Dandao Xuebao/Journal of Ballistics
ISSN: 1004-499X
Vol. 38 No. 1 (2026)

An Automated Pyramid Attention Driven Pest Detection System with Spatial and Texture-Aware Feature Learning

R.Prabha[1*], K. Selvan[2]

[1]Research Scholar, P.G. and Research Department of Computer Science,
J.J. College of Arts and Science (Autonomous), Pudukkottai,
Affiliated to Bharathidasan University, Tiruchirappalli.
Email:prabha.pr@gmail.com

[2]Assistant Professor & Research Advisor, P.G. and Research Department of Computer Science,
J.J. College of Arts and Science (Autonomous), Pudukkottai,
Affiliated to Bharathidasan University, Tiruchirappalli.
Email:kselvanjj75@gmail.com

Corresponding Author: R.Prabha, Research Scholar, P.G. and Research Department of Computer Science, J.J. College of Arts and Science (Autonomous), Pudukkottai, Affiliated to Bharathidasan University. Tiruchirappalli.

Abstract

Pest recognition is a significant feature in the current agricultural system since the level of insect infestation defines the health of crops, their level of yield and general productivity, particularly in vegetable production where plants are more vulnerable at any stage of their development. The identification of the pests must be done as early as possible and properly in order to minimize wastage of crop and unwarranted application of pesticides. Nevertheless, the current approaches of detecting pests have a number of limitations, such as the use of manual examination, the inability to represent features, their low generalization in the absence of different environmental conditions, and the inability to differentiate between similar symptoms of pests. The traditional and standalone approaches do not tend to reflect the complex spatial and textual division of the real field images. In order to overcome these constraints, the present paper presents a new feature extraction and classification model, SPARTA-Net (Spatial Pyramid Attention Representation Texture Aggregation) to effectively detect pests in vegetable crops. The proposed method is a mixture of multi-scale spatial pyramid analysis, feature refining by attention and texture-based representation to depict global and local pest features. The novelty of the method is that it has a hybrid feature aggregation strategy in which the spatial, structural and texture features are combined into a single feature, which improves strength and classification accuracy. Experimental results demonstrate that the suggested model works well with an accuracy of 98%, precision of 97%, recall of 97%, F1-score of 97% and AUC of 99% at a low loss. These findings are indicative of its capability to identify pests in precision agriculture with accuracy and in real-time.

Keywords: Pest Detection, Deep Learning, Agriculture Systems, Vegetable Crops, Image Processing, and Feature Analysis.

1. Introduction

Insect pests in the farming of vegetable crops constitute one of the most important biological limitations, which affect directly the health of plants, their productivity, and stability of the total yield. The worst thing about these pests is that they attack the crops at every stage of growth [1], and this starts with the seedlings and continues till the fully grown plants, thereby destroying the system of the plants in the long run, unless they are prevented and controlled in time. The observable symptoms are the primary visible indicators that can be trained by deep learning models to effectively detect at an early stage and correctly the presence and severity of pests. Such pest-induced symptoms are significant data attributes employed to derive an automatic identification and categorization of the symptoms of pests in the intelligent agriculture and deep learning-based pest detection framework [2]. Manual inspection may be time consuming, labor intensive and may be easily prone to human error especially during the early stages of infestation by pests or where the symptoms of pest infestation closely resembles that of other pest species. Deep learning classifiers that have been trained on large datasets of infected and healthy leaf images, in contrast, can identify more advanced spatial variations, texture shifts, color distortions and shape deformations as a result of insect activity.

Deep learning has similarly been an extremely promising answer to complicated pattern recognition duties in agriculture particularly in identifying and differentiating pests in vegetable harvests. In the conventional farming practices, detection of insect pests and their infestations necessitating human inspection of the leaves and plant structures, which is mostly inconsistent and relies on human judgment. A good substitute is the deep learning models that are learned to automatically extract discriminative features of pest-infected leaf images without having to extract features manually. The models [3] can capture subtle variations in color, texture, shape and pattern of the lesions which the human eye would not detect or would be difficult to detect in all cases. This is especially important in the vegetable crops where in one instance that may have a number of species of pests that would create the same symptoms on the leaves and hence would not be easily differentiated by their appearance.

The deep learning models can significantly increase the accuracy of detection, enable real-time tracking, and encourage precision agriculture operations as they allow the timely identification and classification of pest attacks by exploiting the vast amounts of image-based information. Moreover, traditional machine learning algorithms are based on human-designed features, which are not robust enough to adapt to variations in lighting, leaf orientations, detailed backgrounds, and infestations at different stages. This has resulted in a desperate need of a smart and automated system capable of aiding in the correct identification and classification of insect pests in vegetable crops based on raw image data [4]. The main issue discussed in this paper is the creation of a powerful deep learning-based system that will be able to differentiate between healthy and infested leaves and correctly identify various types of pest infestations in diverse environmental/field conditions.

The reason behind this piece of work is the fact that there is an urgent need to improve the productivity of agriculture and reduction of loss of crops to insect pests in the vegetable fields. Vegetable crops are highly prone to pests attack and any slight delay in unraveling of the issue can cost the farmers a mass loss in yield and economic loss. Besides that, excessive or improper use of pesticides due to improper identification of pests causes environmental degradation, soil contamination and development of pesticide resistant pest species. Therefore, the necessity to develop an intelligent, fast, and efficient pest tracking system that could assist farmers to make the correct decision is significant. The detection systems rely on the deep learning-based pest detection systems which offers a viable solution as it can scan images of crops in an automated, precise, and scalable manner without involving human labor and expert knowledge [5, 6]. The reason why this work should be done is to integrate the application of artificial intelligence in agricultural sector in order to assist in early detection of pests, minimize the quantity of chemicals applied, improve the crop protection strategies and ultimately improve the quality and sustainability of harvest in production of vegetable crop.

The paper will have the following structure: Section 2 will be a comprehensive literature review of the established models and techniques of pest detection and classification in agriculture systems and the weaknesses and strengths of traditional machine learning as well as the latest deep learning-based solutions are summarized. Section 3 includes the description of the proposed work, including a description of the deep learning-based framework, developed to detect pests in vegetable crops, the model architecture, the preparation of data set, the feature learning strategy, and the classification mechanism. Section 4 addresses the experimental results, in which the performance of the proposed model is assessed with the help of common metrics and compared with the current state-of-the-art approaches to show their suitability and excellence. Lastly, Section 5 sums up the paper by providing a conclusion of the main findings and contribution and also provides potential future directions towards enhancing the robustness of the model, its scalability and real time implementation in precision agriculture systems.

2. Literature Review

This section will involve a critical literature review of research on existing research on pest detection and classification within agriculture and vegetable crop systems [7], namely the traditional machine learning approaches, and the innovative deep learning approaches. Various methods to identify insect pests and diagnose crop leaf diseases through image processing, feature extraction and classification algorithms have been proposed over the years. In early research, the handcrafted characteristics of color, texture and shape descriptions along with the traditional classifier were majorly used, and did not perform well in natural field applications that were complex in nature. With the advancement in deep learning especially in convolutional neural networks [8], scholars have shifted to data-driven techniques, which are automatic in learning discriminative features of raw images, and this results to high accuracy and robustness.

Wang, et al [9] introduced an in-depth overview of the recent developments in the field of deep learning algorithms deployment to identify plant diseases and pests, with a focus on the shift to more sophisticated data-driven solutions. They begin by criticizing the classical approaches, particularly that they have been based on handmade aspects and are not resistant to the complex agricultural environment. The review then goes further to analyze the application of various deep learning techniques like image classification, object detection, semantic segmentation, and change detection to successfully detect and locate plant diseases and insect pests. Deb, et al [10] offered an effective deep learning-based methodology that involves use of VGG16 architecture to carry out the automated classification of common potato leaf pests such as aphids, Colorado potato beetles, leafminers, and ladybird beetles. They have created a filtered dataset of 3,000 labeled images of real fields with agriculture and picked healthy and infected leaves of pests in their work, making them pertinent to practice. To apply the transfer learning, the authors fine-tuned the pre-trained VGG16 network, and data augmentation techniques such as rotation, flipping, zooming and contrast adjustment to improve the performance of the models and their generalization.

Bijlwan, et al [11] gave a detailed discussion of various deep learning networks to identify and categorize rice disease and pests and emphasize the relevance of all these networks in the future agricultural activities. The data in this experiment was gathered in detailed data both in the controlled experiment and in the farmer fields at and around Pantnagar, so that, the data is diverse and variably real-life. The authors speak of the performance of different deep learning architectures in detecting different symptoms of different diseases and pest infestations, and that they are demonstrating themselves to be effective at the ability to identify complex visual patterns in rice crops.

In agricultural pest detection and classification, recent studies have started to adopt deep learning models particularly convolutional neural networks (CNNs) due to the ability to automatically derive complex visual representations of crop images [12]. The techniques have proved to have promising potential in the identification of infestation of pests through the patterns of leaf damage such as discoloration, holes, change in texture, and deformity. The accuracy and localization of pests in the field setting, which is also complex, have also been boosted by state-of-the-art algorithms including transfer learning and object detectors (YOLO and Faster R-CNN) and semantic segmentation.

Although these have been developed, a number of challenges still exist in real world deployment. Among the problems is the inconsistency of environmental conditions e.g. lighting, background clutter, occlusion, and variations in image acquisition devices, which vary greatly and greatly impact model performance. In addition, there are many pest species [13], which produce aesthetically similar symptoms on leaves, and it is difficult to identify the classes correctly on the models. Another constraint to the generalization capability of deep learning models is the absence of large, well-labeled datasets, especially per vegetable crop and pest type, which are rare or scarce.

Furthermore, the models that are in existence have a tendency to utilize controlled data (or laboratory data) which is not completely representative of the actual world of an agricultural field. This creates an incongruity between the performance during experiments and practice. Many techniques [14] that concentrate on disease detection, or pest classification, however, few studies have tried to handle both in a unified model. The second problem is the inaccessibility of light models and computationally effective models which can be deployed on the edge devices to facilitate real time monitoring in farms.

In addition, most of the available systems lack the interpretability, and severity estimation, which are very important in decision making in pest management. These shortcomings reveal a major gap in the research to create robust, scalable, and real-time deep learning models capable of detecting and classifying various types of pests under different environmental conditions. Therefore, complex models integrating improved feature extraction, advanced learning strategies [15], and lifelike deployment alternatives are greatly sought after to enhance the pest detection task and permit precision agriculture.

3. Proposed Spatial Pyramid Attention Representation Texture Aggregation (SPARTA-Net) Model

The Spatial Pyramid Attention Representation Texture Aggregation (SPARTA-Net) is a complex and technically well-founded feature extraction and classification model to effectively identify pests in vegetable crops. The main advantage of SPARTA-Net is that it will overcome the shortcomings of the traditional feature extraction methods by combining spatial, textural, and multi-scale representation into a single pipeline which is capable of efficiently describing the intricate visual patterns of the pest infestations. In farm photographs, the destruction wrought by pests can take a variety of forms, as irregular holes, color changes, twisted veins, and texture variations, which vary significantly depending on the kind(s) of pest, and the environmental conditions.

The SPARTA-Net especially is created to reproduce such variations by combining spatial pyramid analysis to extract multi-scale features, attention-seeking mechanisms to focus on important regions, and texture-dependent features to characterize fine-grained patterns of leaf damage. The main strength of the method is in the opportunity to integrate multiple complementary representations of features, including spatial, structural, and textural, to a powerful aggregated feature space which leads to a better accuracy in classification, yet the interpretation is still available. Unlike the more traditional methods such as GLCM or LBP, SPARTA-Net suggests a processing pipeline where features are not only identified but optimized, prioritized and merged based on their importance creating a more discriminative representation of pest features.

The new and unique feature of SPARTA-Net is that it is hierarchical and hybrid which naturally embraces a diversity of sophisticated feature engineering concepts in a single system. The fact that the model uses a spatial pyramid to analyze the input leaf image at different resolutions and scales is in part due to the fact that it is important in both the small (e.g., small spots or small infestations) damage and the large (e.g., large leaf erosion) damage. Second, the addition of an attention mechanism into non-deep learning setting is a unique feature, with statistical weighting or saliency-related techniques highlighting those areas of the image that are more prone to having pest-related features and bring down the impact of background noise and irrelevant data.

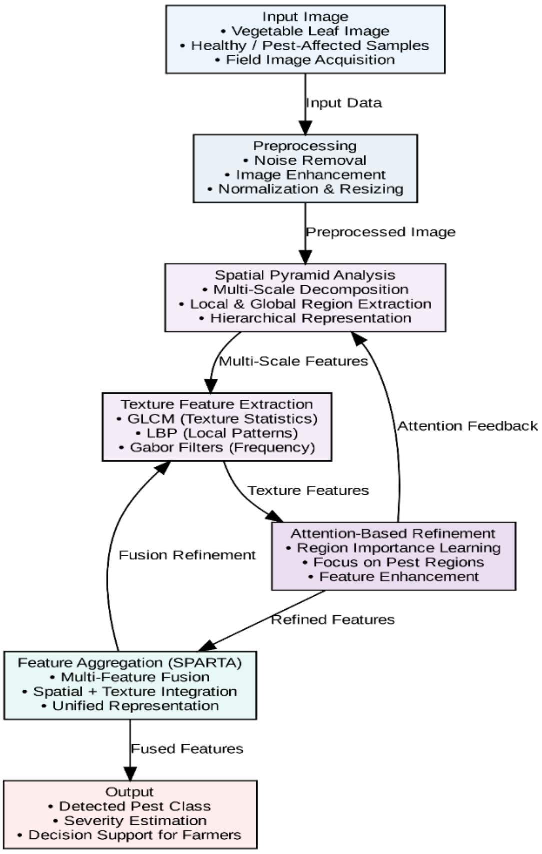

Fig 1. Overview of the proposed work

As shown in Fig 1, the SPARTA-Net workflow begins with the input image that is a raw image of a leaf image of vegetable crops in the real field environment. In the former, in the first step, the preprocessing occurs in which the image is normalized and noise is eliminated to enhance the contrast in order to make the significant features more visible. The resulting processed image is then fed onto the spatial pyramid module where it is further decomposed into different areas at different scales (such as global, regional patches and fine-grained segments). The features are also extracted at each level according to the texture descriptors such as LBP, which extract local patterns, GLCM, which extracts statistical texture properties, and Gabor filters, which extract frequency and orientation. These features that are extracted are a combination of the spatial and structural attributes of pest damage.

The attention module then weights the different regions/features based on their relevance which can be in terms of entropy measures, variance measures or saliency detector. This makes sure that areas with the most salient pest symptoms will play a larger role in the ultimate representation. The weighted features are then sent to the representation stage where the features are transformed to a normalized feature vector space using the normalization and dimensionality reduction algorithms such as PCA or feature scaling. In aggregation step, the attributes of the different scales and descriptors are pooled into a single and large feature vector which is the input to the classification module.

Structurally, SPARTA-Net can be viewed as a series of conceptual layers instead of conventional layers in a neural network, each with its particular role in the feature extraction and classification pipeline. The first layer is known as the preprocessing layer; it is the one that is concerned with image normalization and enhancement. The second layer is the spatial pyramid decomposition layer that decomposes the image into multi-scale regions. The third layer is the feature extraction layer that executes a variety of descriptors such as LBP, GLCM and Gabor filters simultaneously in an effort to extract different features. The fourth layer is the attention layer, which has weighting and selection of the important pest related regions. The fifth layer is the feature representation layer where the features extracted are converted into a structured and compact format. The sixth layer is the aggregation layer which incorporates all the features in a unit of a single vector. Lastly, the classification layer matches the aggregate features to a particular pest classes.

4. Results and Discussions

This section entails an in-depth analysis of the suggested pest detection and classification framework with a thorough analysis of experiments and comparative research. To evaluate the effectiveness of the proposed SPARTA-Net approach, various standard evaluation measures, such as accuracy, precision, recall, F1-score, AUC, and loss, are involved, to provide a comprehensive picture of its performance. Moreover, the results are contrasted with a variety of state-of-the-art deep learning models to demonstrate the gains by the proposed approach. The trends of performance are represented in terms of different graphical illustrations and analysis plots to make sure that the model is robust, reliable, and able to generalize under different conditions.

Fig. 2 shows the accuracy analysis of different state-of-the-art models and the proposed model to detect pests in vegetable crops. Accuracy refers to a model having the capacity to correctly categorize infected and healthy samples of pests. The results indicate that the classical models such as VGG16 and MobileNet possess a relatively low accuracy due to their low capacity to capture all the variations of the leaf structures that are brought about by the pests. The more sophisticated architectures like EfficientNet and Xception show superior performance, because they learn more finer feature representations. Nonetheless, the proposed model has the most accurate performance of 98 percent, indicating its better capacities to simulate discriminative features and adapt to various situations of agricultural images successfully.

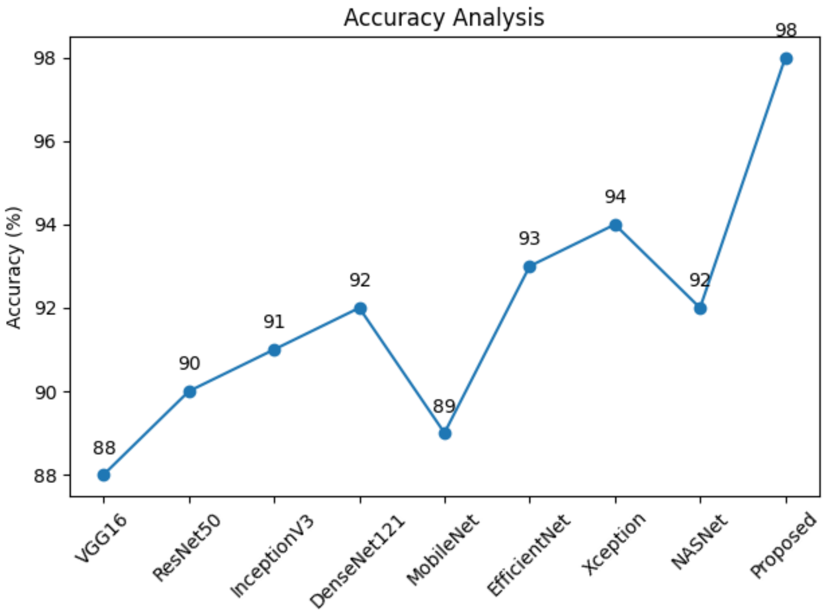

Fig 2. Accuracy analysis with state of the art models

Fig. 3 presents the precision analysis that calculates the reliability of the models in determining the pests accurately without giving false positives. In the identification of the pests, accuracy is crucial since a false alarm would lead to unwarranted use of the pesticides. The results reveal that the existing models are moderately to highly accurate but the proposed model is the most accurate at 97%. This implies that it can easily identify the pest infested regions with high accuracy compared to the healthy regions hence minimizing the occurrence of false alarms and gives more accurate predictions in the real world application in agriculture.

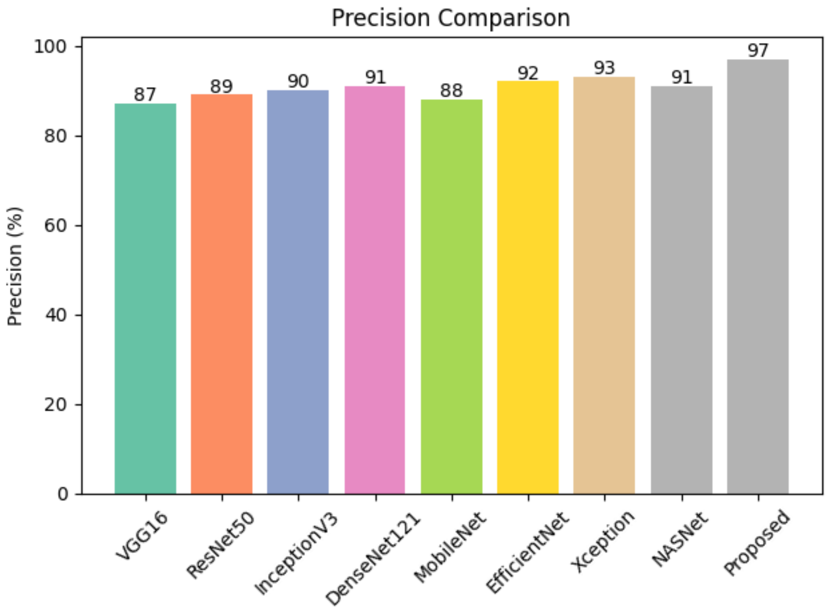

Fig 3. Precision analysis with state of the art model

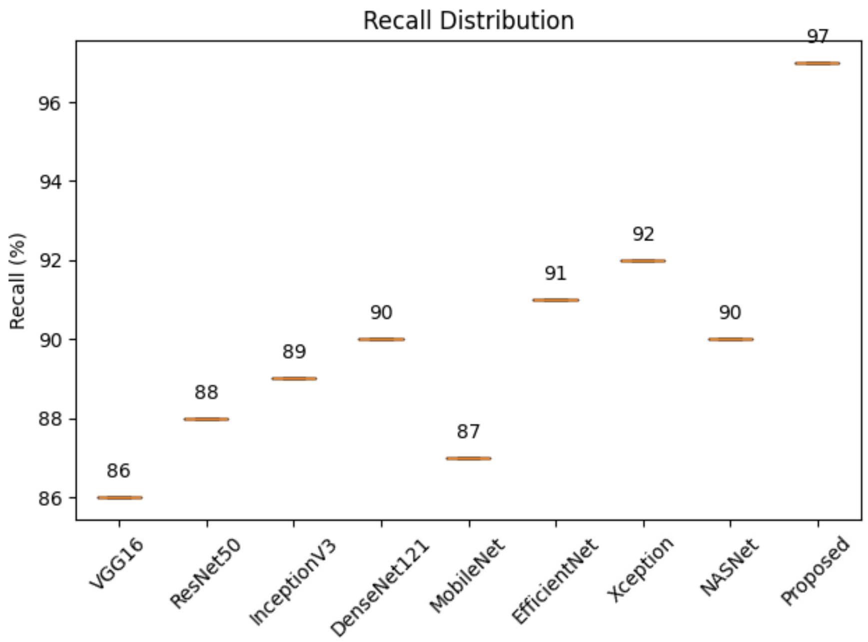

Fig 4. Recall analysis with state of the art model

Fig. 4 shows the recall analysis, which is used to measure the effectiveness of the models to identify all the actual instances of pests included in the dataset. Pest infestation that would result in essential crop losses call for high recall in order to make sure that they are not missed. This analysis shows that simple models yield worse recall as they are not sensitive to delicate pest patterns, in contrast to more sophisticated models. The proposed model can be recalled 97 times, which implies that the model is quite robust to detect nearly all pest cases, whether early or less evident ones.

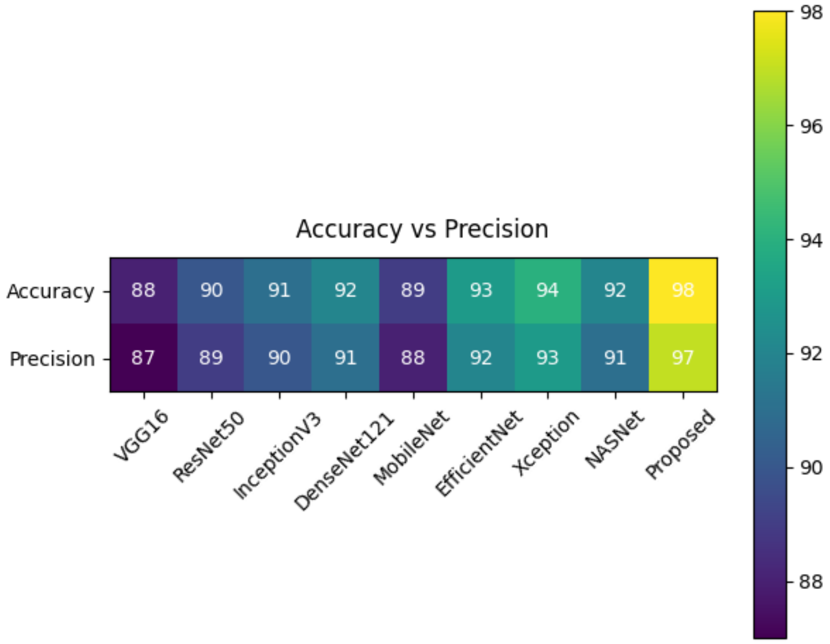

Fig 5. Accuracy Vs Precision with state of the art model

Fig. 5 is a comparative analysis of accuracy and precision, which provides a reasonable view of the overall accuracy and dependability of the predictions. The graph shows that, even though some of the models show a linear relation between these measures, slight variations indicate a tradeoff between an appropriate detection and a false negative. The fact that the proposed model has managed to attain both accuracy and precision at the same time is unique because it has shown that it can balance a high degree of accuracy and reliability, which is paramount in the real world pest detection systems. Fig. 6 shows the correlation of AUC and loss of the models evaluated. AUC is used to measure the ability of the model to differentiate between pest and non-pest classes and loss is used to show the error made in training. The results reveal that the conventional models possess moderate values of AUC and augmented loss that depicts ineffective learning. By comparison, the proposed model performs the best in terms of the highest AUC of 99% and the lowest loss of 0.10, which indicates that it has the best classification and the best learning process with a minimum error in predictions.

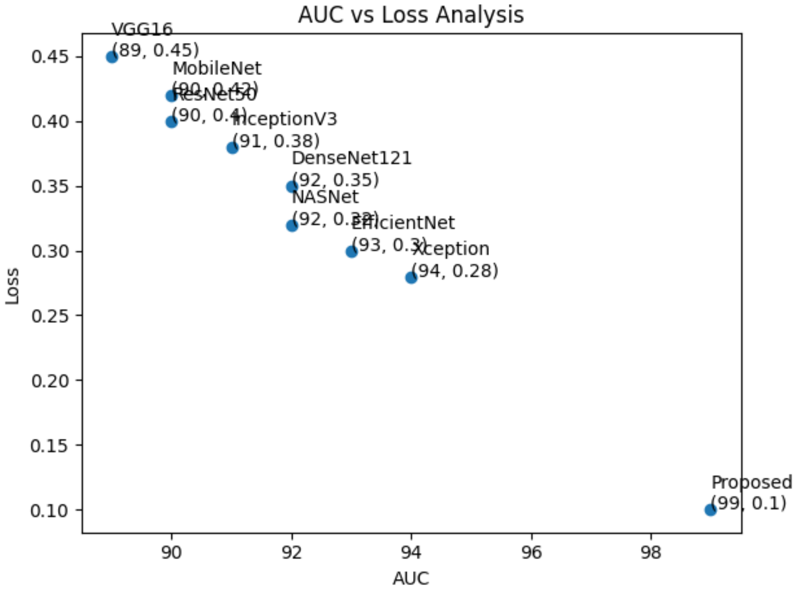

Fig 6. AUC Vs Loss with state of the art model

Fig. 7 represents the F1-score analysis that is the combination of the precision and the recall to give a balanced assessment of the performance of the model. This is particularly a critical measure in the pest detection case, where the accuracy and complete detection are required. The findings show that the models in use provide moderate F1-scores because of the trade-offs between precision and recall, but the proposed model provides the highest F1-score of 97, which confirms the capability to keep an efficient balance between the two important measures. All the models were analyzed using multi-metric analysis in terms of accuracy, precision, recall, F1-score and AUC, which are provided in Fig. 8. It is an in-depth examination that certain models can work in certain measures, but might not be uniform in every instance. The model proposed is in all aspects more robust, stable and well generalized in all measures when compared to the pest detection tasks.

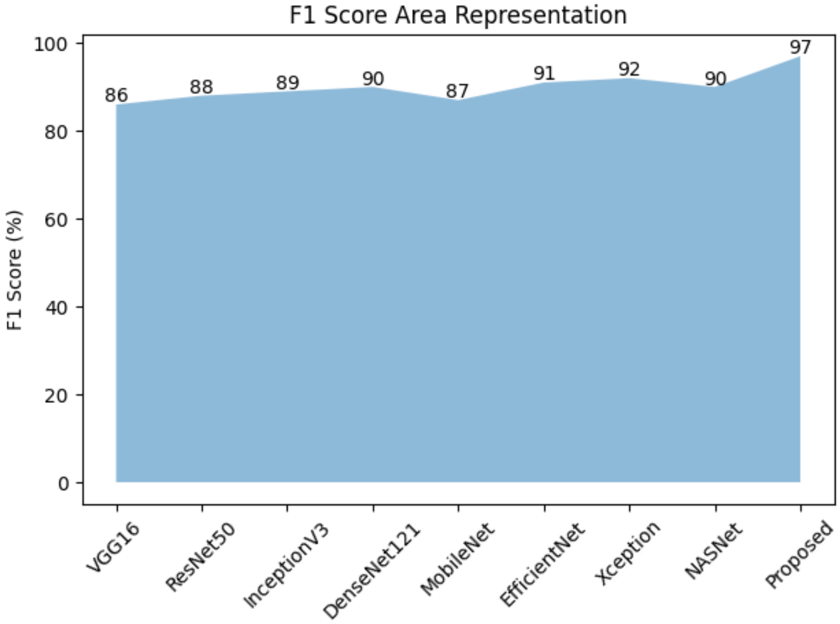

Fig 7. F1-score analysis with state of the art model

Fig. 9 shows the overall performance analysis of the proposed work that summarizes its performance compared to the existing models. The proposed model has demonstrated consistent increases in all the measures of evaluation with the least loss, which implies that there is efficiency in the learning process and the model is being correctly classified. This balanced performance with the optimized results proves the effectiveness of the given approach and its potential to be implemented in the real life conditions of the precision agriculture systems to identify the pests with the required measures of accuracy and timeliness.

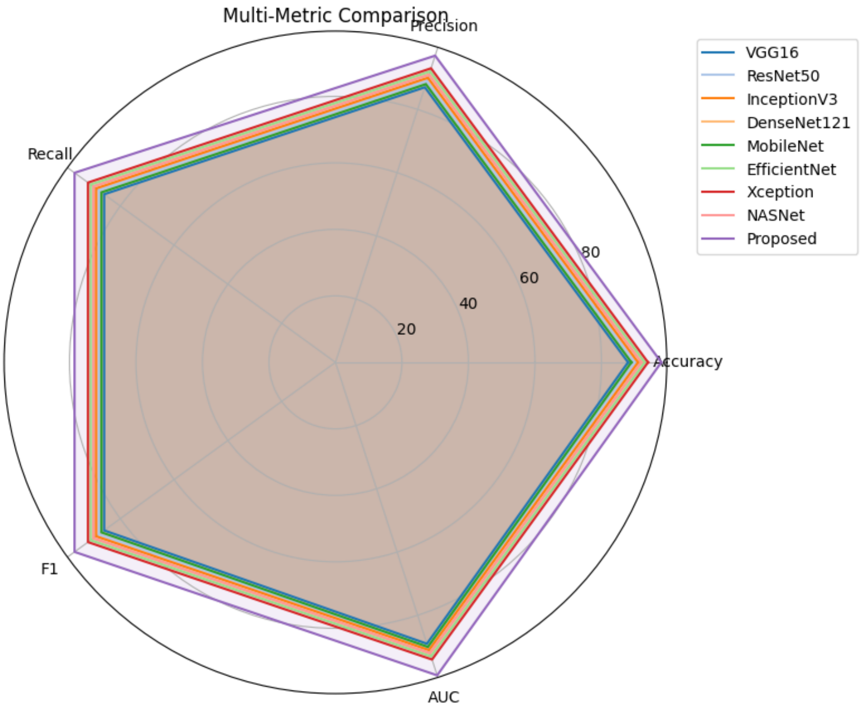

Fig 8. Multi-metric analysis with state of the art model

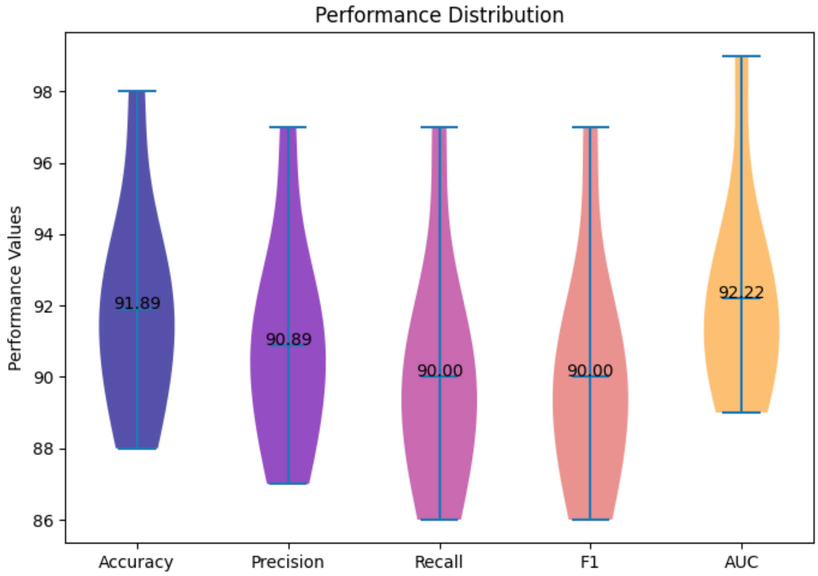

Fig 9. Overall performance analysis of the proposed work

5. Conclusion and Future Work

In this paper, a powerful and efficient system of pest detection and classification of vegetable crops is introduced basing on an efficient scheme of feature extraction and analysis. The provided work proposes SPARTA-Net (Spatial Pyramid Attention Representation Texture Aggregation), which focuses on the opportunity to learn multi-scale spatial features, fine-scale texture variations connected with the pest-infested regions on leaves. The main contribution that has been made in this work is the ability to associate spatial pyramid decomposition and attention-directed feature refinement and texture-based feature representation into one single aggregation framework that enhances the discriminative features learning. The proposed system workflow starts with the acquisition of the input image and then preprocessing is done to improve the image quality and eliminate noise. The processed image is then broken down into various levels of space in a pyramid form with texture features being extracted with the help of descriptors like LBP, GLCM and Gabor filters. Attention mechanism is put in place to emphasize significant areas with pests and the features obtained are transformed and combined into one representation. Finally, a classification module is used to make a prediction based on the preferred feature set of which pest category is predicted. The scale, brightness and background complexity changes can be handled effectively using this hierarchical pipeline. The results of the experiments prove the excellence of the proposed model in comparison to the state-of-the-art techniques. This model is accurate with a precision of 97, recall of 97 and F1-score of 97 with a small loss of 0.10. These results confirm the efficiency, robustness, and the high degree of generalization of the created approach, which can be applied to real-life farming and precision farming devices.

References

[1] M. T. Mallick, D. O. Murty, R. Pal, S. Mandal, H. N. Saha, and A. Chakrabarti, "High-speed system-on-chip-based platform for real-time crop disease and pest detection using deep learning techniques," Computers and Electrical Engineering, vol. 123, p. 110182, 2025.

[2] S. M. Venkateswara and J. Padmanabhan, "Deep learning based agricultural pest monitoring and classification," Scientific Reports, vol. 15, p. 8684, 2025.

[3] S. A. Kumar, F. Shaik, B. Yashwitha, P. Vidyavathi, A. U. Maheswari, and P. Spurthi, "Intelligence pest detection and control in agriculture using computer vision and deep learning," in 2025 International Conference on Intelligent Computing and Control Systems (ICICCS), 2025, pp. 1147-1150.

[4] [4] M. Bilal, A. A. Shah, S. Abbas, and M. A. Khan, "High-Performance Deep Learning for Instant Pest and Disease Detection in Precision Agriculture," Food Science & Nutrition, vol. 13, p. e70963, 2025.
[5] Y. S. Gill, H. Afzaal, C. Singh, G. S. Randhawa, K. Angrish, N. Jaura, et al., "Deep learning driven edge inference for pest detection in potato crops using the AgriScout robot," Computers and Electronics in Agriculture, vol. 244, p. 111492, 2026.
[6] Z. Wang, H.-W. Zhang, Y.-Q. Dai, K. Cui, H. Wang, P. W. Chee, et al., "Resource-efficient cotton network: A lightweight deep learning framework for cotton disease and pest classification," Plants, vol. 14, p. 2082, 2025.
[7] M. Zhong, L. Wei, and H. Mo, "Cotton pest and disease diagnosis via YOLOv11-based deep learning and knowledge graphs: a real-time voice-enabled edge solution," Frontiers in Plant Science, vol. 16, p. 1671755, 2025.
[8] S. Gopalakrishnan, A. Harish, S. Manoj, S. Abirami, and K. Latthisri, "Advanced CNN Architectures for Enhanced Detection of Leukemia Cells in Bone Marrow Images," in 2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), 2025, pp. 166-171.
[9] S. Wang, D. Xu, H. Liang, Y. Bai, X. Li, J. Zhou, et al., "Advances in Deep Learning Applications for Plant Disease and Pest Detection: A Review," Remote Sensing, vol. 17, p. 698, 2025.
[10] N. Deb and T. Rahman, "An efficient VGG16-based deep learning model for automated potato pest detection," Smart Agricultural Technology, vol. 12, p. 101409, 2025/12/01/ 2025.
[11] A. Bijlwan, R. Ranjan, S. Pokhariyal, A. Govind, M. Singh, K. P. Singh, et al., "Enhancing rice disease and insect-pest detection through augmented deep learning with transfer learning techniques," Smart Agricultural Technology, vol. 11, p. 100954, 2025/08/01/ 2025.
[12] R. K. Ray, S. Chakravarty, S. Dash, A. Ghosh, S. N. Mohanty, V. R. R. Chirra, et al., "Precision pest management in agriculture using Inception V3 and EfficientNet B4: A deep learning approach for crop protection," Information Processing in Agriculture, 2025.
[13] H. Lu, B. Dong, B. Zhu, S. Ma, Z. Zhang, J. Peng, et al., "A survey on deep learning-based object detection for crop monitoring: pest, yield, weed, and growth applications: H. Lu et al," The Visual Computer, vol. 41, pp. 10069-10094, 2025.
[14] R. S. Sherryl, S. J. Rosy, P. J. Jayarin, and D. B. David, "Hybrid Deep Neural Architecture for Alzheimer's Stage Prediction Using MRI and Attention-Based Temporal Modeling," in 2025 International Conference on Sustainability, Innovation & Technology (ICSIT), 2025, pp. 1-7.
[15] V. K. Isukapalli, S. Gopalakrishnan, and R. D. Kumar, "HSRG-Net: Hybrid Swin Transformer V2 based Framework with ResNet-50 Feature Extraction and GCN for Bone Marrow Cancer Detection," in 2025 2nd International Conference on Computing and Data Science (ICCDS), 2025, pp. 1-5.