Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

A Robust Deep Learning Framework: Ensemble of YOLOv8 and EfficientNet

Harjeet Kaur *

*Corresponding author, School of Computer Science and Engineering, Lovely Professional University, Phagwara, India. E-mail: harjeet.kaur@lpu.co.in

Deepak Prashar

School of Computer Science and Engineering, Lovely Professional University, Phagwara, India. E-mail: deepak.prashar@lpu.co.in

Vipul Kumar

School of Computer Science and Engineering, Lovely Professional University, Phagwara, India. E-mail: vipul.19845@lpu.co.in

Journal of Information Technology Management, 2025, Vol. 17, Special Issue, pp.32-44.
Published by the University of Tehran, College of Management
doi:%20https://doi.org/10.22059/jitm.2025.102920
Article Type: Research Paper

Received: January 17, 2025
Received in revised form: March 03, 2025
Accepted: June 13, 2025
Published online: August 01, 2025

Abstract

This research work aims to present a robust deep learning framework by devising a deep learning-based ensemble method of YOLOv8 and EfficientNet. The suggested model is evaluated on the dataset collected from Kaggle, comprising 10,000 high-definition images of stems, leaves, and cut fruits of banana and papaya. These images are captured under different lighting conditions and thus expanded to 80,000 images. Authors have proposed an ensemble model comprising YoloV8 and EfficientNet as base deep learning models to enhance prediction and classification performance. Here, authors combine the merits of both models, i.e., speed of YoloV8 and the accuracy of EfficientNet, by putting a majority voting method in place. The final forecast is determined by majority voting, and EfficientNet is given higher significance in the situation of a tie owing to its enhanced accuracy. The proposed model presents a robust solution for agricultural disease management and demonstrates significant improvements in the detection of diseases in papaya and banana, opening avenues for its widespread employment in real life.

Keywords: Deep Learning, EfficientNet, Yolov8, Image Classification, Object Detection, Loss.

Introduction

Food security and optimized practices carry the primary significance in the domain of agricultural environment, which is why minimizing harvest losses is such an important goal. In the case of tropical fruit production, where the complicated dynamics of agricultural methods are combined with problems such as climate events and illnesses, this requirement continues to hold true. Bananas, also known as Musa paradisiaca, and papayas, sometimes known as Carica, are two of the most famous crops that suffer from harvest losses among the tropical fruit species. On the other hand, the introduction of innovative and cutting-edge farming practices, which are driven by cutting-edge data-driven solutions, presents a glimmer of hope (Ramesh & Vydeki, 2020).

The farmer is responsible for manually recognizing, diagnosing, and selecting the measures that were executed. Conventional procedures are experienced, and they rely on the farmer's capacity. The potential for detection systems to reduce losses is becoming more tangible as a result of the proliferation of data-driven Industry 4.0 approaches and agricultural research, as well as instrumental tools such as computer vision models and tailored ensemble models. In spite of this, the importance of analyzing and confirming the efficacy of the many models that are now available, in addition to the creation of pipelines that are specific to each model, cannot be understated in terms of their contribution to the field of botanical pathology (Sharma et al., 2018).

The prime objective of the current work is to conduct experimental evaluation in order to assess and optimize the current state of CNN models, such as YOLOV8 and EfficientNet, in comparison to the experimental model that is tailored specifically to the data pipeline that was integrated (Vijayakumar & Vinothkanna, 2020). To be more specific, the effectiveness of these models for illness identification among products is being investigated. This includes the classification of disease severity and the evaluation of suggestive prediction methods. The purpose of this is to provide crop management methods and harvest practices that have been optimized, with the end goal of encouraging resilience and sustainability within tropical fruit-producing systems (Pooja et al., 2017).

Literature Review

The current section discusses the findings by various researchers who have yielded significant findings in the related domain. Kumar et al. (2023) proposed revolutionary advancements in various industries by employing technological advancements, particularly in the domain of deep learning (DL). This has led to the development of exemplary achievements in the domain of agriculture, particularly in disease identification, crop classification, crop prediction, and many more. Convolutional Neural Networks (CNNs) have demonstrated their efficacy for image recognition tasks (Mangal et al., 2023). Several findings by various researchers have demonstrated that DL models outdo machine learning (ML) models and human experts, as they can efficiently identify plant diseases from images, yielding excellent accuracy. Considering the efficacy of DL, the same can demonstrate promising results towards Disease identification in papaya and banana. Disease detection aims to overcome the challenges raised by numerous environmental factors and fluctuating symptoms by utilising a number of DL models to enhance the robustness and accuracy of disease detection.

In many tropical and subtropical locations, bananas and papayas are important crops that provide both nourishment and revenue (Paymode & Malode, 2022). These crops, however, are enormously susceptible to a number of diseases, which can cause severe output losses and financial harm. Effective illness treatment and prevention depend on early and precise disease identification. Conventional disease detection techniques, which depend on skilled specialists doing manual inspections, are frequently laborious and prone to human error. Autonomous illness detection systems have been made possible by developments in image processing and machine learning. To help farmers in preserving healthy crops and increasing output, Disease Detection integrates DL methods to offer a more dependable and effective solution for identifying common illnesses in papaya and banana plants.

Image categorization problems across several domains have been revolutionized by the use of CNNs and other DL models (Zhou et al., 2021). CNNs are very useful for applications like plant disease identification because of their ability to automatically and adaptively learn the spatial hierarchies of information from input pictures. Several researchers have demonstrated and validated the efficacy of CNN in various scenarios, particularly for diagnosing plant diseases. The efficacy of CNN is strengthened by achieving enhanced accuracy. Disease detection and classification can be carried out by employing a group of DL models to enhance classification accuracy. Ensemble modelling aims to combine the advantages of several base models and thus reduces the chances of misclassification or misidentification. Thus, ensemble modeling enhances the overall system performance, in line with state-of-the-art research demonstrating that ensemble approaches consistently outperform individual deep learning models across various performance metrics such as accuracy, precision, and more.

Despite the advancements in the domain of DL, its employment in plant disease identification has still taken a backseat owing to numerous factors and variables (Chakraborty et al., 2021). The ability to create an efficient and accurate plant disease detection model is convoluted by diversity in disease symptoms (Singh et al., 2023). This may be attributed to various factors, namely species, growth stage, and climatic conditions, etc. Furthermore, diagnosis may also be performed inappropriately due to look-alike visual symptoms similar across numerous diseases. To address these challenges associated with disease detection, numerous DL models are trained using different datasets to contain a diversity of images. This training of models over a dataset enables enhancement of detection generalization, thus improving the performance of models despite fluctuation of ambient factors and symptoms.

Consequently, the model becomes more adept at recognizing variations in disease symptoms, maximizing the accuracy and other significant performance metrics.

During the past few decades, there has been a significant rise in the employment of ensemble learning, which integrates numerous base DL models, achieving an outperformance over individual base models (Bose et al., 2020). This integration of various DL models garners enhancement in forecasting accuracy and resilience. Thus, ensemble learning shrinks the limitations of individual models and limits the impact of misclassification by individual models, enhancing the classification accuracy. Various researchers have demonstrated that ensemble models yield significant improvement in handling complicated classification and regression tasks. Thus, the employment of an ensemble model stands out as a promising solution for disease prediction and classification in papaya and banana plants.

The facet of agriculture may be completely revamped if DL is incorporated to its fullest potential (Shreshtha et al., 2020). DL based disease detection models enable farmers to be provided with real-time and correct information regarding crop conditions, facilitating automated and efficient disease management. Thus, it can effectively be incorporated to detect diseases in papaya and banana plants. Proactive detection of diseases enables farmers to take corrective actions in a proper timeframe, enhancing the quality and quantity of harvest. Furthermore, timely disease diagnosis reduces the need for chemical treatments, thereby optimizing agricultural costs and promoting healthier crops. This underscores the unmatched performance of deep learning in agricultural disease detection.

The effectiveness of DL models has been validated through a number of applications, which have established CNNs as an appropriate choice for farmers towards maintaining healthy crops (Mukherjee et al., 2025). Owing to the efficacy of DL models in agriculture, it has undisputedly huge potential to be an ace choice to enhance the functionalities in various domains, including agriculture (Veeraballi et al., 2020). The efficiency of DL models may be enhanced by devising complex models that have the potential to manage a variety of crops. This ability may be achieved by devising an ensemble model that integrates several base models. Usage of an ensemble model addresses the challenges of disease detection in various plants, including papaya and banana (Ramesh et al., 2018). Subsequent studies might investigate the application of sophisticated methods like federated learning and transfer learning to enhance the performance and flexibility of the models. Furthermore, the creation of intuitive user interfaces and mobile applications can aid in the broader adoption of these technologies by farmers, providing them with the resources they need to improve agricultural yield and health (Ahmed et al., 2019).

Methodology

DL models are trained using ‘N’ images {x_1,x_2,…,x_N} where each x_i contains one or more diseases of banana and papaya. Custom data is collected, consisting of approximately 1,0000 images from the agriculture fields as well as from a variety of web sources. Data augmentation increases the size of the dataset to 80,000. The dataset contains the images of stems, leaves, and cut fruits of several diseases of the banana and papaya plants. Images are taken in the shade of a plant as well as in direct sunshine using high-quality mobile phone cameras and a professional-grade DSLR camera. In this paper, an ensemble of two DL models, that is, YoloV8 and EfficientNet, has been used. The ensemble method is basically used to boost the classification and prediction performance by overcoming the limitations of one model using another. YoloV8 and EfficientNet models are explained as follows:

EfficientNet: EfficientNet is a family of convolutional neural network (CNN) architectures where the base architecture consists of a stack of convolutional layers with depth-wise separable convolutions, followed by batch normalization and swish activation functions (Behera et al., 2021). It uses efficient building blocks consisting of depth-wise convolutions followed by point-wise convolutions with expansion and squeeze-and-excitation blocks to enhance model representation capacity while keeping computational cost low. The architecture of EfficientNet is shown in Figure 1.

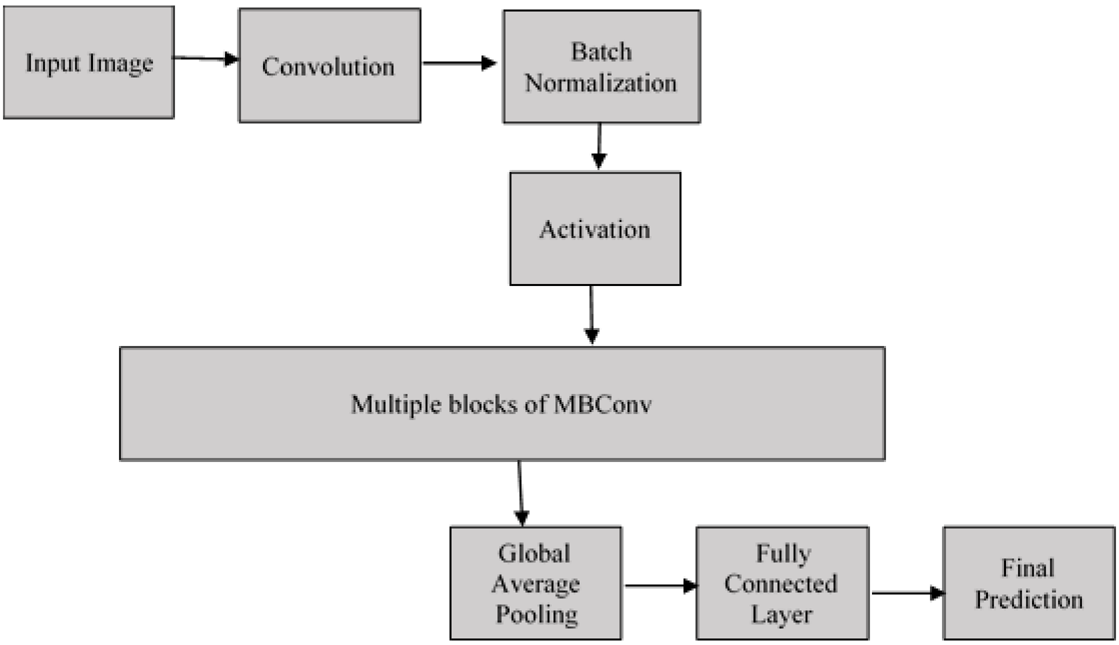


Fig. 1. Architecture of EfficientNet

YoloV8: YOLOV8 has improved architecture for object detection over previous versions, comprising modified CSPDarknet53 architecture (Sharma et al., 2023). CSPDarknet53 consists of a series of convolutional layers to extract features from the input image, which are thereafter passed to the head of YOLOV8 for predictions. It comprises multiple convolutional layers followed by fully connected layers. Its prime achievement is the inclusion of a self-attention mechanism that allows the network to focus on crucial parts of the feature maps, leading to better detection accuracy.

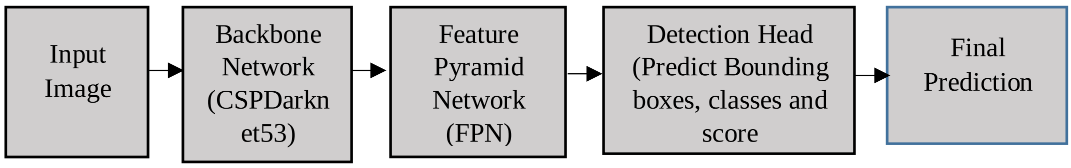


Fig. 2. Architecture of EfficientNet

Ensemble: An Ensemble of YoloV8 and EfficientNet has been implemented using a majority voting scheme (Islam et al., 2020). EfficientNet delivers high accuracy; its focus on efficiency might not translate directly to real-time object detection tasks, where speed is crucial, while YOLOV8's architecture and design choices prioritize speed and real-time performance (Nikhitha et al., 2019). Both models have been trained on the same dataset. For making the predictions, the same input is passed through both models. Both models generate bounding boxes, class labels, and confidence scores independently. Then, a majority vote among the class labels predicted by both models is conducted. Finally, the class label with the highest confidence score is assigned. In case of a tie, results of EfficientNet models are considered as they have better accuracy than YoloV8. The working of the flow diagram of the proposed model is given in Figure 3.

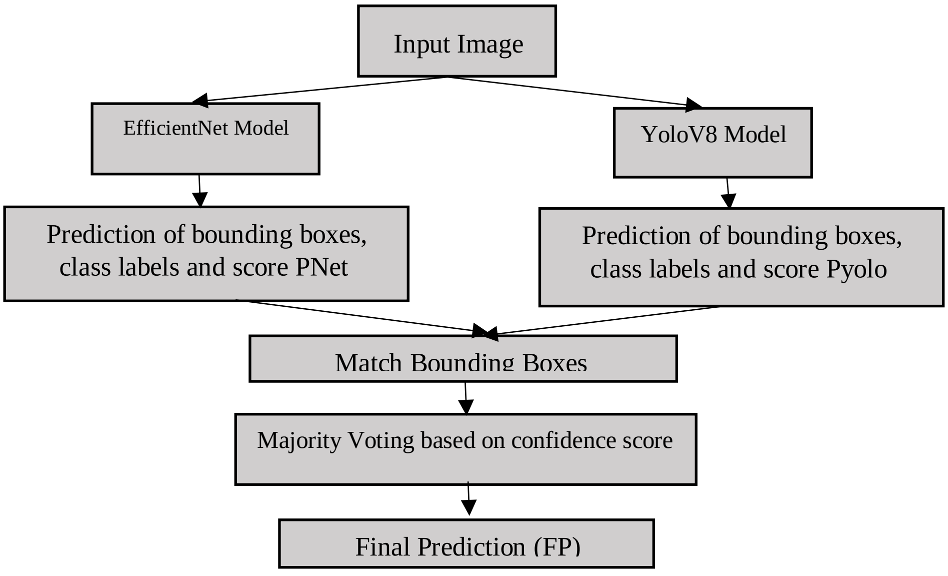


Fig. 3. Flow Diagram of Proposed System

Mathematical Formulations

First step is to predict the class probabilities for object detection:

$$\boldsymbol{Prob(I) = \{c_i, b_i, s_i : i = 1, 2, \dots, N\}} \quad (1)$$

Next is to apply EfficientNet for feature extraction, the output will be given by:

$$\boldsymbol{Out_{conv}(x) = \sigma(W * x + b).} \quad (2)$$

The final classification is given using the softmax activation function. The Diseasedet is the ensemble of YOLOv8 and EfficientNet, and the classification is given using a majority voting scheme, with EfficientNet given priority in case of a tie, as given in Equation (3)

$$\boldsymbol{Final\_Pred = argmax.(\sum_{i=1}^{n} \delta(c_i = c).(s_i + wt_i))} \quad (3)$$

The primary objective is to maximize the accuracy and reduce the loss function, which can be defined as the summation of classification loss (〖Loss〗 _class) and object detection loss 〖Loss〗 _objdect. Mathematically, it is represented in Equation (4):

$$\boldsymbol{Loss_{Total} = \alpha Loss_{class} + \beta Loss_{Objdet}} \quad (4)$$

Classification loss is defined using Binary Cross Entropy Loss as given in Equation (5):

$$\boldsymbol{Loss_{class} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log y_i + (1 - y_i)\log(1 - y_i]} \quad (5)$$

Object detection loss is calculated as smooth L1 loss as given in Equation (6):

$$\boldsymbol{Loss_{objdect} = \sum_{i=1}^{N} smooth_{l1}(b_i - \hat{b}_i)} \quad (6)$$

The symbols are defined in Table 1

Table 1. Symbols defined in the Equations

Symbol	Meaning
$\boldsymbol{\delta}$	Indicator Function
$\boldsymbol{w}$	Weight to resolve the tie in the final classification
$\boldsymbol{I}$	Input Image
$\boldsymbol{c_i}$	Class Label
$\boldsymbol{b_i}$	Bounding box coordinates
$\boldsymbol{s_i}$	Confidence score
$\boldsymbol{W}$	Weight matrix of the convolutional kernel
$\boldsymbol{x}$	Input Feature Map
$\boldsymbol{b}$	Bias term
$\boldsymbol{\sigma}$	Sigmoid Activation Function

The pseudocode of the proposed method is given below:

Input: Image (I)

Output: Classification Class (C)

1. Load YOLOv8 and Efficient Net
2. Preprocess $\boldsymbol{I}$
3. $\boldsymbol{P\_Yolo}$= Prediction from YOLOv8
   $\boldsymbol{P\_Eff}$= Prediction from EfficientNet
4. Yolo_classes, Yolo_Score= Extract($\boldsymbol{P\_Yolo}$)
   Eff_classes,Eff_scores=Extract($\boldsymbol{P\_Eff}$)
5. Score_dict={}
6. If class not in Score_dict:
   Score_dict[class]=0
   Score_dict[class]+=Yolo_scores[class]+w*Eff_scores[class]
7. C=max(score_dict, key=score_dict.get)
8. Return C

Results and Discussion

The proposed approach is validated using experimental evaluation using the dataset collected from Kaggle. As discussed, the considered dataset consisted of 10000 high-quality images, which were augmented to generate 80,000 images. The research work is carried out Tensorflow API. The proposed model uses EfficientNet and YOLOV8 as base classifiers. During experimental evaluation, a learning rate of 0.001 for EfficientNet is used with 'adam' optimizer and Swish activation function (Upadhyay et al., 2025). Similarly, YOLOV8 considers a batch size of 64 with Leaky ReLU activation function with 'adam' optimizer and learning rate of 0.001. Various simulations are done to prove the efficacy of the suggested technique.

a) Comparative Analysis of Various Models

The effectiveness of the proposed model is compared with baseline classifiers using various performance metrics, namely recall, mean average precision (mAP), and F1-score (Mangla et al., 2022). The motive behind using mAP is that it indicates the effectiveness of object detection tasks. The obtained performance metrics are illustrated in Table 2 for EfficientNet, YOLOV8, and the proposed model. Table 2 signals the supremacy of the proposed model over EfficientNet and YOLOV8 by achieving mAP and recall of 98.5% and 95.2%. It can be concluded from Table 1 that the proposed ensemble model outperforms EfficientNet and YOLOV8 by 4% and 7.2% respectively. Similarly, recall of the proposed model is higher than recall of EfficientNet and YOLOV8 by 2.7% and 5.1%. Precision vs. recall curve of EfficientNet, YOLOV8, and the proposed model is illustrated in Figure 4.

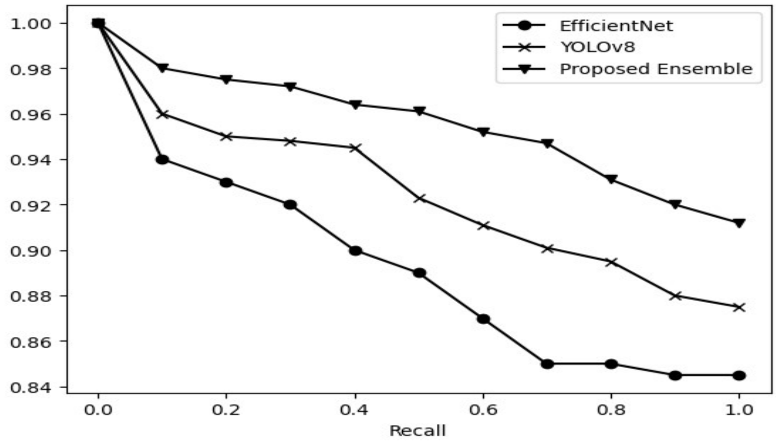


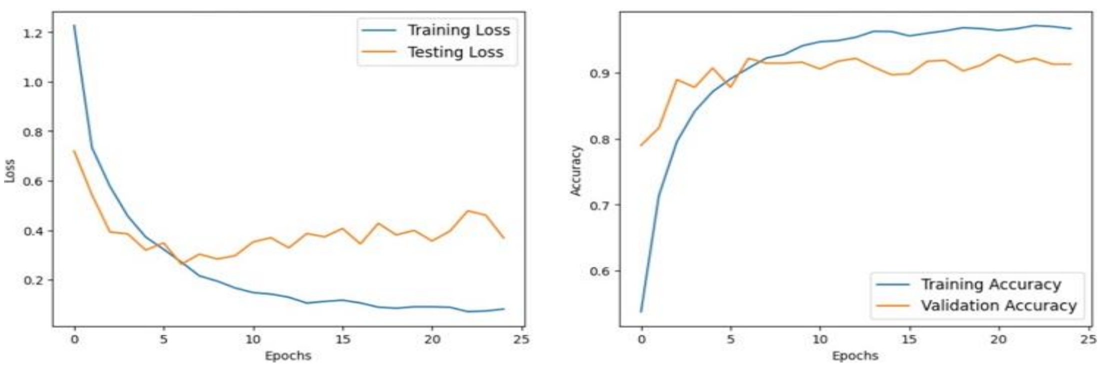
Fig. 4. Precision vs Recall Curve of EfficientNet, YOLOV8, and Proposed Model

Table 2. Comparative analysis of various Models on FLIR Thermal Dataset

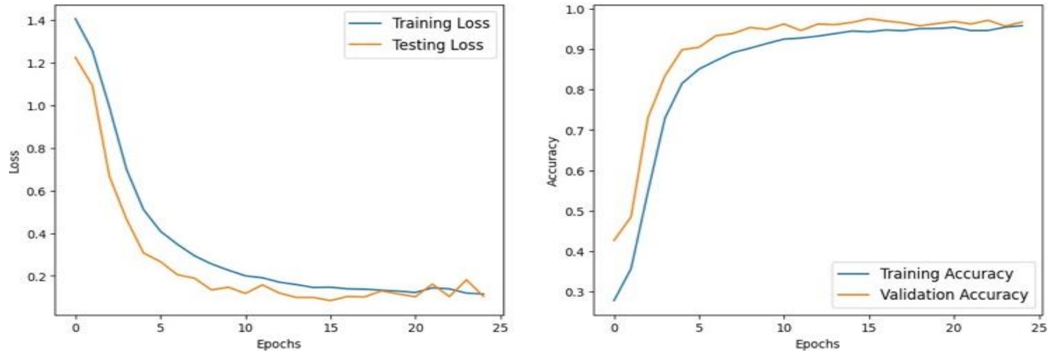
Model	Recall (%)	mAP (%)	F1-Score (%)
EfficientNet	92.5	94.5	93.5
YOLO$_{V8}$	90.1	91.3	90.7
Proposed Ensemble	95.2	98.5	96.8

b) Comparative Analysis of History Graphs

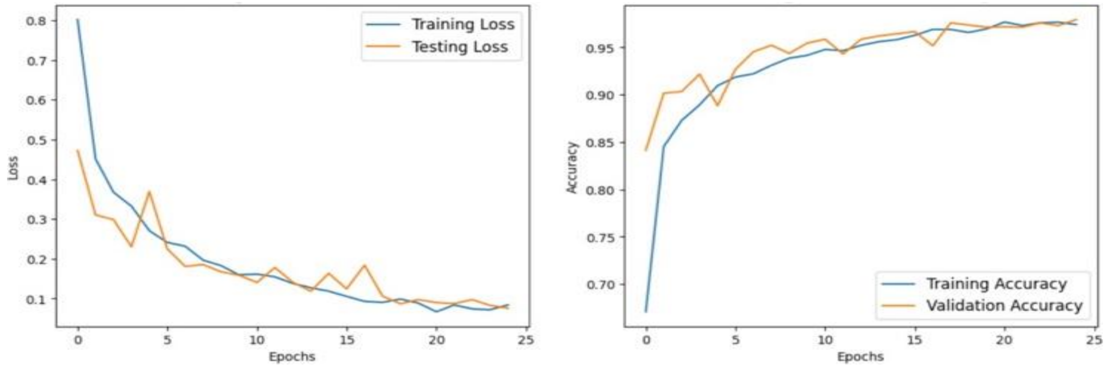
DiseaseDet uses cutting-edge DL methods to solve the crucial problem of precisely identifying illnesses in harvests of papaya and bananas. For the same, authors have used EfficientNet and YOLOV8. The motivation behind the usage of these models is their well-established effectiveness in object classification. The abilities of both models are integrated to harness their power and efficiency to achieve overall accuracy enhancement. The efficiency of the suggested model is compared and validated using historical graphs. The obtained results are illustrated in Figure 5, which clearly illustrates that the ensemble model outperforms individual models in terms of training and validation loss and accuracy.



History Graph of EfficientNet



History Graph of YOLOV8



History Graph of Proposed Model.

Fig. 5. Comparative Analysis of History Graphs of all Models

Thus, the ensemble model capitalises on the merits of both base models, namely EfficientNet and YOLOV8, for identifying diseases of bananas and papayas. This experimental analysis advocates the merging of several base DL models to achieve better performance in various applications, including agricultural.

Conclusion

Current research establishes the requirement of timely detection of disease in banana and papaya so as to facilitate farmers in taking proactive actions. This preventive approach enables a time and cost-effective solution while achieving a chemical-free and maximal harvest. For the same, authors have proposed an ensemble model that integrates multiple base models for disease identification. The proposed ensemble model integrates YOLOV8 and EfficientNet through a majority voting classifier, thus combining the merits of both models in disease identification. The proposed model is simulated on the dataset from Kaggle, which is augmented to 80,000 photos (from originally 10,000 photos). This data augmentation ensures that the model yields promising performance despite varying input. The majority voting classifier in the proposed model limits the limitations of both models while leveraging their efficiency, thus guaranteeing performance enhancement over base models. The obtained results establish DiseaseDet as a trustworthy model for disease prediction, advocating its widespread employment in real life. Current research may be extended further to include a wider range of plant diseases, not limited to only papaya and banana.

Acknowledgments

This study is a part of a PhD thesis. Acknowledgment is bestowed on honorable supervisors and examiners.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ahmed, K., Shahidi, T. R., Alam, S. M. I., & Momen, S. (2019, December). Rice leaf disease detection using machine learning techniques. In *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)* (pp. 1–5). IEEE.

Behera, S. K., Rath, A. K., & Sethy, P. K. (2021). Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *Information Processing in Agriculture, 8*(2), 244–250.

Bose, B., Priya, J., Welekar, S., & Gao, Z. (2020, December). Hemp disease detection and classification using machine learning and deep learning. In *2020 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)* (pp. 762–769). IEEE.

Chakraborty, S., Paul, S., & Rahat-uz-Zaman, M. (2021, January). Prediction of apple leaf diseases using multiclass support vector machine. In *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)* (pp. 147–151). IEEE.

Islam, M. A., Islam, M. S., Hossen, M. S., Emon, M. U., Keya, M. S., & Habib, A. (2020, November). Machine learning based image classification of papaya disease recognition. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1353–1360). IEEE.

Kumar, K. N., Lahza, H., Sreenivasa, B. R., Shawly, T. A., Alsheikhy, A. A., Arunkumar, H., & Nirmala, C. R. (2023). A novel cluster analysis-based crop dataset recommendation method in precision farming. *Computer Systems Science & Engineering, 46*(3), 3239–3260.

Mangal, A., Garg, H., & Bhatnagar, C. (2023). Assessing the performance of Co-Saliency Detection method using various deep neural networks. *Journal of Information Technology Management, 15*(Special Issue: Intelligent and Security for Communication, Computing Application (ISCCA-2022)), 23–34.

Mangla, M., Shinde, S. K., Mehta, V., Sharma, N., & Mohanty, S. N. (Eds.). (2022). *Handbook of research on machine learning: Foundations and applications*. CRC Press.

Mukherjee, R., Ghosh, A., Chakraborty, C., De, J. N., & Mishra, D. P. (2025). Rice leaf disease identification and classification using machine learning techniques: A comprehensive review. *Engineering Applications of Artificial Intelligence, 139*, 109639.

Nikhitha, M., Sri, S. R., & Maheswari, B. U. (2019, June). Fruit recognition and grade of disease detection using Inception V3 model. In *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1040–1043). IEEE.

Paymode, A. S., & Malode, V. B. (2022). Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. *Artificial Intelligence in Agriculture, 6*, 23–33.

Pooja, V., Das, R., & Kanchana, V. (2017, April). Identification of plant leaf diseases using image processing techniques. In *2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)* (pp. 130–133). IEEE.

Ramesh, S., & Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm. *Information Processing in Agriculture, 7*(2), 249–260.

Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In *2018 International Conference on Design Innovations for 3Cs: Compute, Communicate, Control (ICDI3C)* (pp. 41–45). IEEE.

Sharma, N., Mangla, M., Iqbal, M. M., & Mohanty, S. N. (2023). Deep learning framework for identification of skin lesions. *EAI Endorsed Transactions on Pervasive Health and Technology, 9*(1).

Sharma, P., Berwal, Y. P. S., & Ghai, W. (2018, November). KrishiMitr (Farmer's Friend): Using machine learning to identify diseases in plants. In *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)* (pp. 29–34). IEEE.

Shrestha, G., Das, M., & Dey, N. (2020, October). Plant disease detection using CNN. In *2020 IEEE Applied Signal Processing Conference (ASPCON)* (pp. 109–113). IEEE.

Singh, M. K., & Kumar, A. (2023). Cucumber leaf disease detection and classification using a deep convolutional neural network. *Journal of Information Technology Management, 15*(Special Issue: Intelligent and Security for Communication, Computing Application (ISCCA-2022)), 94–110.

Upadhyay, A., Chandel, N. S., Singh, K. P., Chakraborty, S. K., Nandede, B. M., Kumar, M., ... & Elbeltagi, A. (2025). Deep learning and computer vision in plant disease detection: A comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review, 58*(3), 1–64.

Veeraballi, R. K., Nagugari, M. S., Annavarapu, C. S. R., & Gownipuram, E. V. (2020). Deep learning-based approach for classification and detection of papaya leaf diseases. In *Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA 2018) held in Vellore, India, December 6–8, 2018, Volume 1* (pp. 291–302). Springer International Publishing.

Vijayakumar, T., & Vinothkanna, M. R. (2020). Mellowness detection of dragon fruit using deep learning strategy. *Journal of Innovative Image Processing, 2*(1), 35–43.

Zhou, C., Zhang, Z., Zhou, S., Xing, J., Wu, Q., & Song, J. (2021). Grape leaf spot identification under limited samples by fine grained-GAN. *IEEE Access, 9*, 100480–100489.

Bibliographic information of this paper for citing:

Kaur, Harjeet; Prashar, Deepak & Kumar, Vipul (2025). A Robust Deep Learning Framework: Ensemble of YOLOv8 and EfficientNet. *Journal of Information Technology Management,* 17 (Special Issue), 32-44. https://doi.org/10.22059/jitm.2025.102920