

Artificial Intelligence in Entomology: Smart Detection, Identification and Management of Insect Pests

Dr. Indu

Assistant Professor, Department of Zoology, Government Degree College Nanauta, Saharanpur, Uttar Pradesh, INDIA

Corresponding Author: dr.indusingh37@gmail.com



www.jrasb.com || Vol. 5 No. 4 (2026): August Issue

Received: 10-08-2026

Revised: 24-08-2026

Accepted: 30-08-2026

ABSTRACT

Artificial intelligence can expand the frequency and scale of insect surveillance, but its agricultural value depends on the decisions supported by its predictions. This review examines image recognition, automated traps, optical sensing, uncertainty estimation and their integration into pest management. Published studies demonstrate progress in identifying insects under controlled and field conditions, including recognition across thousands of species and evaluation on independent greenhouse observations. However, image classification accuracy, trap counts and crop protection outcomes represent different quantities. Dataset imbalance, unfamiliar species, repeated observations and changing environmental conditions complicate their interpretation. A critical synthesis identifies three requirements for dependable deployment: taxonomically defensible reference data, validation across sites and seasons, and explicit links between observations and management thresholds. Six original figures distinguish reported measurements from conceptual and mathematical illustrations. Developments available through 2026 support systems that combine automated surveillance with expert verification and ecological knowledge. Evidence of sustained economic and environmental benefit remains more limited than evidence of recognition performance.

Keywords- Artificial Intelligence; Entomology; Insect Pest Detection; Pest Identification; Machine Learning; Deep Learning; Computer Vision; Automated Insect Traps; Optical Sensing; Precision Agriculture; Smart Pest Management; Image Recognition; Automated Surveillance; Uncertainty Estimation; Integrated Pest Management (IPM); Agricultural Monitoring; Insect Classification; Pest Prediction; Crop Protection; Digital Entomology.

I. INTRODUCTION

Insect pests reduce crop yield and quality through feeding and disease transmission. Their effects occur within a broader crop-health problem: frequently cited global loss estimates combine insects with pathogens and other damaging organisms and therefore must not be interpreted as insect-only losses [1,2]. Climate-dependent changes in insect metabolism, development and distribution further complicate surveillance. Mechanistic projections indicate that warming can increase insect-related losses in major grain crops, although the magnitude varies geographically and among crops [3].

Integrated pest management (IPM) provides the appropriate framework for evaluating artificial

intelligence (AI). Its purpose is to prevent economically important damage while limiting ecological disruption, rather than to eliminate every insect encountered. Economic thresholds, representative sampling and conservation of natural enemies remain essential [4, 5]. Automated recognition contributes information to this framework; it does not supply the complete management decision.

Computer vision can shorten the interval between observation and identification, while connected sensors can generate repeated records over periods that are impractical for manual inspection [6, 7]. Nevertheless, automation also multiplies errors when unreliable predictions are accepted without scrutiny. This review therefore asks how AI can produce dependable pest information, what current evidence actually establishes,

and which additional tests are necessary before detection becomes demonstrably useful management.

Four endpoints organize the assessment: correct taxonomic identity, accurate population information, appropriate management decisions and beneficial field outcomes. Evidence at one endpoint supports, but does not prove, the next. This distinction is especially important when a study reports image-level performance while its practical motivation concerns reduced crop damage. The review treats that transition as an empirical question requiring explicit validation.

II. REVIEW APPROACH AND EVIDENCE BASE

2.1 Scope and selection

This is a critical narrative review, not a systematic review or meta-analysis. Targeted web searches of scholarly records and publisher pages were completed on 11 September 2026, using combinations of insect pest, artificial intelligence, identification, smart

trap, counting, uncertainty and integrated pest management. Reference tracing supplemented these searches. Priority was given to peer-reviewed studies, original benchmark descriptions and foundational methodological papers; one explicitly labelled multisensor preprint supplies emerging technical context. English-language sources were selected for relevance to detection, identification, monitoring or decisions. No original field experiment was conducted, and no exhaustive screening count or pooled effect size is claimed.

2.2 What the datasets represent

IP102 introduced more than 75,000 images across 102 pest categories, together with approximately 19,000 bounding-box-annotated images. Its long-tailed distribution makes rare-category learning a substantive challenge [8]. Pest24 contains 25,378 annotated images across 24 pest categories and emphasizes small, densely distributed objects obtained with an automated trapping and imaging device [9]. Figure 1 compares dataset scale without implying equivalent sampling designs or interchangeable performance tests.

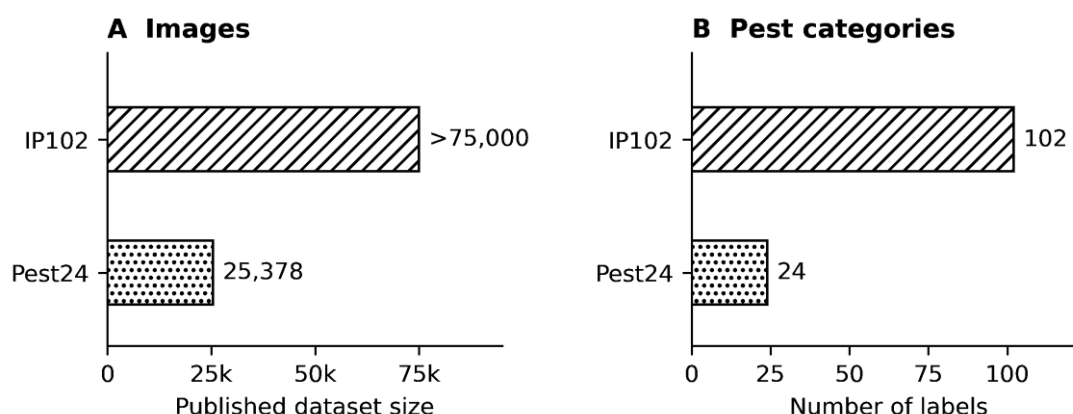


Figure 1. Benchmark scale and task coverage. Original visualization of published dataset descriptions [8, 9]. IP102 is shown as more than 75,000 images, its reported rounded lower bound; Pest24 contains 25,378. Categories are benchmark labels, not a measure of geographic coverage. IP102 combines classification data with a smaller detection subset, whereas Pest24 emphasizes multi-object detection. These dimensions do not measure model quality.

A useful dataset must represent life stage, sex, pose, damage, illumination, crop background and image quality. Larvae and adults may differ more than adults of separate species. Conversely, cryptic taxa may not possess visible diagnostic characters in an ordinary photograph. Expert labels should record the finest defensible taxonomic rank and retain links to vouchers or other confirmation where available. A model cannot recover characters absent from the image.

Training examples should also include predators, parasitoids, pollinators and non-insect debris. Otherwise, a classifier trained only on pests may assign a pest identity to any unfamiliar object. Dataset documentation should specify provenance, annotation procedures, intended uses, licensing and known

omissions [10]. Images derived from the same specimen, observation or trap sequence must remain together during data partitioning to prevent misleadingly favourable evaluation [11].

Class imbalance has both statistical and biological consequences. Common insects can dominate an aggregate score, while uncommon invasive pests receive little training coverage. Class weighting and balanced sampling can improve attention to rare labels, but neither creates genuinely new biological variation. Additional images should therefore be selected to fill documented gaps rather than merely increase the total. Synthetic images require independent validation because plausible appearance does not establish diagnostic fidelity.

III. ALGORITHMS FOR IDENTIFICATION AND DETECTION

3.1 Matching the model to the biological question

Classification assigns a label to an image or crop; detection localizes individual insects; segmentation delineates their outlines; tracking links observations through time; and density estimation approximates counts when individuals overlap. These tasks require different annotations and error measures. Figure 2 distinguishes

their outputs. A correct image label does not establish that all insects have been counted, while an accurate total does not guarantee correct species composition.

Convolutional neural networks learn visual features from examples. Residual connections facilitate the training of deep networks, and transfer learning adapts pretrained representations to smaller entomological datasets [12, 13]. Cotton-pest research illustrates this approach with field images, but its reported performance applies to the studied categories and evaluation design [14]. Such studies support feasibility without establishing universal identification reliability.

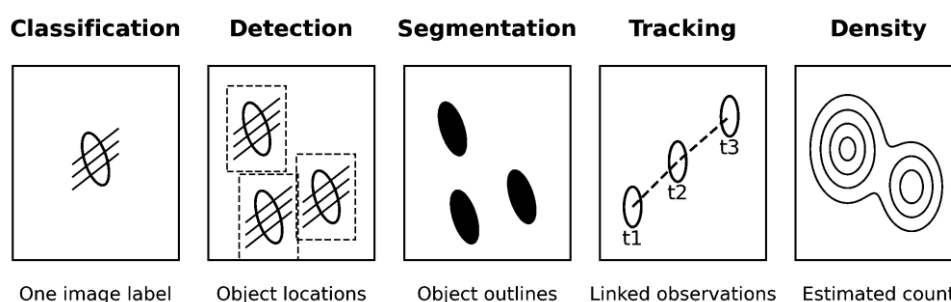


Figure 2. Five distinct outputs of insect image analysis. Original schematic showing classification, detection, segmentation, tracking and density estimation. Marks represent abstract insects, not measured specimens. The distinction is synthesized from recognition, detection, counting and tracking studies [8, 15, 16, 17]. Each output answers a different biological question and requires its own validation.

3.2 Small objects and efficient computation

Faster R-CNN uses region proposals before classification, whereas YOLO predicts object locations and classes within a unified detector [18, 19]. Focal loss reduces the training dominance of easy examples, which is relevant to backgrounds containing few insects [20]. However, architectural refinements cannot compensate for an insect occupying too few pixels to reveal diagnostic traits. Higher-resolution imaging, carefully validated tiling and suitable camera geometry can matter as much as model depth.

Tomato research demonstrated joint recognition of visible diseases and pests, illustrating the possibility of combining crop and organism observations [21]. Symptom recognition nevertheless requires caution: similar discoloration or damage may have different biological causes. For crowded aphid populations, Gao and colleagues combined object detection with density estimation, addressing a situation in which individual localization becomes increasingly difficult [16]. Density estimates should preserve information about counting uncertainty and cannot automatically resolve mixed species.

Lightweight networks such as MobileNetV2 make local computation more feasible, while edge-based pest-detection studies evaluate deployment on constrained hardware [22, 23]. Vision transformers provide another representation strategy, but their suitability depends on data volume, image resolution and computational budget [24]. The appropriate comparison

is between complete systems under the same operating conditions, including missed detections, latency and energy demand.

For practical model selection, a compact benchmark should retain an expert-inspected test set and compare candidate methods under identical preprocessing. Improvements should be examined by insect size and category, not only as a single average. A model that raises overall performance while missing early instars may be less suitable for prevention. The test set should remain untouched during parameter selection to preserve its evidential value.

IV. SMART MONITORING ACROSS AGRICULTURAL ENVIRONMENTS

4.1 Connected cameras and traps

Automated moth detection showed how images collected inside field traps could support repeated identification and counting [15]. Greenhouse sensor networks subsequently combined environmental measurements with insect imagery [25]. Such systems can reduce reporting delays, but hardware maintenance remains unavoidable. Dirty lenses, saturated adhesive surfaces, condensation and interruptions in power or communication can create changes in apparent abundance unrelated to pest biology.

Rustia and colleagues evaluated a cascaded greenhouse system that detected candidate objects, rejected non-insect material and classified insects into four groups. Mean F1 scores were 0.92 on a separate six-month image dataset and 0.90 on data from a different greenhouse; corresponding mean counting accuracies

were 0.91 and 0.90 [26]. Figure 3 preserves the authors' metric labels. These results are informative because they extend evaluation beyond a single image split, although they do not establish crop-loss reduction or performance on every greenhouse pest.

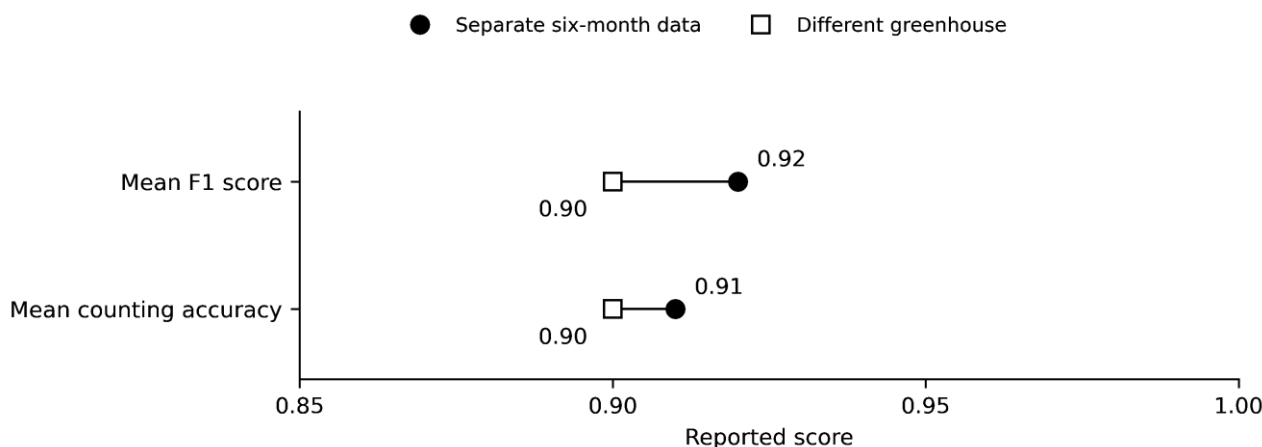


Figure 3. Reported performance under two greenhouse evaluation conditions [26]. Original paired-dot plot of mean F1 score and mean counting accuracy: separate six-month data, 0.92 and 0.91; different-greenhouse data, 0.90 and 0.90. The horizontal axis is restricted to 0.85–1.00 for readability. Connecting lines aid comparison; they do not indicate a longitudinal trajectory. No confidence intervals are inferred from these summary values.

4.2 Observation bias and complementary sensors

Trap catches measure attraction and encounter as well as abundance. Lure specificity, flight activity, temperature and trap position influence which insects are observed. Changes in catches cannot therefore be converted directly into insects per plant without local calibration. Sampling theory remains relevant even when identification is automated [27]. Tracking can reduce repeated counting of insects that remain visible, as demonstrated in an automated light-trap system for moths [17].

Optical sensors provide complementary information. Rydhmer and colleagues measured near-infrared signatures of flying insects and reported a Spearman correlation of 0.61 between daily sensor observations and conventional trap counts during a four-week comparison [28]. The correlation supports related activity patterns, not interchangeability of instruments or species-level identification. An emerging multisensor approach combines imagery, wingbeat information and environmental measurements, although its initial small, imbalanced dataset limits generalization [29].

Monitoring can also support conservation. The FAIR-Device illustrates nonlethal, semi-automatic field imaging, while BIODISCOVER links controlled specimen imaging with identification and biomass estimation [30, 31]. Their applications differ from commercial pest scouting, but they show why a monitoring system should retain useful information about beneficial and non-target organisms. Broader sensing

developments offer complementary coverage rather than a single universal instrument [32].

V. RELIABILITY BEYOND BENCHMARK ACCURACY

5.1 Independent evaluation and meaningful metrics

Field evaluation should withhold entire sites, seasons or devices, as appropriate to the intended deployment. Random image splits can place near-duplicates or closely related observations on both sides of the evaluation, exaggerating generalization [11]. Broader machine-learning benchmarks demonstrate that distribution shifts can materially degrade performance [33]. Entomological assessment must therefore report which changes were tested, rather than describe a model simply as robust.

Metrics should follow the biological task. Identification requires per-class precision and recall, macro-averaged performance and confusion between economically consequential taxa. Detection requires average precision with a stated intersection-over-union criterion. Counting requires absolute error, systematic bias and performance across density ranges. Neither accuracy nor mean average precision should be substituted for the other. Precision–recall analysis is especially informative where positive observations are uncommon [34].

5.2 Rare pests and uncertain predictions

The prevalence of a target affects the usefulness of an alarm. For prevalence p , sensitivity Se and

specificity Sp , positive predictive value is $Se \times p / [Se \times p + (1 - Sp) \times (1 - p)]$. In an illustrative population with 1% prevalence, 95% sensitivity and 95% specificity, only about 16.1% of positive predictions would be correct.

Increasing specificity to 99% raises this value to approximately 49.0%. These are mathematical scenarios, not measurements from a pest trial (Figure 4).

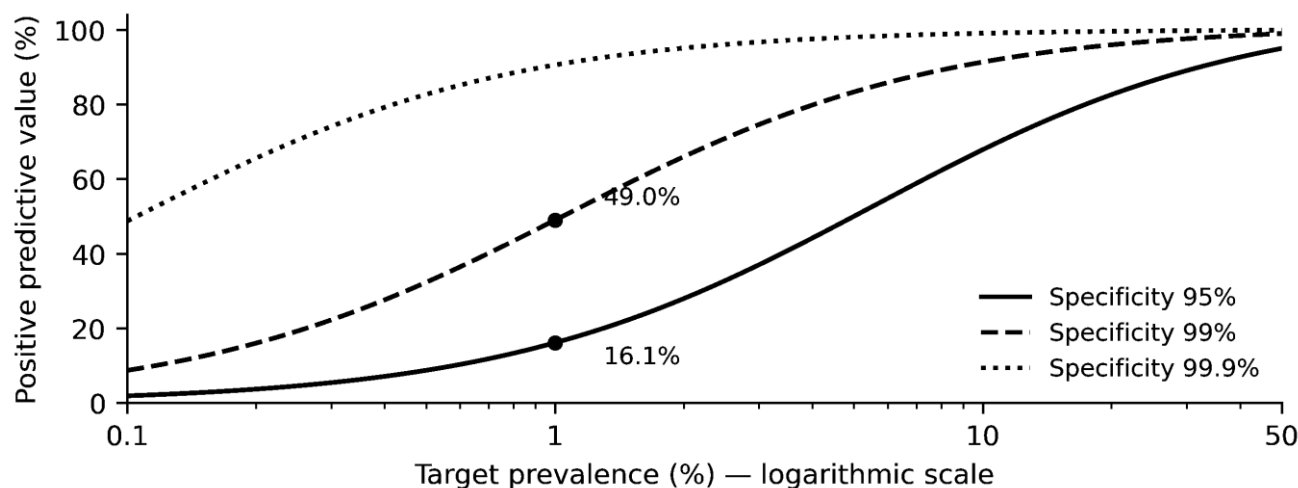


Figure 4. Why rare-target surveillance requires exceptional specificity. Original Bayesian calculation of positive predictive value across target prevalence, with sensitivity fixed at 95% and specificity set to 95%, 99% or 99.9%. Assumptions are hypothetical, rates are held constant, and the logarithmic prevalence axis spans 0.1–50%. The two labelled points correspond to the 1% examples in the text. This figure does not report an observed detector's performance.

A confidence score is not automatically a calibrated probability. Calibration should be assessed on independent data and revisited when operating conditions change [35]. Out-of-distribution methods offer a means to reject unfamiliar inputs, as examined specifically for insect classification [36]. Hierarchical prediction can retain a defensible family or genus assignment when a species-level answer is unsupported. Badirli and colleagues investigated unknown-taxon identification, while BugNet extended hierarchical insect-monitoring workflows in 2026 [37, 38].

Operationally, rejection transfers work to an expert. Evaluation must report both the proportion of observations accepted automatically and error among accepted predictions; high accuracy obtained by rejecting most records may be costly to maintain. Uncertainty intervals for counts and validation metrics should respect the sampling unit, for example through site- or trap-level resampling. This proposed reporting practice protects against artificial precision from treating correlated images as independent evidence.

Errors also have unequal consequences. Confusing a harmless species with a regulated invader can trigger expensive follow-up, while overlooking the invader can delay containment. Evaluation should consequently specify which confusions matter most and measure their frequency. Appropriate operating

thresholds depend on the surveillance objective and review capacity; no single confidence threshold is defensible for all insect-management tasks.

6 From surveillance to integrated pest management

6.1 Linking observations to an intervention threshold

Pest presence alone does not justify intervention. Decisions depend on species identity, damaging life stage, population trajectory, crop susceptibility, expected injury and the contribution of natural enemies. The classical economic injury level can be expressed as $EIL = C / (V \times I \times D \times K)$, where C denotes management cost, V crop value, I injury per insect, D damage per unit injury and K the proportion of injury prevented [39]. Consistent units and locally estimated parameters are required. The economic threshold initiates action early enough to prevent the injury level from being reached.

AI may estimate components of this decision, but detector confidence must not be confused with outbreak probability. A forecast requires time-indexed observations, weather, crop stage and a clearly defined prediction horizon. Phenological models can provide biological structure, while statistical learning can capture additional associations. Whether such a hybrid improves decisions must be tested against simpler baselines, including historical timing, recent counts and established forecasting rules.

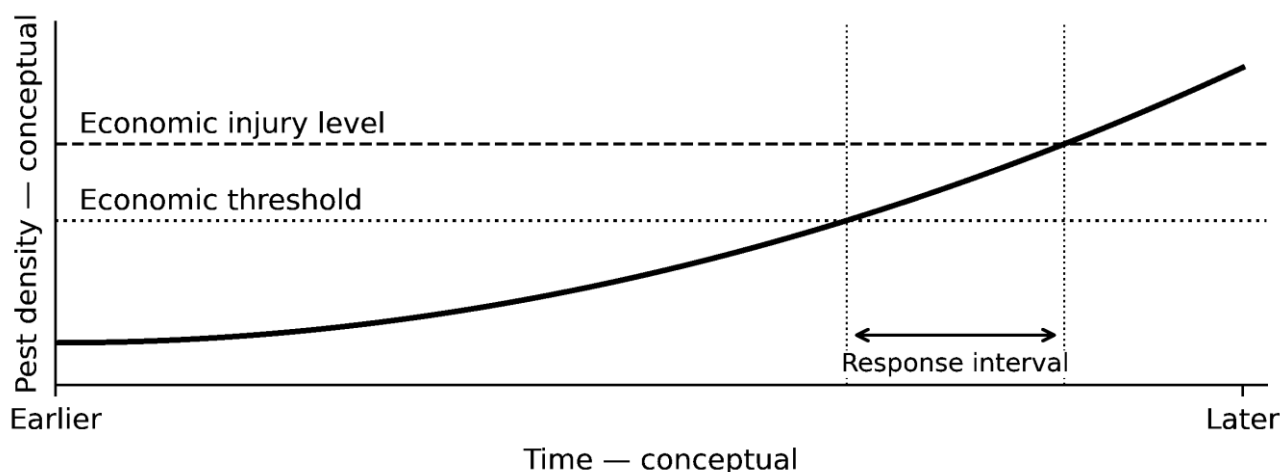


Figure 5. Conceptual relationship between monitoring and management timing. Original schematic of a rising pest population, an economic threshold and a higher economic injury level. The interval illustrates why reporting and response delays matter. Both axes are qualitative; the curve, threshold positions and timing are not field observations or recommendations for any crop. The economic interpretation follows established IPM theory [4, 27, 39].

6.2 Preserving ecological function

A useful recommendation system should consider prevention, host-plant resistance, cultural measures and biological control before selecting an intervention compatible with the local IPM programme [5]. This requires recognizing beneficial organisms and anticipating non-target consequences. A global synthesis links biodiversity to crop-production services, supporting the inclusion of natural enemies and pollinators in agricultural monitoring [40]. Identifying every visible insect as a threat would undermine this purpose.

Spatial information can direct additional scouting towards suspected hotspots and help compare management zones. Yet an apparent hotspot may reflect trap placement or sensor failure. Confirmation should precede consequential action. Post-intervention observations should examine injury, pest recovery, natural-enemy activity and yield, rather than count fewer detections as sufficient evidence of success.

Economic evaluation must include equipment, connectivity, maintenance, annotation and expert review. The appropriate outcome is incremental benefit relative to existing practice, after all relevant costs. Environmental evaluation should consider treatment frequency alongside hazard, non-target effects and resistance implications.

Recent discussion of AI-supported pest monitoring emphasizes its potential contribution to resilient agriculture [41]; the stronger claim that a particular system delivers that benefit requires comparative field evidence.

A prospective trial could compare conventional scouting with AI-assisted IPM across independently replicated sites and seasons, documenting recommendations as well as actions actually taken. This is a proposed evaluation design, not a result of the present review. Its purpose is to reveal whether faster information changes management sufficiently to justify the system.

7 Implementation and accountable operation

7.1 A complete operational system

Deployment joins sensing, quality control, identification, ecological interpretation and follow-up. Figure 6 presents an operational design in which uncertain observations return to specialists and verified corrections improve subsequent models. Each record should retain time, location, device status, taxonomic confidence and model version. Otherwise, apparent changes in abundance may be caused by a software update rather than an ecological event.

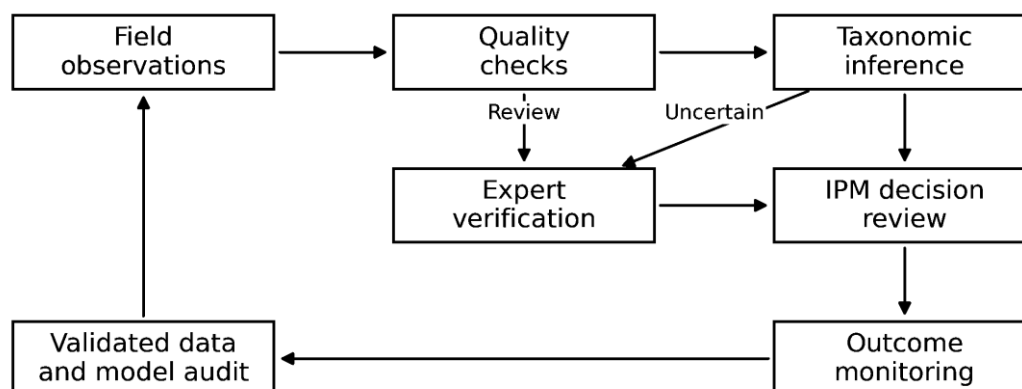


Figure 6. Proposed architecture for accountable AI-assisted pest management. Original workflow linking field observations, image quality checks, taxonomic inference and management review. A separate expert pathway handles ambiguous records, while outcome monitoring feeds validated information back into the system. The diagram is a design synthesis, not a tested product. It draws on connected sensing, uncertainty and documentation research [25, 36, 42].

Local processing can maintain basic detection where connectivity is intermittent, with later synchronization preserving the audit trail. However, device-specific validation is necessary: compression, quantization and reduced resolution may particularly impair tiny-insect recognition [23]. Reported inference speed should identify the hardware and include acquisition and transmission delays where they affect management. A rapid neural-network calculation cannot compensate for an unavailable camera.

7.2 Explanations and human oversight

Grad-CAM highlights image regions contributing to a neural-network prediction, while SHAP attributes model output to input features [43, 44]. These tools may reveal dependence on trap edges, labels or background vegetation. Their visual plausibility does not demonstrate biological causation. Explanations should therefore be paired with controlled background changes, independent observations and taxonomic inspection. A convincing image overlay is not a substitute for an error analysis.

Model cards can document intended use, evaluation conditions, limitations and excluded applications [42]. Farmers and advisers also need understandable alerts: the suspected taxon, evidence quality, reason for review and recommended next observation. A system should distinguish missing observations from a verified absence of pests. These interface requirements follow from the consequences of uncertainty, rather than from model architecture alone.

Access and governance influence adoption. Small farms may face disproportionate costs for connectivity, repairs and specialist support. Cooperative monitoring and extension services offer possible routes to shared infrastructure, but these arrangements require local evaluation. Data ownership, image reuse and location access should be agreed before collection. Long-term

accountability also requires a responsible party for correcting false alerts, documenting model changes and withdrawing unsupported recommendations.

Maintenance should include periodic expert audits of both accepted and rejected observations. Reviewing only uncertain predictions can conceal confidently wrong classifications and distort subsequent training data. A new model should be compared with the deployed version on a fixed reference set and recent field observations before replacement. Retaining earlier versions permits reconstruction of decisions when a later investigation identifies systematic errors.

VI. DEVELOPMENTS THROUGH 2026 AND RESEARCH PRIORITIES

8.1 Transferable representations and expert knowledge

InsectNet demonstrates a global-to-local strategy combining large-scale self-supervised learning with expert-verified regional data. Its published study covers 2,526 insect species and includes uncertainty-aware prediction [45]. This is a significant advance in breadth, but recognition across many taxa does not establish equal performance for every life stage, location or camera. The value of regional adaptation depends on whether local validation captures the insects and conditions that matter operationally.

Vision-language learning provides another route to transferable representations [46]. BioCLIP uses biological images and taxonomic information to learn representations across the tree of life, offering a foundation for more specialized identification tasks [47]. Promptable segmentation systems can assist annotation and object delineation, although their general-purpose success does not by itself establish reliable segmentation

of tiny, overlapping insects [48]. These methods should reduce repetitive annotation while preserving expert control over diagnostic labels.

By 2026, BugNet demonstrated a scalable approach to assembling insect imagery and training hierarchical classifiers [38]. The contemporary identification literature also identifies persistent difficulties involving dataset coverage, affordability and translation from controlled conditions to field use [49]. Increasing model size therefore remains only one possible response to the problem. Better sampling, meaningful rejection and cheaper maintenance may produce greater practical gains.

Language models could help advisers retrieve descriptions and interpret monitoring summaries. Retrieval-augmented generation provides a technical mechanism for consulting external material [50]. In pest management, its proposed use should be restricted to traceable, maintained agronomic sources, with uncertainty and location dependence made explicit. Fluent text cannot establish a diagnosis or replace confirmation of current crop-specific guidance.

8.2 Priorities for stronger evidence

The next generation of studies should prioritize developmental stages and understudied taxa; independent tests across regions and seasons; standardized counting and uncertainty metrics; and prospective comparisons of crop, economic and ecological outcomes. Shared datasets should retain original observation groups, negatives and diagnostic limitations. Monitoring programmes should record beneficial insects sufficiently well to evaluate whether management protects their functions. These priorities follow from the recurring separation between technically successful recognition and demonstrated decision value.

The evidence base also limits this review. Targeted English-language selection can miss relevant regional work, and publication incentives may favour successful models over failed deployments. Incompatible taxonomies, image acquisition procedures and metric definitions prevent a defensible pooled accuracy estimate. Consequently, the figures present selected published measurements or explicitly identified illustrations, rather than a ranking of technologies. Conclusions concern recurring methodological requirements, not a quantified average benefit across agriculture.

VII. CONCLUSION

AI has made insect surveillance more scalable through image recognition, automated trapping and complementary sensors. Published research supports useful detection and identification under defined conditions, with important advances in transfer learning and uncertainty handling through 2026. Reliable pest management nevertheless requires an additional chain of evidence connecting observations to representative population estimates, economically justified thresholds

and measured outcomes. The most defensible role for AI is consequently an accountable component of IPM: it accelerates observation, exposes uncertainty, directs expert attention and supports decisions whose agricultural and ecological consequences are independently evaluated.

REFERENCES

- [1] Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3, 430–439. <https://doi.org/10.1038/s41559-018-0793-y>
- [2] Oerke EC. (2006). Crop losses to pests. *The Journal of Agricultural Science*, 144, 31–43. <https://doi.org/10.1017/S0021859605005708>
- [3] Deutsch CA, et al. (2018). Increase in crop losses to insect pests in a warming climate. *Science*, 361, 916–919. <https://doi.org/10.1126/science.aat3466>
- [4] Stern VM, Smith RF, van den Bosch R, Hagen KS. (1959). The integrated control concept. *Hilgardia*, 29, 81–101. <https://doi.org/10.3733/hilg.v29n02p081>
- [5] Barzman M, et al. (2015). Eight principles of integrated pest management. *Agronomy for Sustainable Development*, 35, 1199–1215. <https://doi.org/10.1007/s13593-015-0327-9>
- [6] Høye TT, et al. (2021). Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences*, 118, e2002545117. <https://doi.org/10.1073/pnas.2002545117>
- [7] Preti M, Verheggen F, Angeli S. (2021). Insect pest monitoring with camera-equipped traps: strengths and limitations. *Journal of Pest Science*, 94, 203–217. <https://doi.org/10.1007/s10340-020-01309-4>
- [8] Wu X, Zhan C, Lai YK, Cheng MM, 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*, 8787–8796. <https://doi.org/10.1109/CVPR.2019.00899>
- [9] Wang QJ, et al. (2020). Pest24: A large-scale very small object data set of agricultural pests for multi-target detection. *Computers and Electronics in Agriculture*, 175, 105585. <https://doi.org/10.1016/j.compag.2020.105585>
- [10] Gebu T, et al. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. <https://doi.org/10.1145/3458723>
- [11] Roberts DR, et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929. <https://doi.org/10.1111/ecog.02881>

- [12] He K, Zhang X, Ren S, Sun J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- [13] Thenmozhi K, Reddy US. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, 164, 104906. <https://doi.org/10.1016/j.compag.2019.104906>
- [14] Alves AN, Souza WSR, Borges DL. (2020). Cotton pests classification in field-based images using deep residual networks. *Computers and Electronics in Agriculture*, 174, 105488. <https://doi.org/10.1016/j.compag.2020.105488>
- [15] Ding W, Taylor G. (2016). Automatic moth detection from trap images for pest management. *Computers and Electronics in Agriculture*, 123, 17–28. <https://doi.org/10.1016/j.compag.2016.02.003>
- [16] Gao X, Xue W, Lennox C, Stevens M, Gao J. (2024). Developing a hybrid convolutional neural network for automatic aphid counting in sugar beet fields. *Computers and Electronics in Agriculture*, 220, 108910. <https://doi.org/10.1016/j.compag.2024.108910>
- [17] Bjerger K, et al. (2021). An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors*, 21(2), 343. <https://doi.org/10.3390/s21020343>
- [18] Ren S, He K, Girshick R, Sun J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28. <https://arxiv.org/abs/1506.01497>
- [19] Redmon J, Divvala S, Girshick R, Farhadi A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- [20] Lin TY, Goyal P, Girshick R, He K, Dollár P. (2017). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision*, 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>
- [21] Fuentes A, Yoon S, Kim SC, Park DS. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9), 2022. <https://doi.org/10.3390/s17092022>
- [22] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [23] Albanese A, Nardello M, Brunelli D. (2021). Automated pest detection with DNN on the edge for precision agriculture. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 11(3), 458–467. <https://doi.org/10.1109/JETCAS.2021.3101740>
- [24] Dosovitskiy A, et al. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *International Conference on Learning Representations*. <https://openreview.net/forum?id=YicbFdNTTy>
- [25] Rustia DJA, Lin CE, Chung JY, Zhuang YJ, Hsu JC, Lin TT. (2020). Application of an image and environmental sensor network for automated greenhouse insect pest monitoring. *Journal of Asia-Pacific Entomology*, 23(1), 17–28. <https://doi.org/10.1016/j.aspen.2019.11.006>
- [26] Rustia DJA, et al. (2021). Automatic greenhouse insect pest detection and recognition based on a cascaded deep learning classification method. *Journal of Applied Entomology*, 145, 206–222. <https://doi.org/10.1111/jen.12834>
- [27] Binns MR, Nyrop JP. (1992). Sampling insect populations for the purpose of IPM decision making. *Annual Review of Entomology*, 37, 427–453. <https://doi.org/10.1146/annurev.en.37.010192.002235>
- [28] Rydhmer K, et al. (2022). Automating insect monitoring using unsupervised near-infrared sensors. *Scientific Reports*, 12, 2603. <https://doi.org/10.1038/s41598-022-06439-6>
- [29] Tschalkner M, et al. (2024). Multisensor data fusion for automatized insect monitoring (KInsecta). *arXiv preprint, arXiv:2404.18504*. Not peer-reviewed in the cited version. <https://arxiv.org/abs/2404.18504>
- [30] Chiavassa JA, et al. (2024). The Field Automatic Insect Recognition-Device—A non-lethal semi-automatic Malaise trap for insect biodiversity monitoring: Proof of concept. *Ecology and Evolution*, 14(12), e70642. <https://doi.org/10.1002/ece3.70642>
- [31] Årje J, et al. (2020). Automatic image-based identification and biomass estimation of invertebrates. *Methods in Ecology and Evolution*, 11, 922–931. <https://doi.org/10.1111/2041-210X.13428>
- [32] van Klink R, et al. (2022). Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution*, 37(10), 872–885. <https://doi.org/10.1016/j.tree.2022.06.001>
- [33] Koh PW, et al. (2021). WILDS: A benchmark of in-the-wild distribution shifts. *Proceedings of Machine Learning Research*, 139, 5637–5664. <https://proceedings.mlr.press/v139/koh21a.html>

- [34] Saito T, Rehmsmeier M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- [35] Guo C, Pleiss G, Sun Y, Weinberger KQ. (2017). On calibration of modern neural networks. *Proceedings of Machine Learning Research*, 70, 1321–1330. <https://proceedings.mlr.press/v70/guo17a.html>
- [36] Saadati M, et al. (2024). Out-of-distribution detection algorithms for robust insect classification. *Plant Phenomics*, 6, 0170. <https://doi.org/10.34133/plantphenomics.0170>
- [37] Badirli S, Picard CJ, Mohler G, Richert F, Akata Z, Dundar M. (2023). Classifying the unknown: Insect identification with deep hierarchical Bayesian learning. *Methods in Ecology and Evolution*, 14, 1515–1530. <https://doi.org/10.1111/2041-210X.14104>
- [38] Grele A, Richards LA. (2026). BugNet: A rapid and scalable pipeline for automated insect monitoring using hierarchical data. *Frontiers in Ecology and Evolution*, 14, 1750931. <https://doi.org/10.3389/fevo.2026.1750931>
- [39] Pedigo LP, Hutchins SH, Higley LG. (1986). Economic injury levels in theory and practice. *Annual Review of Entomology*, 31, 341–368. <https://doi.org/10.1146/annurev.en.31.010186.002013>
- [40] Dainese M, et al. (2019). A global synthesis reveals biodiversity-mediated benefits for crop production. *Science Advances*, 5(10), eaax0121. <https://doi.org/10.1126/sciadv.aax0121>
- [41] Vinatier T, Pérez-López E. (2026). AI for accurate insect pest monitoring: A path toward resilient agriculture. *PLOS Sustainability and Transformation*, 5(1), e0000216. <https://doi.org/10.1371/journal.pstr.0000216>
- [42] Mitchell M, et al. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. <https://doi.org/10.1145/3287560.3287596>
- [43] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- [44] Lundberg SM, Lee SI. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774. <https://arxiv.org/abs/1705.07874>
- [45] Chiranjeevi S, et al. (2025). InsectNet: Real-time identification of insects using an end-to-end machine learning pipeline. *PNAS Nexus*, 4(1), pgae575. First published online 27 December 2024. <https://doi.org/10.1093/pnasnexus/pgae575>
- [46] Radford A, et al. (2021). Learning transferable visual models from natural language supervision. *Proceedings of Machine Learning Research*, 139, 8748–8763. <https://proceedings.mlr.press/v139/radford21a.html>
- [47] Stevens S, et al. (2024). BioCLIP: A vision foundation model for the tree of life. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19412–19424. https://openaccess.thecvf.com/content/CVPR2024/html/Stevens_BioCLIP_A_Vision_Foundation_Model_for_the_Tree_of_Life_CVPR_2024_paper.html
- [48] Kirillov A, et al. (2023). Segment anything. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4015–4026. https://openaccess.thecvf.com/content/ICCV2023/html/Kirillov_Segment_Anything_ICCV_2023_paper.html
- [49] Chakrabarty S, et al. (2026). Application of artificial intelligence in insect pest identification—A review. *Artificial Intelligence in Agriculture*, 16(1), 44–61. <https://doi.org/10.1016/j.aiia.2025.06.005>
- [50] Lewis P, et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474. <https://arxiv.org/abs/2005.11401>