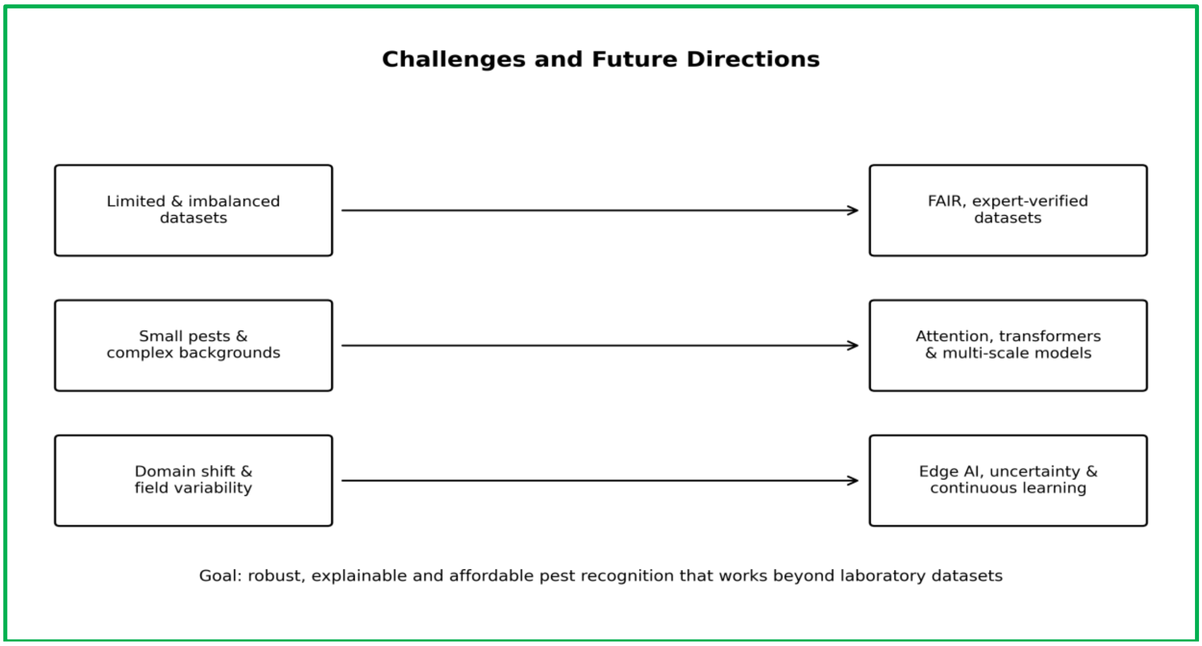

Figure 4. Major limitations of current deep learning pest-recognition systems and corresponding research priorities.

Future Perspectives

Future research should move toward biologically informed and field-validated deep learning. Datasets should be developed with entomologists and should include verified species labels, multiple life stages, diverse environments and standardized metadata. FAIR data practices and reproducible benchmarking can improve comparison among models and reduce duplication of effort.

Multimodal systems represent another promising direction. Combining RGB images with hyperspectral information, thermal data, environmental sensors or trap counts may improve detection under conditions where visible appearance alone is ambiguous. However, additional sensors increase cost and complexity, so research should evaluate whether the added information produces meaningful gains under realistic farm conditions. Continual learning and uncertainty estimation may improve safety and reliability. Instead of forcing every image into a known pest category, future systems should be able to flag low-confidence or out-of-distribution images for human verification. This human-in-the-loop approach is particularly important for rare pests, new invasive species and unfamiliar field conditions. The integration of deep learning with climate-based pest forecasting, remote sensing and precision agriculture could create comprehensive pest early-warning systems. Automated detection can provide the biological observation, while weather and crop models provide the environmental context. Together, these tools could support more timely, targeted and sustainable pest management.

Conclusion

Deep learning is rapidly transforming the identification and classification of agricultural insect pests by enabling automated analysis of images from field scouting, traps, smartphones and remote-sensing platforms. CNNs, object detectors, attention mechanisms, vision transformers and hybrid models have expanded the range of possible applications from simple species classification to real-time detection and counting. Public datasets such as IP102 have provided valuable benchmarks and accelerated methodological development. However, practical agricultural deployment remains more difficult than achieving high accuracy on curated datasets. Small insects, complex backgrounds, class imbalance, visually similar species, life-stage variation, geographic domain shifts and limited computational resources continue to constrain performance. The next generation of systems should therefore emphasize representative expert-verified datasets, independent field validation, lightweight models, explainability and uncertainty-aware predictions. When integrated with IoT monitoring, weather information, crop phenology and IPM decision rules, deep learning can become an important component of precision pest management. Continued collaboration among entomologists, computer scientists, agronomists, extension personnel and farmers will be necessary to ensure that technological advances produce reliable, affordable and environmentally responsible benefits in agricultural production.

References

1. Wu, X., Zhan, C., Lai, Y.-K., Cheng, M.-M. and Yang, J. (2019). IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8787–8796. https://doi.org/10.1109/CVPR.2019.00899
2. Islam, M.M., Akter, M., Uddin, M.S., Rahman, M.A. and Habib, M.T. (2026). Machine Vision–Based Insect Recognition in Agriculture: A Comprehensive Review. Engineering Reports, 8(4): e70729. https://doi.org/10.1002/eng2.70729
3. Ejaz, M.H., Bilal, M. and Habib, U. (2025). Crop Pest Classification Using Deep Learning Techniques: A Review. arXiv preprint arXiv:2507.01494.
4. Ung, H.T., Ung, H.Q. and Nguyen, B.T. (2021). An Efficient Insect Pest Classification Using Multiple Convolutional Neural Network Based Models. arXiv preprint arXiv:2107.12189.
5. Rafi, M.H., Mahjabin, M.R. and Rahman, M.S. (2023). An Efficient Deep Learning-based Approach for Recognizing Agricultural Pests in the Wild. arXiv preprint arXiv:2310.16991.
6. Banerjee, S., Mallick, T., Chakroborty, A., Saha, H.N. and Takur, N.T. (2025). Evaluation of State-of-the-Art Deep Learning Techniques for Plant Disease and Pest Detection. arXiv preprint arXiv:2508.08317.