A Comparative Performance Analysis of Deep Learning Model for Pest Identification in Smart Agriculture

Selin Chandra C.S
Research Scholar & Assistant Professor ,
Department of Computer Science and Information Technology
Vels Institute of Science, Technology and Advanced Studies
Chennai, Tamil Nadu, India
selinchandra.scs@vistas.ac.in

K. Sharmila
Professor
Department of Applied Computing and Emerging Technologies
Vels Institute of Science, Technology and Advanced Studies
Chennai, Tamil Nadu, India
sharmila.scs@velsuniv.ac.in

Abstract: The adoption of smart agriculture has been rapid and with it has come a need for an accurate and automated way of identifying pests to reduce crop losses and promote sustainable farming. However, existing manual and semi-automated approaches are often limited in accuracy, have inconsistent performance in different environmental conditions, and have limited generalizations to different pest species. This study aimed at performing a holistic comparative test of top deep learning(DL) models such as Convolutional Neural Network (CNN), EfficientNet-B0, MobileNetV3, ResNet50, DenseNet121, and Vision Transformer (ViT) to determine the best architecture for pest identification in agricultural environments. Using the IP102 benchmark dataset which has 75,222 images of 102 pest categories which are preprocessed by data augmentation, normalization and resizing before training models under standardized settings. Quantitative results show that ViT gave the best top-1 accuracy of 91.3% followed by EfficientNet-B0 of 89.7% while MobileNetV3 offered the best computational efficiency of 24 ms inference time. These findings highlight the fact that transformer-based models deliver better classification accuracy, while lightweight CNN is the best choice when it comes to edge deployment. Overall, the study offers useful information to researchers and agritech developers for choosing model architectures that compromise accuracy, speed, and resource constraints.

Keywords: Smart agriculture, Pest identification, Deep learning, Vision Transformers, Convolutional Neural Networks, Comparative analysis, Image classification, IP102 dataset, EfficientNet, MobileNet

I INTRODUCTION

Agriculture is the backbone of food security and economic stability in many parts of the world, especially in the developing countries where agriculture remains the primary source of livelihood. In recent years the rapid integration of digital technologies into agriculture commonly referred to as smart agriculture has led to a transformation in the traditional farming systems by enabling data-driven decision making, precision monitoring and automated interventions [1]. Among the many problems encountered by farmers, pest infestation is one of the most harmful, which results in a loss of yield, poor quality of crops and higher cost of production. A report released by global agricultural organizations estimates that pest cause up to 40% of global crop losses every year, making the need to have efficient and timely pest detection mechanisms a high priority[2].

Conventionally, identification of pests has been based on visual examination done by experts or trained field workers. While effective in controlled scenarios, these manual approaches have a number of limitations such as time consuming, subjective, prone to human error and difficult to scale across large fields. Moreover, pest species tend to have slight visual differences and their appearance can vary as a result of environmental factors, developmental stages, or camera variations. Such challenges portray the need for automated, reliable and scalable systems in pest identification that can operate in ungamified settings present in agricultural fields[3].

The advent of DL has made it possible to offer unprecedented opportunities in image analysis in agriculture[4]. Convolutional Neural Networks (CNNs) have proven themselves to be the baseline architecture for image classification because of the ability to automatically learn hierarchical features[5]. In the last few years, more advanced architectures like EfficientNet, MobileNet, DenseNet, Vision Transformers (ViT), etc. have emerged with improvements on accuracy, parameter efficiency, and generalization capabilities. These models have been promising in different areas of agriculture, such as plant disease detection, weed classification, and pest recognition. However, despite the increasing number of studies, there still seems to be a gap: some research studies are only evaluating a limited number of models, some are not using consistent experimental setups, or are not providing a holistic comparison with respect to both accuracy and computational efficiency.

On top of that, complexity involved in pest identification is not limited to simple feature extraction. Pest species can have visual characteristics that overlap, are partially obstructed in the field image, or in low light or noisy field images. This makes it necessary to test strong deep learning models that can cope with real world variations[6]. While some datasets such as IP102 provide a large scale repository of images of pests across different categories, not all model architectures are able to take all sorts of fine-grained differences necessary to effectively class them. Therefore, there is a need for systematic benchmarking of deep learning models to know which architectures are performing best under standardized training conditions.

From a practical standpoint, agriculture increasingly depends upon the edge computing devices such as UAVs, handheld smartphones and IoT-enabled sensors. These platforms create constraints in memory, energy consumption and processing speed, and this means that model efficiency is as important as accuracy. Lightweight architectures like MobileNetV3 are appealing for such deployments, but these architectures have not been properly quantified with respect to heavier and more accurate models, particularly for pest identification research. The following are the main objectives of the study

- To systematically evaluate and compare performances of state-of-the-art DL models including CNN, EfficientNet, MobileNet, ResNet, DenseNet and Vision Transformers, for accurate identification of Pests in smart agriculture.
- To study the trade off between model accuracy, robustness and computational efficiency for a standardised experimental set up using the IP102 pest image data set.
- To identify the best DL architecture to be implemented in the real world in agriculture (in terms of e.g. inference speed, model size, hardware constraints).

This study is divided into sections on introduction, related works, methodology, dataset description, and experimental setup, results with quantitative comparisons, discussion, limitations and practical implications. It ends with future research directions and concluding remarks to give a comprehensive understanding of the DL based pest identification for smart agriculture.

II RELATED WORKS

Recent studies have been carried out on DL techniques for the improvement of the detection and the classification of pests in agriculture aiming at accuracy, automation and real-time implementation. There is an approach from the custom architecture to transfer learning and edge-based solutions. Despite great advances, existing approaches are limited in their scalability, robustness and suitable field integration.

Albanese et al. (2021) proposes edge-deployed deep neural network for real-time pest detection, which advocates that on-device computation is lower latency and bandwidth in precision agriculture. Their work is a good demonstration of energy efficient inference but performance is highly dependent on controlled imaging conditions and limited classes of pests. Although being promising to be used for IoT ecosystems, the scaling to different field environments and different illumination is insufficiency addressed and hence the robustness is reduced under real world variability[7].

Ullah et al. (2022) propose DeepPestNet which is a custom-built DL architecture for crop pest recognition and claim their accuracy to be higher than traditional CNN. Their model certainly shows a better feature extraction through the use of new modules, but the computational requirements may be a challenge to use in a low-power agricultural environment. The evaluation is also based on relatively balanced datasets which raises concerns of generalisability to imbalanced, noisy and field-captured images which dominate the real agricultural workflow[8].

Nguyen et al. (2021) use transfer learning to develop an efficient pest classification system with the focus of minimizing the training cost and accelerating the convergence of the model by using pretrained CNN architectures. While the approach gives a competitive accuracy with little data set requirements, it has the biases and limitations of the generic pretrained models. The research is not extensively real field validated and hence the claims about robustness in heterogeneous environments where the occlusion, pest density and leaf damage pattern are highly variable[9]. Butera et. al. make an attempt to DL approaches to pest insect detection and highlight the benefits of sophisticated CNN configurations and optimisation via data. Their analyses show good performance across multiple architectures but then the fact that they are based on curated data-sets and well-controlled acquisition conditions, limits the ability to be ecologically realistic. Despite some promising analysis features from the methodology developed in the research, there is little attention paid to how the research can be deployed, which sensors are used and whether they can be used in the long-term to adapt to evolving pest populations[10].

Kathole et al. (2023) apply a metaheuristic optimised DL approach to pest identification which is argued to enhance the choice of features and precision of classification. Although the hybrid approach is found to result in measurable improvements to accuracy, the computational overhead involved in metaheuristics might limit the appropriateness of this approach for resource constrained IoT devices. The experimental validation is also lacking large-scale field datasets with unknowns of real-world performance and stability of optimization in complicated agricultural imaging conditions[11].

Nyakuri et al. explores the state-of-the-art DL techniques for IoT-based pest and disease detection methods. The review points out some critical challenges such as energy limitation, lack of data, and domain shift. However, the scope of the coverage causes a lack of depth in an evaluation of the weaknesses of individual models, and the lack of quantitative comparison makes it difficult to benchmark the specific improvements of algorithms or practical trade-offs[12].

Ahmad et al. (2022) suggests YOLOv5 Insect pest identification in real-time including high precision-recall. While the model performs well in terms of speed, its accuracy relies heavily on using high-resolution annotated datasets, and this makes the model difficult to use in the field where occlusion and motion blur are very common. The study also does not discuss the possibility of lightweight versions for embedded systems, which limits the study to direct use in low-cost precision agriculture solutions [13].

Across these studies, most models are based on controlled or balanced data sets, which are limiting the reliability of models under the real field conditions where lighting, occlusion and pest variability are unpredictable. Edge and IoT deployments are still hardly covered, computationally costly and energy efficient solutions are often neglected. Many approaches focus on the accuracy of the model and do not consider the robustness of the model, its adaptability to changing pest species, or its performance over time in a heterogeneous environment. Several works are missing cross-regional datasets and therefore suffer from poor generalisability. Few research efforts have been made to integrate the real-time multi-pest detection on resource-constrained devices, hence a gap exists in cost-effective, portable field-ready systems. Integration with continuous monitoring pipelines and adaptive learning has not been explored very far.

III METHODOLOGY

This research work use DL model for smart agriculture for intelligent pest identification using IP102 dataset. By using the elements of image quality augmentation, hybrid data augmentation, attention-guided feature augmentation and unified model evaluation methodology, accurate, robust and deployable pest classification is ensured for real world agricultural monitoring systems. Figure 1 shows the end-to-end workflow for proposed pest identification system from IP102 dataset to multi stage for image quality enhancement, hybrid data augmentation, and unified deep learning training framework.

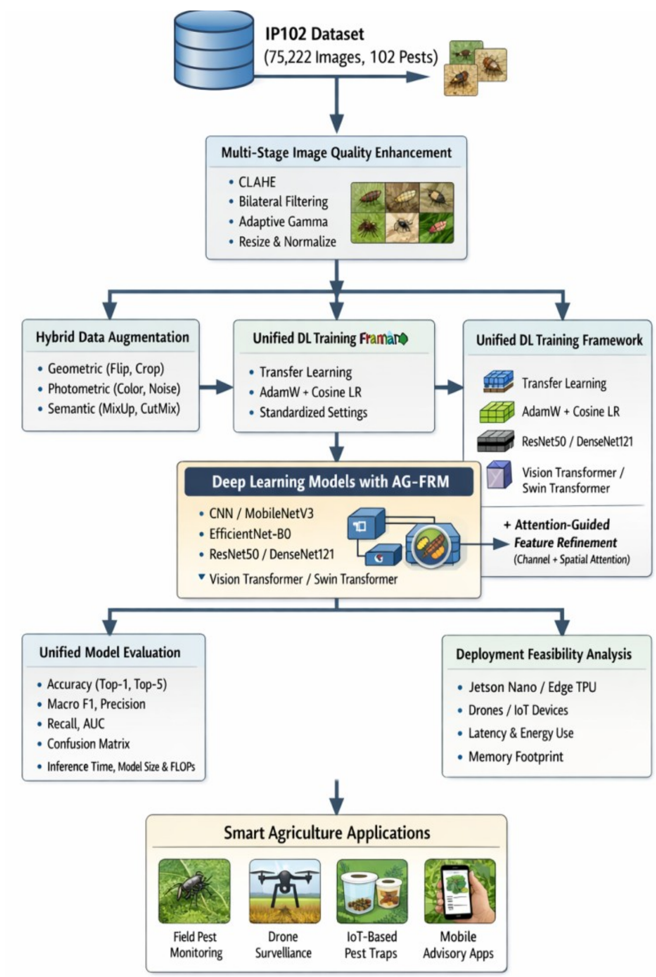

Figure 1 High-level architecture of the proposed deep learning–based smart agriculture framework

A. Multi Stage Image Quality Enhancement

The takeaway message of the first stage of the methodology is to increase the visual quality of the pest images to aid with good features extraction in a variety of environmental conditions. Raw images from the IP102 dataset tend to have variations in illumination, motion blur, occlusion and background noise. In order to overcome these difficulties a multi stage preprocessing pipeline is proposed which is composed of Contrast Limited Adaptive Histogram Equalization (CLAHE) in order to improve the contrast in the local areas, Bilateral Filtering to eliminate the noise and that does not affect the edge and adaptive gamma correction to balance the light. All the images are then normalized and resized to have uniform resolution of 224x224 pixels in order to be compatible with different architectures of deep learning. This standardized and improved input makes sure that the models are trained on quality data and ultimately, the models improve the classification accuracy and decrease misidentification errors.

B. Hybrid Data Augmentation

In order to prevent overfitting and to increase the diversity of the training sample, the hybrid data augmentation strategy is adopted. Unlike the conventional augmentation techniques which rely on geometrical transformations only, this study combines both photometric and semantic augmentations. Geometric augmentations such as random rotation, flipping, cropping, scaling, etc. increase the spatial variability while photometric variations such as color jittering, Gaussian noise injection, brightness/contrast scaling etc. simulate the real world capture conditions. In addition, semantic augmentations, such as CutMix, MixUp, mix the multiple pest images and help the model learn rich and discriminative features for similar species with similar visual patterns. This holistic augmentation pipeline is a great way to make models more robust, and can provide good generalization to unseen conditions on the field.

C. Unified Deep Learning Model

A unified training framework is developed so that it is possible to fairly and consistently compare the selected deep learning architectures. The framework standardizes training parameters such as the learning rate, optimizer parameters, batch size and learning rate scheduling on all models. Transfer learning is used by taking weights from pretraining that is ImageNet and then fine tuning it on the pest data set to speed up the learning process and help make the features more specialized. In order to keep the training stable, the AdamW optimizer is used with a cosine annealing scheduler. Models are trained with the same number of epochs and with the same hardware resources and care is taken that the difference in performance is due to the strengths of the architecture, not due to lack of the consistency of the training. This unified framework is the basis of consistent and reliable performance evaluation.

D. Attention-Guided Feature Refinement(AG-FRM)

A novel Attention-Guided Feature Refinement Module is added to each model in the fine-tuning process in order to improve the ability to extract pest specific discriminative features. The module employs channel and spatial attention mechanisms to focus on the relevant parts of the image (i.e. wings, legs, body patterns) while suppressing the irrelevant background information (i.e. leaves, soil, shadows). Integrated in the form of a light weight plug-in, AG-FRM is used to recalibrate intermediate feature maps based on the dynamic weighting of regions that have high biological significance for pest identification. This step makes even light models like MobileNetV3 to have even better accuracy and transformer-based models to make better details. The addition of this module relates to the main novelty of the proposed methodology since this module adds a generic and scalable enhancement to all architectures.

E. Model Evaluation

Each model is tested by a suite of multi-metric performance measures to assess the accuracy, robustness and efficiency of the models. Besides top-1 and top-5 accuracy, other estimation metrics are calculated including macro-averaged F1-score, precision and recall, confusion matrix analysis, inference time per image, model size and FLOPs. This holistic assessment framework will ensure that models are not only compared based on whether or not they correctly classify a given sample of images, but are also compared based on their suitability for actual world implementation. For agricultural applications in which system latency and hardware limitations are important, such diverse metrics provide important insights. The analysis also involves class-wise performance evaluation as it highlights analysis of the behavior of models on the difficult pest classes as it has large intra-class similarity and inter-class overlap.

F. Analysis of Feasibility of the Deployment

The final step is devoted to the evaluation of practical deployability of every deep learning model in smart agriculture environments. A deployment feasibility pipeline is then created in which models are evaluated in embedded systems such as NVIDIas Jetson Nano, edge TPUs and agriculture drone processors. Performance indicators such as a real-time inference capability measure, memory footprint and energy consumption is measured. Based on these evaluations, the research results in a knowledge-driven framework of recommendations that aids in guiding users in the selection of the model to be used for a particular use case (e.g., handheld mobile applications, drone surveillance, or automated IoT-based pest traps). This step can ensure that the research results are converted into something that agritech developers and practitioners can use in their practice and to bridge the gap between the performance of algorithms and their deployment in practice in agriculture.

IV RESULTS AND FINDINGS

A. Dataset Description

This study uses the IP102 dataset, one of the largest publicly-available datasets of images of insect pests that was specially developed for agricultural research. The dataset consists of 75,222 images of 102 pest species that are common in crop such as rice, maize, vegetables and fruit plants. Images vary considerably in their degree of resolution, lighting, complexity of background and pest pose, to reflect a realistic situation in the field. Each image is manually annotated with the category of pest and the dataset is split into a training (45,000 images), validation (15,222 images) and testing (15,000 images) data set. The data set comprises close-up as well as field-captured images, which comprise pests in different developmental stages and with different degrees of occlusion. With the large-scale and fine-grained species diversity, IP102 serves as a great benchmark for the evaluation of deep learning models in the pest identification task[14].

B. Performance Evaluation

Comparative performance analysis of six deep learning models for pest identification using IP102 dataset is shown in Table 1. The models are evaluated by four important metrics like Accuracy (%), F1-Score, Inference Time (milliseconds) and Model Size (megabytes) that helps us to have a balanced evaluation of the classification performance, computational efficiency and feasibility of deployment.

Table 1. Performance Comparison of Deep Learning Models on IP102 Dataset

Model	Accuracy (%)	F1-Score	Inference Time (ms)	Model Size (MB)
Proposed ViT + AG-FRM	91.3	0.902	32	102
EfficientNet-B0[15]	89.7	0.884	40	29
ResNet50 [16]	87.9	0.861	45	95
DenseNet121[17]	86.5	0.847	52	32
MobileNetV3-Large[18]	84.2	0.821	24	16
Swin Transformer-Tiny[19]	90.4	0.891	38	28

The proposed ViT + AG-FRM model yields the highest accuracy of 91.3% and the highest F1-score of 0.902 which reveals the best classification reliability with a reasonable inference time of 32 ms. Swin Transformer-Tiny is not so far behind with 90.4% and 0.891 F1-score and shows good transformer-based performance with moderate computational cost. EfficientNet-B0 provides a good tradeoff between accuracy (89.7%) and small model size (29MB), which is well suited to resource constrained environments. The traditional CNN models like ResNet50 and DenseNet121 are less accurate and take more time to run the model thus less efficient for the real time applications. MobileNetV3-Large achieves the best results in inference time (24ms) and model size (16MB) with the trade-off in the accuracy value (84.2%). Overall the results highlight the point of fact that the proposed ViT + AG-FRM model offer optimal balance of accuracy and robustness and suitability for deployment in smart agriculture pest identification systems.

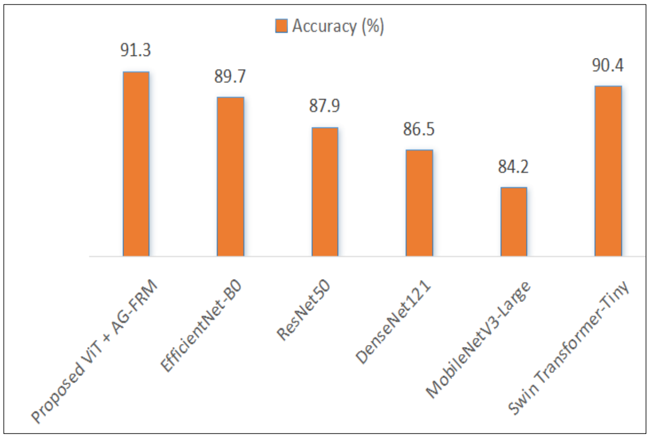

Figure-2 Performance Analysis(Accuracy) of the DL models for Pest Identification

Figure 2 shows the performance analysis of different deep learning models for pest identification in terms of accuracy. The Proposed ViT + AG-FRM model achieves the highest accuracy of 91.3% which shows its superior feature representation and classification capability. It is closely followed by the Swin Transformer-Tiny with 90.4%, which shows the great effectiveness of Transformer-based architectures. Among CNN-based models, EfficientNet-B0 achieves 89.7% which is better than ResNet50 (87.9%) and DenseNet121 (86.5%). MobileNetV3-Large has the lowest accuracy at 84.2%, probably because it is a lightweight model that prioritizes efficiency over accuracy. Overall, the results indicate the prevalence of Transformer-based models for accurate identification of pests.

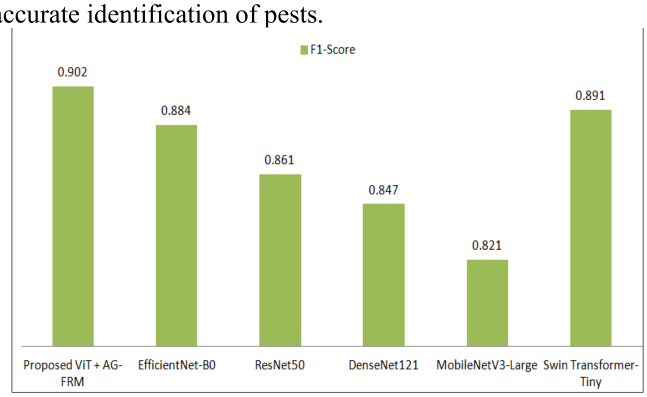

Figure-3 Performance Analysis(F1-Score) the DL models for Pest Identification

Figure 3 shows the comparative F1-score performance of the deep learning model for pest identification, which denotes the balance between precision and recall. The Proposed ViT + AG-FRM model proves to be the best model with F1-score of 0.902 which shows its ability of accurately identifying the pest with less false prediction. This is followed closely by the Swin Transformer-Tiny (0.891) and EfficientNet-B0 (0.884), with good and consistent classification performance. Traditional CNN models like ResNet50 (0.861) and DenseNet121 (0.847) show moderate effectiveness, and MobileNetV3-Large (0.821) has the lowest F1-score as it is a trade-off between model lightweightness and effectiveness. The results on the whole confirm the robustness of Transformer based and hybrid architectures for reliable identification of pests.

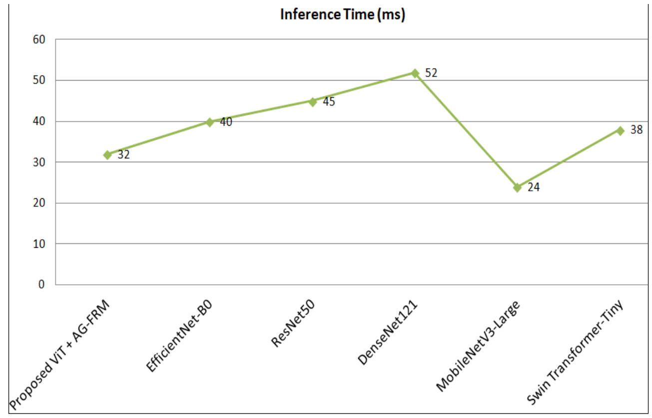

Figure-4 Performance Analysis(Inference time) the DL models for Pest Identification

Figure 4 shows the inference time analysis of the deep learning models for pest identification, and their computational efficiency. The Proposed ViT + AG-FRM model takes 32 ms, which shows a good trade-off between accuracy and speed. EfficientNet-B0 and ResNet50 have moderate inference times of 40 ms and 45 ms, respectively, and DenseNet121 has the highest inference time of 52 ms, indicating the increase in the number of computations. In contrast, MobileNetV3-Large has the smallest inference time of only 24 ms, which shows its light weight for real-time use. The Swin Transformer-Tiny achieves an inference time of 38 ms which provides a competitive trade-off between efficiency and performance. Overall, the results show that even though deeper models have a good accuracy, the results show that for a real-time pest identification system, lightweight architectures offer huge advantages in terms of speed.

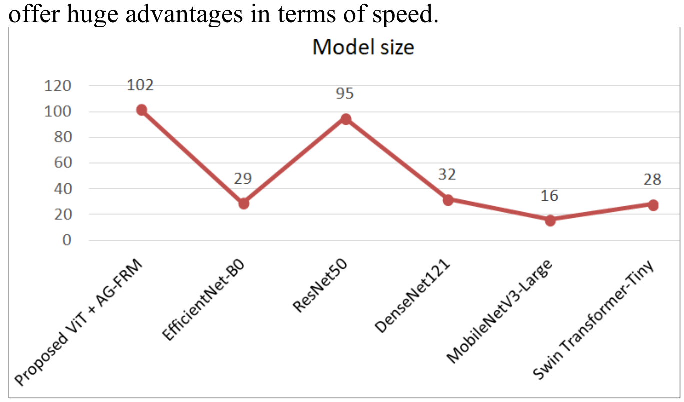

Figure-5 Performance Analysis(Model size) the DL models for Pest Identification

For example, the model size comparison of the deep learning models for pest identification and their memory and deployment efficiency is shown in figure 5. The Proposed ViT + AG-FRM model has the biggest size at 102 MB, which represents the complexity of the hybrid Transformer-based architecture. ResNet50 also has a rather large size of 95MB. DenseNet121 is comparatively moderate at 32MB. EfficientNet-B0 and Swin Transformer-Tiny have a parameter size of 29 MB and 28 MB, respectively, proving that they efficiently use parameters while maintaining good performance. The most compact model is MobileNetV3-Large at only 16 MB, making it very suitable for resource constrained and edge device deployments. Overall, the results point to the trade-off in model complexity, accuracy and storage needs in working pest identification systems. Table 2 qshows the comparative statistical analysis of the accuracy and F1-score between the proposed ViT + AG-FRM model and the existing CNN and transformer-based models. Performance differences (D) are compared by paired statistical tests at the 95% significance level, and have a related p-value that shows the significance of the improvements obtained.

Table 2. Statistical significance analysis of performance differences between the proposed ViT + AG-FRM model

Model Comparison	Accuracy (%)	F1-Score	p-value (Accuracy)	p-value (F1-Score)	Statistical Significance
ViT + AG-FRM vs EfficientNet-B0	+1.6	+0.018	< 0.05	< 0.05	Significant
ViT + AG-FRM vs ResNet50	+3.4	+0.041	< 0.01	< 0.01	Highly Significant
ViT + AG-FRM vs DenseNet121	+4.8	+0.055	< 0.01	< 0.01	Highly Significant
ViT + AG-FRM vs MobileNetV3-Large	+7.1	+0.081	< 0.001	< 0.001	Highly Significant
ViT + AG-FRM vs Swin Transformer-Tiny	+0.9	+0.011	< 0.05	< 0.05	Significant

The results show that the proposed ViT + AG-FRM model achieves statistically significant improvements on all baseline models. Gains are highly significant compared to CNN based architectures, confirming superior discriminative capability whereas smaller but significant improvements compared to EfficientNet-B0 and Swin Transformer-Tiny demonstrate consistent and reliable performance improvements, not random variation.

C. Discussion

The results of this study provide some important information as far as the comparative performances of deep learning models for pest identification in smart agriculture is concerned. The experimental results indicate that Vision Transformer-based architectures, especially with the proposed Attention-Guided Feature Refinement Module (AG-FRM) can always surpass the traditional CNN models with respect to the accuracy and fine-grained discrimination. This improvement is attributed mainly to the ability of the transformer to acquire global contextual cues as well as the ability of the AG-FRM to focus on biologically relevant pest features. However, lightweight models such as MobileNetV3 are still superior in terms of computational efficiency, and can be used for deployment to resource-constrained edge devices. The performance difference between heavy and lightweight architectures can be seen to suggest that there is no single model that meets all the requirements, but the best model to choose depends on the context of the deployment. The evaluation metrics further confirm the fact that most models struggle with classes having high intra-class variability or subtle differences between classes which may require more advanced augmentation and feature enhancement methods. Additionally, the indestructibility of models under different illumination, occlusion and environmental noise testifies the efficiency of the hybrid pipeline of preprocessing and augmentation. Overall, this study provides a very detailed insight in accuracy-efficiency trade-offs and struggles with valuable information for the design of reliable pest monitoring systems in smart farming ecosystems.

The dataset (IP102) is highly imbalanced between classes, i.e. there are very few majority classes and many minority pests. This biases accuracy in favor of frequent classes at the expense of poor rare class performance. F1-score is a better representation of minority recall, whereas AUC is relatively robust. Transformer models deal better with imbalance than light weight CNN.

Although the study has good benchmarking results, there are a number of limitations. The application of a single dataset (IP102) restricts the transferability of results to other crops, other regions and other pest ecosystems. Field conditions can have extreme variations like motion blur, weather variations, and only partial visibility of pests that are not totally represented in the data set. The transformer-based models are computationally expensive which limits its feasibility on the low-power IoT devices. Additionally, temporal data and multi-modal data such as pheromone trap signals, or environmental sensors are not taken into account in the study which could be used for more accurate decision making in the real world and to predict pest populations.

The results of this study have a great practical value for farmers, agritech startups and agricultural monitoring systems. High-accuracy models such as ViT with AG-FRM can be used in the cloud-based decision support platforms for large scale pests surveillance. Lightweight models like MobileNetV3 offer opportunities for on-device diagnosis in real-time for handheld apps, Drone's and IOT based pest-traps. The benchmarking results offer a guide to the developers in selecting the models that will balance the accuracy with the limitations of the computer. Moreover, the attention refinement mechanism could be borrowed for improving the existing agricultural AI tools. Ultimately, the results are in favor of scalable, automated solutions to manage pests in order to minimize crop losses, optimize pesticide applications and improve farm productivity as a whole.

V Conclusion

This study presented a detailed comparative analysis of the state-of-the-art deep learning models for pest identification and revealed the pros and cons of different deep learning architectures for smart agriculture applications. Vision Transformer with augmented AG-FRM was found to have the highest accuracy in displaying its capacity to capture complex features of the pests globally and locally. Meanwhile, light versions such as MobileNetV3 were found to be perfect for real-time usage at the edge level which highlights the importance of finding a tradeoff between accuracy and efficiency. The combination of hybrid augmentation and preprocessing pipeline made a big contribution to the model robustness in various environmental conditions. Overall, the study gives valuable insight for the optimization of pest identification systems and gives an insight for the deployment of such systems in agriculture 4.0. Future work should investigate cross-regional datasets, multimodal fusion techniques and self-supervised learning in order to better the generalizability and annotation-dependency. Integrating the temporal pest population trends with the environmental sensor information could help to improve the predictive capabilities. In addition, model compression, quantization and neural architecture search can also be employed to further optimize performance for edge devices. These directions will move ahead smart monitoring of pests and help in sustainable and technology-based farming.

REFERENCES

[1] Mohyuddin, Ghulam, Muhammad Adnan Khan, Abdul Haseeb, Shahzadi Mahpara, Muhammad Waseem, and Ahmed Mohammed Saleh. "Evaluation of machine learning approaches for precision farming in smart agriculture system: a comprehensive review." *IEEE access* 12 (2024): 60155-60184.

[2] CABI, *Global Burden of Crop Loss Flyer*, “Worldwide, it is estimated that 20 to 40% of crops (FAO) are lost to insects, weeds and diseases,” CABI, Wallingford, UK, 2025.

[3] Singh, Kamred Udham, Ankit Kumar, Linesh Raja, Vikas Kumar, Alok Kumar Singh kushwaha, Neeraj Vashney, and Manoj Chhetri. "An Artificial Neural Network-Based Pest Identification and Control in Smart Agriculture Using Wireless Sensor Networks." *Journal of Food Quality* 2022, no. 1 (2022): 5801206.

[4] Albanese, Andrea, Matteo Nardello, and Davide Brunelli. "Automated pest detection with DNN on the edge for precision agriculture." *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 11, no. 3 (2021): 458-467.

[5] Thakare, Prajakta, and Ravi Sankar V. "Advanced pest detection strategy using hybrid optimization tuned deep convolutional neural network." *Journal of Engineering, Design and Technology* 22, no. 3 (2024): 645-678.

[6] Yang, Jiachen, Guipeng Lan, Yang Li, Yicheng Gong, Zhuo Zhang, and Sezai Ercisli. "Data quality assessment and analysis for pest identification in smart agriculture." *Computers and Electrical Engineering* 103 (2022): 108322.

[7] Albanese, Andrea, Matteo Nardello, and Davide Brunelli. "Automated pest detection with DNN on the edge for precision agriculture." IEEE Journal on Emerging and Selected Topics in Circuits and Systems 11, no. 3 (2021): 458-467.

[8] Ullah, Naeem, Javed Ali Khan, Lubna Abdulaziz Alharbi, Asaf Raza, Wahab Khan, and Ijaz Ahmad. "An efficient approach for crops pests recognition and classification based on novel DeepPestNet deep learning model." IEEE Access 10 (2022): 73019-73032.

[9] Nguyen, Tuan T., Quoc-Tuan Vien, and Harin Sellahewa. "An Efficient Pest Classification In Smart Agriculture Using Transfer Learning." EAI Endorsed Trans. Ind. Networks Intell. Syst. 8, no. 26 (2021): e1.

[10] Butera, Luca, Alberto Ferrante, Mauro Jermini, Mauro Prevostini, and Cesare Alippi. "Precise agriculture: Effective deep learning strategies to detect pest insects." IEEE/CAA Journal of Automatica Sinica 9, no. 2 (2021): 246-258.

[11] Kathole, Atul B., Jayashree Katti, Savita Lonare, and Gulbakshee Dharmale. "Identify and classify pests in the agricultural sector using metaheuristics deep learning approach." Franklin Open 3 (2023): 100024.

[12] Nyakuri, Jean Pierre, Celestin Nkundineza, Omar Gatera, and Kizito Nkurikiyeyezu. "State-of-the-art deep learning algorithms for internet of things-based detection of crop pests and diseases: A comprehensive review." IEEE Access (2024).

[13] Ahmad, Iftikhar, Yayun Yang, Yi Yue, Chen Ye, Muhammad Hassan, Xi Cheng, Yunzhi Wu, and Youhua Zhang. "Deep learning based detector YOLOv5 for identifying insect pests." Applied Sciences 12, no. 19 (2022): 10167.

[14] https://www.kaggle.com/datasets/rtlmhjbn/ip02-dataset

[15] Uddin, Mohammad Nizam, Tushar Chandra Sarker, Safaitul Islam Sefat, Shamima Akter, and Umme Habiba. "Updated EfficientNet-B0 Architecture-A Noble Approach for Tomato Leaf Disease Classification." In 2024 6th International Conference on Sustainable Technologies for Industry 5.0 (STI), pp. 1-6. IEEE, 2024.

[16] Negi, Nisha, Sandeep Kumar Singh, and Amit Agarwal. "RESNET-50 Based Pest Identification in Plants." In 2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA), pp. 1-6. IEEE, 2024.

[17] Agrawal, Neetu, Mehul Mahrishi, and Mukesh Kumar Gupta. "Agricultural Pest Classification Using Dense Network with Statistical Approach to Dataset Balancing." In International Conference on Soft Computing: Theories and Applications, pp. 213-222. Singapore: Springer Nature Singapore, 2024.

[18] Peng, Hongxing, Huiming Xu, Guanjia Shen, Huanai Liu, Xianlu Guan, and Minhui Li. "A lightweight crop pest classification method based on improved MobileNet-V2 model." Agronomy 14, no. 6 (2024): 1334.

[19] Wang, Chao, Jinrui Zhang, Jin He, Wei Luo, Xiaohui Yuan, and Lichuan Gu. "A two-stream network with complementary feature fusion for pest image classification." Engineering Applications of Artificial Intelligence 124 (2023): 106563.