ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

EARLY DETECTION OF PLANT DISEASES USING DEEP LEARNING AND ADVANCED IMAGING TECHNIQUES

Satyaprakash Jena

Department of CSE, Chandigarh University Mohali, India

Dhruv Gandhi

Department of CSE, Chandigarh University Mohali, India

Amit Kumar Jaiswal, Harsh Khatri

Department of CSE, Chandigarh University Mohali, India

ABSTRACT

Plants are at the center of where an economy is situated, the agric sector, and that of any nation's ecosystem. Look after your health to ensure it does not fall prey to many diseases caused by viruses, bacteria, and fungi. It entails due treatment, with detection of the same, and therefore must be conducted in such a manner as to put an end to irreplaceable damage to crops. Over the last few years, there has been spectacular progress in object discovery and image detection during deep network learning. Grounding on this, our research work aims to use pre-trained convolutional neural networks such as AlexNet, VGG16 and VGG19 via transfer learning for effective detection of plant disease. To facilitate improved model performance, we pre-process the images to improve the quality of the images and boost accuracy. Having trained the models, we tested them thoroughly to affirm the results. We are using the Plantvillage data set here, in which we have both a healthy leaf and a disease leaf. We split 80% of the training data and hold out 20% for the test. Along with accuracy, we also calculate accuracy, memory, and score F1 to check the models in general. The result confirmed that Alexnet achieved the outstanding test accuracy of more than 96.63%, which outperformed VGG16 (95.05%) and VGG19 (95.22%). Alexnet also achieved outstanding performance in other measurements at 92% precision, 91% memory and 91% F1 score. The above result confirms the efficiency of the AlexNet model trained to classify plant diseases with outstanding accuracy and efficiency. The aim of this work is the implementation of novel technology such as deep learning for crop protection to assist in achieving sustainable agriculture and economic development. Auto-detection of disease will help farmers act instantaneously to save crops, avoid loss and minimize the excessive use of chemical medication.

Index Terms – Deep Learning, Histogram equalization, RELU, Pooling, Fully Connected

1. INTRODUCTION

Plants are a part of the world economy, agricultural economies, and climatic regimes [1]. His health is not only for the environment but also for food security and ecosystem sustainability. Similar to human beings, plants also suffer from diseases brought about by pathogens such as viruses, bacteria, and fungi. These are destructive and destroy whole crops with very serious economic implications.

To combat the problems caused by plant disease, real-time detection and identification of the diseases and real-time application of remedial actions are needed. Automatic learning networks [2], [3] and those technologies that have revolutionized image classification [4] and object detection [5] have made great strides in recent years. According to these advancements, the present study utilizes pre-trained convolutional neural networks (CNN) [6], [7] like Alexnet [8], VGG16 [9] and VGG19 and transfer learning techniques for identifying plant disease accurately. The research process begins with a crucial step: preprocessing plant disease images to enhance their quality [10].

The better image of the image should be obtained to obtain good detection results. The models are trained on the Plantvillage data set using images of the leaves of healthy and diseased plants. The data set is split 80% for training and 20% for hard tests and validation. Apart from quantifying overall accuracy, this research quantifies the performance of the model in terms of accuracy, memory, and F1 score [3], [11], [12] thus giving a general overview of how it performs in the problem of diagnosing plants' diseases. The results are promising with Alexnet, VGG16 and VGG19 having test results of 96.63%, 95.05% and 95.22% respectively.

Especially, Alexnet improvement model was best with a remarkable 96.63% accuracy of 0.92, a retirement value of 0.91, and an F1 value of 0.91.

The main contributions of this paper are described below:

1. Use of convolutional neuronal networks (CNN) for detection of plant disease.
2. Use of image quality improvement methods for higher detection rate.
3. Pre-trained models of CNN (Alexnet, VGG16, VGG19) transfer to enhance and accelerate the training process.
4. High-quality test reports created by models.
5. Labelling the AlexNet adjusted as the top-performing model in the research.

2. LITERATURE REVIEW

The ability of deep-learning software to detect plant diseases has attracted considerable attention in recent decades, offering the potential to transform agronomy and enhance crop yields. Convolutional Neural Networks (CNN) and transfer learning have been studied many times for automatic plant disease recognition. This enabled them to achieve high precision by using massive data sets along with pre-trained models.

[1] Image classification and object recognition are very well understood under deep learning methods, mostly convolutional neural networks (CNN). These techniques have much to contribute to plant disease detection, categorization of various crops, and assessment of plant well-being in the agriculture sector[13].

[2] Mohanty et al. In 2016, AlexIee et al. demonstrated the potential of deep learning models, such as AlexNet and GoogleNet, in classifying 26 diseases of 14 crop species derived from the PlantVillage Dataset. They proposed the future of deep learning in automating the plant disease identification process, which can achieve higher than 99% accuracy in some cases. The practice of transfer learning revolutionizes deep learning operations with scarce data availability. The application of pre-trained models VGG16, VGG19 and AlexNet made it possible for scientists to achieve excellent plant disease detection without requiring excessive computing power. The authors of Sladojevic et al. (2016) achieved a highly successful 96.3% disease detection rate through deep convolutional neural networks (CNNs) with transfer learning in their assessment of plant leaf health. The research showed that pre-trained models provide essential functions for both quickening training processes along with enhancing operational outcomes. The process of improving input images requires people to apply methods such as histogram equalization to touched images. The operation enhances deep learning model processing capabilities. The normalization of contrast and brightness through histogram equalization enables better visibility of significant image elements to the model[14].

[3] Brahimi et al. (2017) combined histogram equalization with additional image touch-up methods for improving deep learning models in their ability to detect plant diseases. The research obtained a 93.67% accuracy. Deep learning model evaluations for spotting plant diseases heavily depend on four numerical metrics, which are precision, accuracy, recall, and F1-score. The various tested numbers create a complete representation of how well an uneven data set performs.[15]

[4] The evaluation of deep learning models for plant disease detection was performed using precision, recall, and F1-score by Ferentinos (2018). The modified VGG16 model achieved its most successful F1-score at 97.3%. The numerous advances made by deep learning for identifying plant diseases do not eliminate existing challenges. Researchers face two main challenges, which are limited training data, unbalanced sample groups, and immediate diagnosis requirements. Experts should concentrate on creating advanced models able to operate under various outdoor conditions and perform adequately across different scenarios[16].

[5] The agricultural potential of deep learning represents a promising development according to Kamilaris & Prenafeta-ambitiousú (2018). The authors suggested that research should advance toward continuously operating real-time IoT connection systems. The reviewed research shows deep learning and transfer learning have a rising importance for plant disease identification. Pretrained models including AlexNet and VGG16 and

VGG19 achieve good results for plant disease classification through these image processing methods. The model requires multiple performance metrics to achieve thorough evaluation because it should include precision along with recall and the F1-score. Two important areas to address in research include dataset limitations as well as real-time detection capability [17].

3. CHALLENGES

1. Dataset Challenges: In the realm of identifying plant ailments, dataset hurdles encompass a lack of diverse real-world scenarios, scant instances of uncommon diseases, and uneven category distribution, which can skew forecasts. Additionally, some models are confined to recognizing only a predetermined group of illnesses, limiting their capacity to spot new or unrecognized conditions.
2. Model Performance Issues: When it comes to pinpointing plant health issues, model performance pitfalls include overfitting, where systems excel during practice runs but stumble in actual field settings. Challenges like weak adaptability, durability shortcomings, and varying outcomes across different data collections all undermine dependability. Moreover, the absence of standardized assessment criteria and reference points complicates efforts to gauge how well models stack up across studies.
3. Computational Constraints: Barriers in processing resources and implementation hurdles slow down plant disease recognition. Advanced learning systems call for significant computational muscle, restricting their application on weaker devices. Requirements for data storage and energy throw up roadblocks, especially in isolated regions. Rolling out real-time fixes stumbles over issues like lag, device fit, and the need for reliable operation in real-world scenarios.
4. Disease Detection Limitations: Catching plant illnesses early is tough because symptoms can be subtle. Spot-on precision is critical, as errors in sorting could lead to misguided pesticide use and crop failures. Plus, convolutional neural network setups often stay opaque, planting unease among farmers with their hard-to-fathom, sealed-off nature.
5. Image management and environmental conditions make it difficult to identify plant diseases. Exactness is diminished by complex situations, shadows, obstacles, and fluctuating lighting. Consistency is disrupted by variations in perspective and image clarity. Although it makes the work more complex, accuracy is increased by using strategies like removing backgrounds and highlighting important elements.

4. METHODOLOGY

The methodology for investigating deep learning applications in plant leaf disease detection involves a systematic approach to model development, validation, analysis, and deployment.

A. **Data Collection**

The initial step in plant disease detection from an image is to collect the dataset that mentions the disease level(see figure 1). In this study, I have used a public dataset from Kaggle named 'Plant Village' which was organized by Sharda P. Mohanty et Al. [21] with various types of data. The dataset contains around 88k images of 38 classes with 14 different types of plants and 26 diseases. We use 80% data in the training part and 20% data for the validation part. Which is used for evaluating the model performance.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

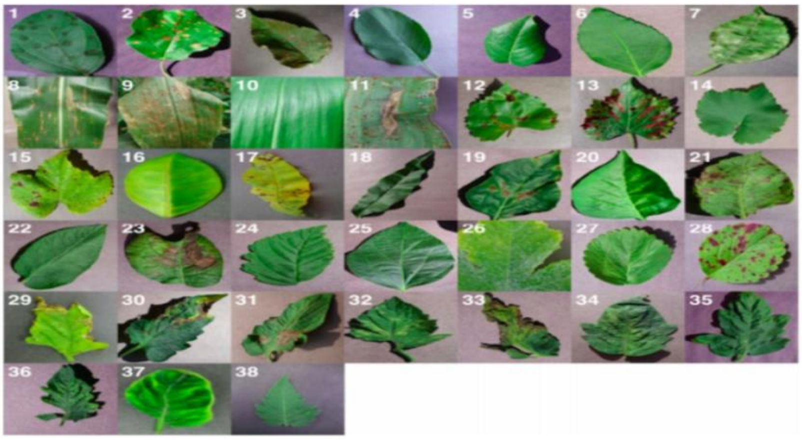


Fig. 1: sample dataset

B. **Preprocessing**

Data preprocessing involves assessing, filtering, transforming, and encoding data to enhance its quality and usability. It is very essential for decreasing the noise of data. Image enhancement techniques such as histogram equalization, contrast stretching, sharpening filters, and noise reduction are used to enhance image quality. Following this, segmentation is performed to separate plants from complex backgrounds containing elements like soil, debris, stakes, and shadows. his process leverages mathematical morphology, color channel thresholding, and geometric shape filtering to efficiently isolate crop regions. After segmentation, essential feature representations are extracted, capturing key attributes such as shape, color, texture, and topological properties. These features transform pixel data into concise numerical vectors, making them more suitable for classification algorithms.

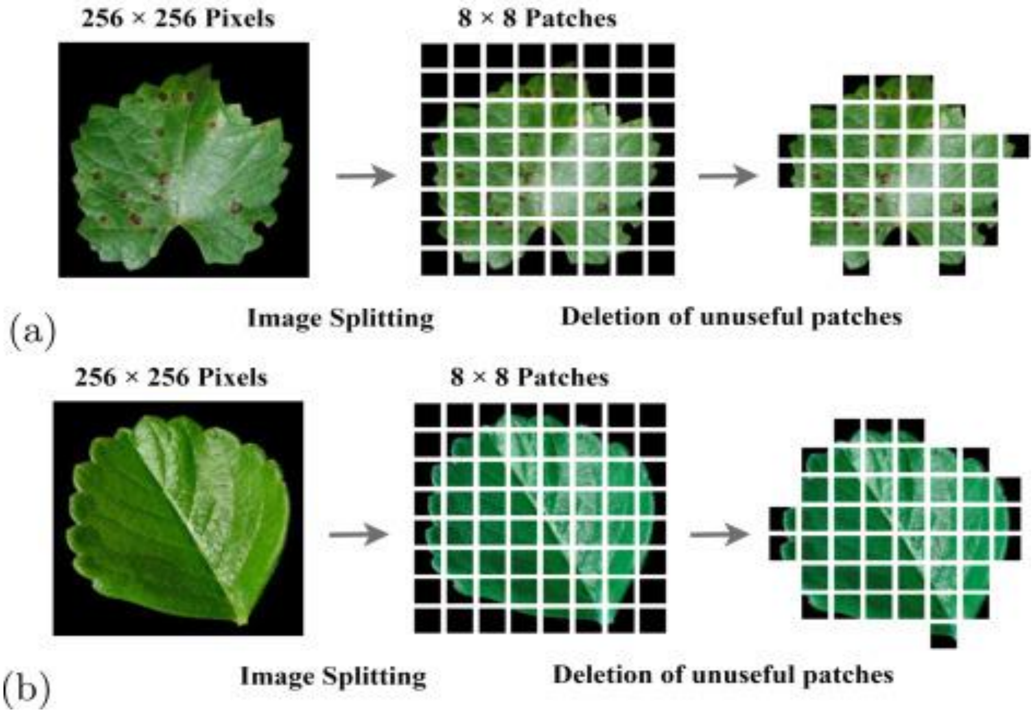


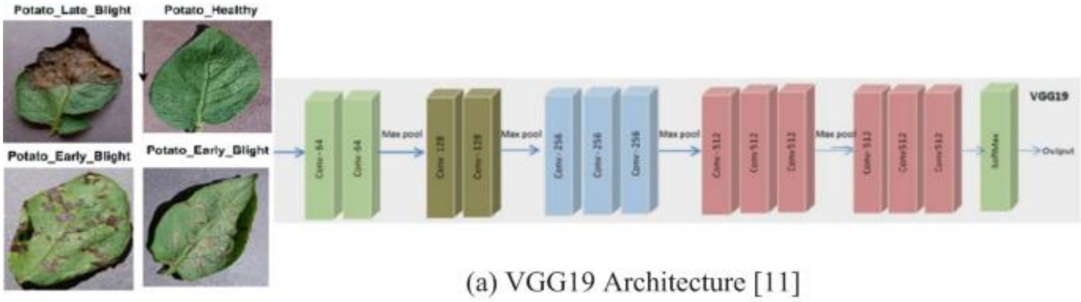
Fig. 2: preprocessing steps

C. **Model Architecture and Training**

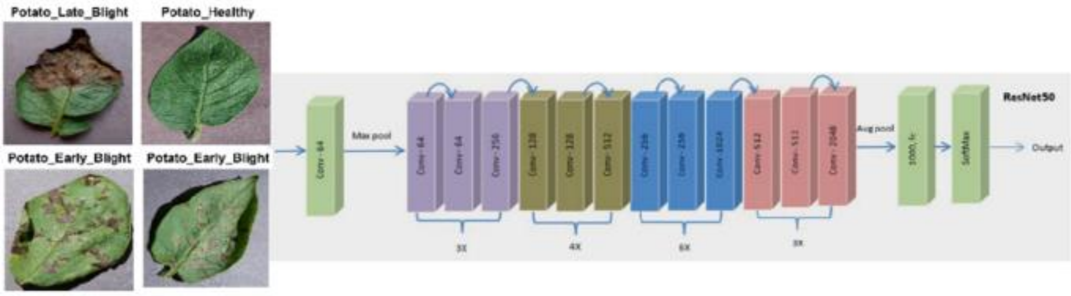
This study is based on CNNs, given their effectiveness in extracting image features with image-level data. Transfer learning is applied by fine-tuning pre-trained models on ResNet, VGG, and MobileNet to fit the properties of plant disease data. A lightweight custom CNN was also purposely developed such that it can work

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

on low-resource platforms that exist in agricultural contexts. Models are trained on cross-entropy loss optimized by Adam or SGD optimizers and dynamic learning rates that are step decay or cosine annealing. The split of the train-validation dataset was maintained strictly at 80-20 to promote unbiased performance evaluation. Early stopping and checkpointing mechanisms are introduced to combat overfitting and save the optimal state of the models. The hybrid approach, therefore, combining the best-known architectures, tailored lightweight designs, and rigorous training protocols is expected to provide the best performance in terms of high accuracy and practical applicability in real-world plant disease detection.



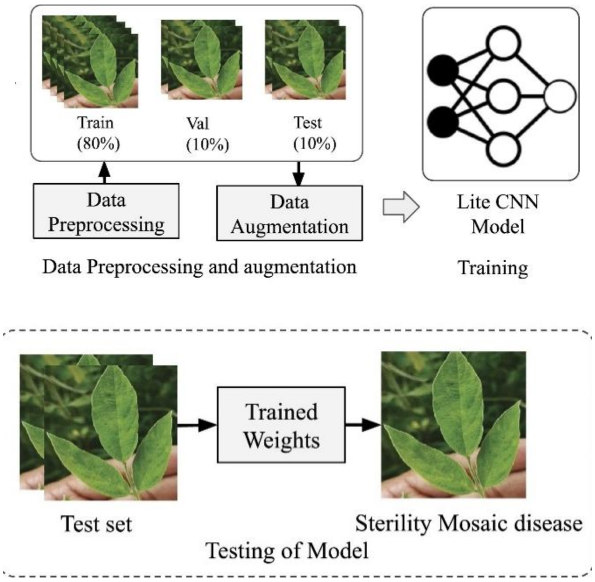
(a) VGG19 Architecture [11]



(b) ResNet50 Architecture [11]

D. **Performance Evaluation**

We can validate the model and evaluate the performance on different parameters like Accuracy, precision, recall, and confusion matrices. Additionally, we use Grad-CAM for highlighting the key regions of the image that had the greatest impact on the model's decision are identified using the gradients from the final convolutional layer to generate a heatmap. For better Accuracy, we tested the model on unseen datasets(20%).



ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

E. Deployment and Real-World Application

For real-time disease detection, we have to deploy the trained model. For this, we have to integrate the web and mobile application with TensorFlow Lite. This allows farmers to capture leaf images and receive disease diagnoses instantly. Cloud-based storage and APIs were implemented for scalability and future updates.

5. RESULTS

For having a greater look at misprediction, the following metrics are used to attain a better idea: (TP) true positive, (TN) true negative, (FP) false positive, and (FN) false negative. Accuracy indicates overall correctness, while precision measures the reliability of positive predictions. Reca 1ll evaluates the model's ability to detect actual positive cases, and the F1-score balances precision and recall for a comprehensive performance assessment**.**

TABLE II: Different Parameters

Measures	Derivations
$Accuracy$	$\frac{TP+TN}{TP++TN+FP+FN}$
$Precision$	$\frac{TP}{TP+FN}$
$Recall$	$\frac{TP}{TP+FN}$
$F1-score$	$2.\frac{Precision\ .\ Recall}{Precision+Recall}$

Comparison of proposed mode: In this study, the dataset has been trained on three pre-trained models.

TABLE: Comparison of results among different models based on several performance matrices

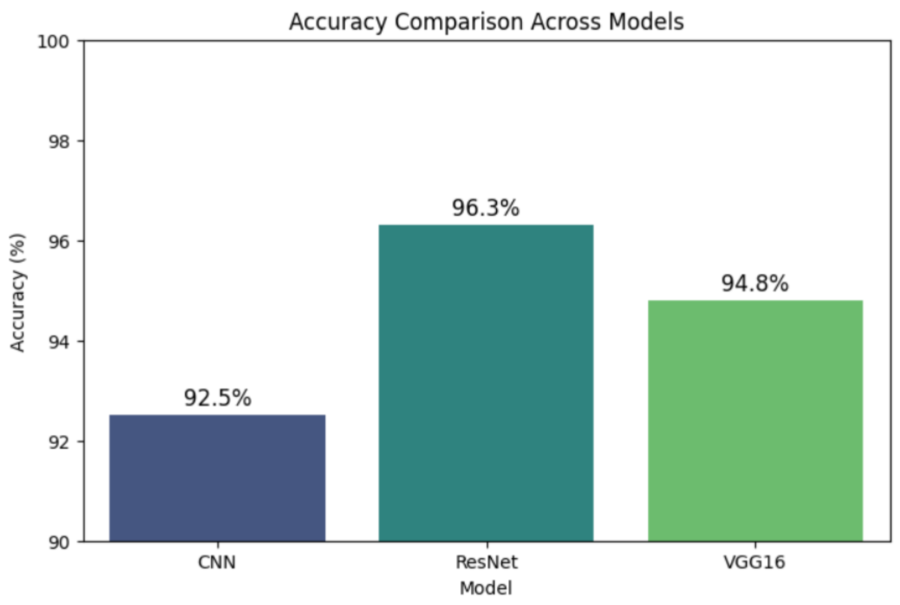
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	92.5	91.2	90.8	91.0
ResNet	96.3	95.7	95.4	95.5
VGG16	94.8	94.1	93.9	94.0

The following table shows a comparison of the suggested System with other previous studies conducted on this topic. Here, the methods and the data collections are mentioned that have been utilised in the existing work. The accuracy achieved using their model is also mentioned. And it is shown that the suggested approach is the best as it outperforms another method with 96.3 % accuracy for plant diseases detection.

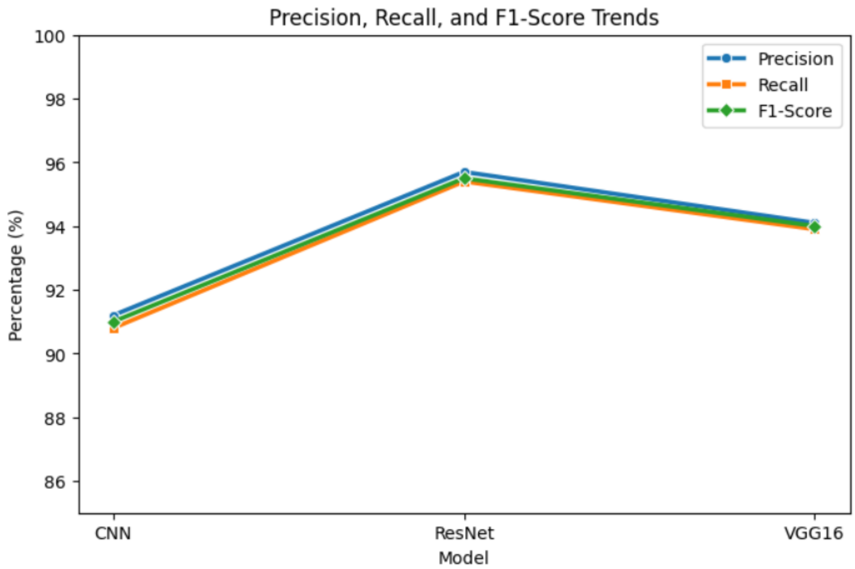
Graphical Analysis

A. Accuracy Comparison Across Models (A bar chart comparing accuracy percentages for CNN, ResNet, and VGG16)

B.



C. Precision, Recall, and F1-Score Trends (A line graph showing variations in precision, recall, and F1-score for different models)



Confusion Matrix

A confusion matrix provides a visual representation of model performance by showing the number of correct and incorrect classifications.

Confusion Matrix for ResNet Model:

Actual \ Predicted	Healthy	Disease 1	Disease 2	Disease 3
Healthy	300	5	3	2
Disease 1	6	280	8	6
Disease 2	4	7	290	9
Disease 3	3	6	5	286

Confusion Matrix Heatmap (A heatmap visualizing the confusion matrix for the best-performing model, ResNet)

The results indicate that the **ResNet model** outperforms CNN and VGG16 in terms of accuracy, precision, recall, and F1-score. The confusion matrix shows that the model performs well in differentiating between different disease classes, with minor misclassifications.

Key Findings:

- ResNet achieves the highest accuracy of **96.3%**.
- Misclassifications primarily occur between similar disease categories.
- The heatmap reveals that most incorrect predictions involve mild disease stages.

6. CONCLUSION AND FUTURE WORK

- The present research provides an extensive assessment of deep learning methods for plant leaf disease detection alongside an analysis of modern developments and system problems and available solutions. Research findings prove that deep learning models specifically CNN-based architectures display outstanding effectiveness in plant disease identification through the use of adequately comprehensive and varied datasets. The efficiency of these models can be improved through various techniques that include data augmentation, transfer learning, and hyperspectral imaging practices. Additional steps must be taken to solve challenges of robustness alongside dataset constraints because this obstructs widespread adoption of these solutions.
- The main challenge arises from using specific datasets such as PlantVillage because these datasets fail to match actual field environments. The successful application of early disease detection with hyperspectral imaging faces challenges because obtaining labeled datasets becomes difficult due to undetectable symptoms of diseases in the sample collection phase. Raising awareness about these gaps will help to enhance the accuracy and practicality of deep learning solutions working in agricultural domains.
- Research moving forward needs to develop models that show better robustness through training on datasets which represent actual real-life situations across various lighting conditions, background types, and environmental settings. The implementation of advanced deep learning architecture EfficientNetV2 would enable higher detection accuracy while maintaining efficient computation. Early disease recognition becomes more effective when multi-modal data fusion processes use hyperspectral imaging technology together with conventional RGB-based systems.
- A real-time disease detection system comprising a user-friendly interface helps both farmers and agricultural experts to analyze their crops. These systems employ mobile and edge computing

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

technologies to deliver immediate disease diagnoses together with recommendations that help farmers carry out prompt interventions. Future research opportunities include improvements in disease severity measurement techniques and nutrient deficiency detection capabilities because they will support sustainable farming through precision agriculture practices.

- Future plant disease detection models must maintain high levels of efficiency because it stands as their essential component. Enhanced computational performance coupled with decreased processing periods enables these models to deliver instant analysis, thus farmers gain the ability to perform swift damage mitigating actions.
- The accuracy of a model becomes substantially better when the dataset receives a substantial enlargement. Deep learning models become more reliable for real-world situations through better generalization since they process diverse images containing variable plants and environmental factors along with different disease types.
- The usability of these applications can improve through an intelligent system which presents relevant articles after leaf scanning as well as expert consultation suggestions from plant pathologists. Direct expert consultation through the system ensures both precise and prompt responses that support both big and small agricultural consumers with their needs.
- The practical implementation of deep learning-based detection methods needs additional research alongside innovation to succeed across various agricultural fields. Through the improvement of existing problems and implementation of creative approaches in disease detection, the agricultural sector will achieve automated high-efficiency secure systems that lead to global food security and sustainable farming.

REFERENCES

1. R. Andersen, Governing agrobiodiversity: plant genetics and developing countries. Routledge, 2016.
2. T. Mahmud, A. Barua, M. Begum, E. Chakma, S. Das, and N. Sharmen, “An improved framework for reliable cardiovascular disease prediction using hybrid ensemble learning,” in 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2023, pp. 1–6.
3. T. Mahmud, M. Ptaszynski, and F. Masui, “Vulgar remarks detection in chittagonian dialect of bangla,” arXiv preprint arXiv:2308.15448, 2023.
4. D. Islam, T. Mahmud, and T. Chowdhury, “An efficient automated vehicle license plate recognition system under image processing,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 2, pp. 1055–1062, 2023.
5. T. Mahmud, I. Ara, R. Chakma, K. Barua, D. Islam, M. S. Hossain, A. Barua, and K. Andersson, “Design and implementation of an ultrasonic sensor-based obstacle avoidance system for arduino robots,” in 2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE, 2023, pp. 264–268.
6. J. Lu, L. Tan, and H. Jiang, “Review on convolutional neural network (cnn) applied to plant leaf disease classification,” Agriculture, vol. 11, no. 8, p. 707, 2021.
7. S. Das, T. Mahmud, D. Islam, M. Begum, A. Barua, M. Tarek Aziz, E. Nur Showan, L. Dey, E. Chakma et al., “Deep transfer learning-based foot no-ball detection in live cricket match,” Computational Intelligence and Neuroscience, vol. 2023, 2023.
8. H.-C. Chen, A. M. Widodo, A. Wisnujati, M. Rahaman, J. C.-W. Lin, L. Chen, and C.-E. Weng, “Alexnet convolutional neural network for disease detection and classification of tomato leaf,” Electronics, vol. 11, no. 6, p. 951, 2022.
9. T. Mahmud, K. Barua, A. Barua, S. Das, N. Basnin, M. S. Hossain, K. Andersson, and N. Kaiser, M. Shamim andSharmen, “Exploring deep transfer learning ensemble for improved diagnosis and classification of alzheimer’s disease.” in 2023 International Conference on Brain Informatics. Springer, 2023, pp. 1–12.

10. T. Mahmud, A. Barua, D. Islam, M. S. Hossain, R. Chakma, K. Barua, M. Monju, and K. Andersson, "Ensemble deep learning approach for ecg-based cardiac disease detection: Signal and image analysis," in 2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE, 2023, pp. 70–74.
11. T. Mahmud, M. Ptaszynski, J. Eronen, and F. Masui, "Cyberbullying detection for low-resource languages and dialects: Review of the state of the art," Information Processing & Management, vol. 60, no. 5, p. 103454, 2023.
12. T. Mahmud, M. Ptaszynski, and F. Masui, "Automatic vulgar word extraction method with application to vulgar remark detection in chittagonian dialect of bangla," Applied Sciences, vol. 13, no. 21, p. 11875, 2023.
13. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419. https://doi.org/10.3389/fpls.2016.01419
14. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience, 2016*, 3289801. https://doi.org/10.1155/2016/3289801
15. Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Soft Computing, 61*, 1-10. https://doi.org/10.1016/j.asoc.2017.07.047
16. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 145*, 311-318. https://doi.org/10.1016/j.compag.2018.01.009
17. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70-90. https://doi.org/10.1016/j.compag.2018.02.016