The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

From Color Histograms to Deep Convolutional Models: A Review of Potato Leaf Disease Classification Techniques and Limitations

Syed Ishak Sha Ali[1], Sree Raja Mounika[2], A.S. Venkata Praneel[3]
[1,2,3]Department of Computer Science and Engineering,
GST, GITAM (Deemed to be University), Visakhapatnam, AP, India.
syedishakshaali4588@gmail.com[1], sreerajamounika013@gmail.com[2],
pankired@gitam.edu[3]
Corresponding Author *: pankired@gitam.edu

Abstract

Early blight, late blight, and other potato leaf diseases may threaten productivity and world food safety. Treatment of illness is effective if it is based on accurate classification and early detection. Deep learning techniques are contrasted in this paper with traditional algorithms for the classification of potato leaf diseases. Legacy techniques, such as manually designed features such as color histograms and Gray Level Co-occurrence Matrix (GLCM), and legacy machine learning classifiers have so far provided useful insights. Still, they are suffering from the drawbacks of robustness, scalability, and the ability to adapt to real-world scenarios. Whereas newer deep models like Custom Convolutional Neural Networks (CNNs), VGG16, and AlexNet have been shown to achieve significantly superior performance in recognizing disease patterns from raw imaging data. These models obviate the need for human feature engineering by learning hierarchical representations to generalize over different illumination, textures, and backgrounds. The popularly used PlantVillage dataset offers a strong baseline against which to compare and measure accuracy, model complexity, and models' generalizability. Earlier research has shown that light-weight and computationally humane models like Custom CNNs and AlexNet are equally good, if not superior, in terms of classification accuracy as deep architectures, provided they are properly fine-tuned and aided with transfer learning. This article discusses avenues through which plant disease detection systems can be further enhanced and enumerates key significant research limitations, such as restricted access to diverse, real-world data and more adaptive deployment strategies. Future research would have to focus on developing robust models capable of functioning in field environments, working with effective training pipelines, and integrating explain ability tools so that realistic decision making is supported for precision agriculture.

Keywords:
Potato Lead Disease, Deep Learning, Convolutional Neural Network, PlantVillage Dataset, Image Classification, Precision Agriculture.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

1. Introduction

Potato is the world's most important food crop for global food security and rural economies. Its high yield potential and ability to excel across a wide range of agro-climatic environments make it the cornerstone of production farming systems in most parts of the world. Nevertheless, potato cultivation is exceptionally sensitive to various foliar diseases like early blight and late blight, which can have drastic effects on quantity and quality. In Figs. 1 & 2, we have given some images of the Late blight-affected and Early blight–affected potato leaf samples. Such diseases are an ongoing menace to farmers and result in yearly heavy financial losses. Precise and rapid diagnosis of such disorders is essential for establishing effective control and management mechanisms. Conventional visual inspection techniques, although a common practice, are still subjective, time-consuming, and unsuitable for large-scale agricultural applications. Computer vision and artificial intelligence have evolved special approaches for computer-aided disease diagnosis. Of these, deep learning methods, namely Convolutional Neural Networks (CNNs), have shown remarkable effectiveness in leaf disease detection by learning complex visual patterns from raw image information directly [3]. This paper provides an exhaustive overview of deep learning-based strategies towards potato leaf disease classification based on the utilization of models like AlexNet, VGG16, and some custom CNN architectures [4,5]. The review measures these models based on their performance, training, the data used, and how they can be used in various deployment scenarios. By analyzing existing studies, the research aims to present typical trends, determine the current challenges, and outline potential research topics for the development of disease detection technology for precision agriculture.

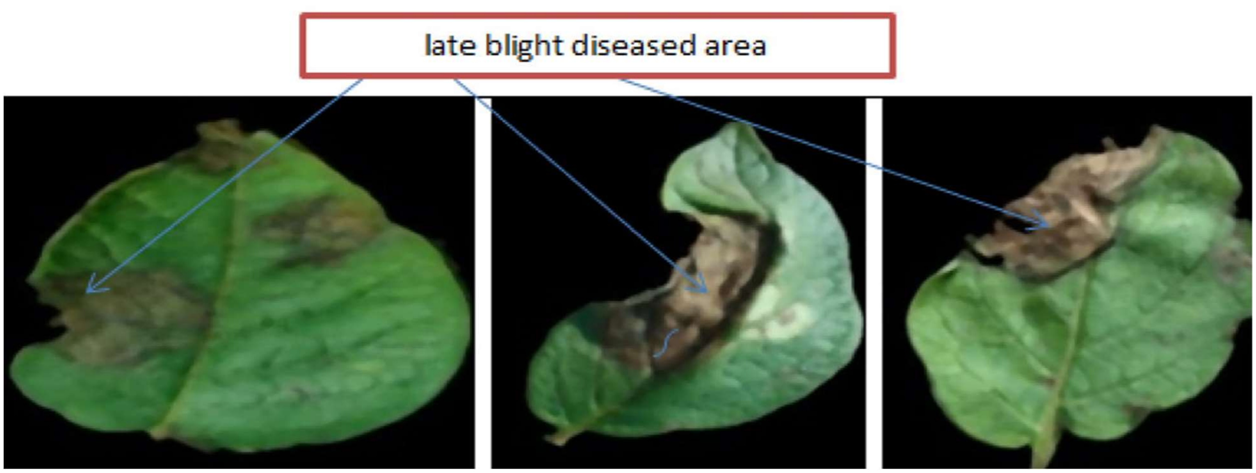

Fig 1: Late blight-affected potato leaf samples (Image source [3]).

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

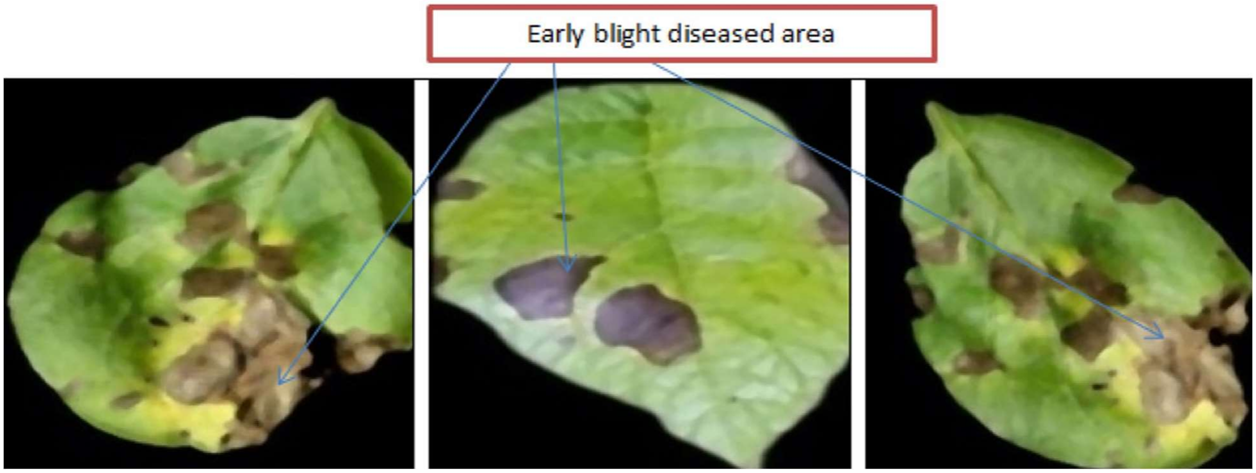

Fig 2: Early blight-affected potato leaf samples (Image Source [3]).

2. Traditional Approaches

The conventional potato leaf disease diagnosis technique to date included manual visual inspection, hand-drawn feature extraction, and conventional machine learning classifiers. Visual inspection includes detection of any visual disease symptom, e.g., color deviation, lesion pattern formation, or variation in surface texture, visually. While simple to use, these are highly variable and subjective, especially when used over large fields or in different environments. To achieve better consistency, researchers made use of feature-based approaches using the assistance of image processing algorithms like thresholding algorithms including Gray Level Co-occurrence Matrix (GLCM) [2] to extract texture features like contrast and homogeneity regions and color histograms to determine the intensity value distribution of the colors of the leaf image. These features were then applied to machine learning classifiers like Support Vector Machines (SVMs), highly appropriate for high-dimensional classification, and Random Forests averaging many decision trees to provide better results. Although very promising initially, such rudimentary approaches have inherent limitations. These require slow and tedious feature engineering, are highly reliant on subject matter expertise, and are generally unstable under changing real-world conditions like illumination, background texture, and image noise. As such, they are being gradually replaced by deep learning approaches, i.e., Convolutional Neural Networks [3,4], which are capable of learning advanced features from raw image data without the requirement of expert knowledge.

2.1. Manual Inspection:

In conventional farming, potato leaf disease diagnosis was largely dependent on the human eye. The test relies on the observation of the leaf characteristics such as discoloration, lesions, and defoliation patterns to ascertain the occurrence and type of disease. Although inexpensive and simple, the technique is open to human error and the skill and reliability of the observer. While sufficient for small-scale farming, eye diagnosis is cumbersome and unreliable when done outside in an open field setting. Accuracy is degraded by environmental heterogeneity, observer fatigue, and irregular illumination conditions. These constraints restrict the use of it for early detection and discourage its use in precision agriculture, where diagnosis has to be quick and scalable [3].

2.2. Feature Extraction Using GLCM and Color Histograms:

In order to justify automated detection, initial work in the detection of leaf diseases was based on manually developed feature extraction methods. One of the most dominant techniques used is the Gray Level Co-occurrence Matrix (GLCM) [2], which extracts texture-based features like contrast, correlation, energy, and homogeneity from gray-level images. Texture features assist in separating healthy from infected regions of the leaf by detecting surface irregularities. In addition to these, color histograms were used to quantify the intensity distribution for colors on the leaf surface to assist in detecting color transformation-based disease. Although these features, manually extracted, work well in controlled environments, their performance is typically compromised with their use in field-extracted images [3]. Variation in illumination, scaling, and intricate background changes decrease their robustness and restrict their use in real-world scenarios, calling for more adaptive learning-based methods.

2.3. Classical Machine Learning Classifiers:

After feature extraction, common classification issues in potato leaf disease detection were typically addressed by utilizing traditional machine learning approaches. Among all these, Support Vector Machines (SVMs) gained popularity because they are able to deal with high-dimensional spaces of features and achieve good classification performance [2]. SVMs operate by finding the best hyperplane that maximizes the margin between many classes; hence, they are appropriate for handling many classes of disorders. Similarly, Random Forest classifiers that combine predictions from many decision trees have been used due to their strength and ability to manage noisy data. These algorithms performed satisfactorily when executed under controlled conditions. However, because of the widespread use of hand-designed features and vulnerability to specific dataset attributes, they were not extremely efficient in advanced real-world farming settings [3]. These traditional methods are thus increasingly being replaced by deep learning algorithms that can learn and generalize over raw images without intermediaries [4].

3. Deep Learning Approaches

Over the past decade, deep learning has made the detection of potato leaf disease accessible to the creation autonomous, scalable, and precise classifier systems. Deep models such as Convolutional Neural Networks (CNNs) [1], in contrast to hand-designed feature-based traditional systems, can learn high-level, hierarchical representations of raw image data with minimal or no human involvement. This enables models to generalize more across varied image conditions such as lighting, size, and background noise. Deep learning methods also limit domain-specific feature engineering and are therefore more generalizable and versatile to various types of farm photos. They have outperformed the existing methods in potato leaf disease classification with high accuracy and robustness on public and customized data sets. The extensive utilization of models such as VGG16 and AlexNet [2,3] and the development of Custom CNN architectures has also spurred their use in agriculture applications. Such models, in some instances, benefited from tools such as transfer learning and data augmentation, provide a suitable basis for scalable disease detection. This paper compares most commonly employed deep structures to classify potato leaf disease by structure performance, and flexibility across deployment environments.

3.1. Convolutional Neural Networks (Custom CNN):

Convolutional Neural Networks (CNNs) are the go-to for the majority of deep learning-based plant disease classification problems. Compared to their counterparts using traditional handcrafted feature-based approaches, CNNs learn automatically hierarchical feature representations of raw pixel images by composing convolutional, pooling, and fully connected layers stacked on top of each other. For potato leaf disease detection, lightweight, domain-specific Custom CNN structures have been proposed for image sets like PlantVillage [1,2] or in-field images. They consist of two or four convolutional layers and one or two fully connected layers with high accuracy and low computational cost. Their efficiency and simplicity make them extremely suitable for mobile and real-time farm deployment. Despite being highly efficient Custom CNNs, their performance is affected by variability of images. They fail to generalize to field images with cluttered background, non-uniform lighting, or occlusion unless using adequate data augmentation and regularization [5].

3.2. VGGNet:

VGGNet or VGG16 is yet another prevalent architecture employed for classifying plant diseases [1,6]. Its basic and regular structure depends upon stacked repetition of 3×3 convolution layers with max-pooling and dense fully connected layers. VGG16 proved to be effective in extracting abstract spatial information from leaf images and is one of the best choices for classification. VGG16 is widely used in transfer learning tasks, with the network pre-trained with weights on huge datasets like ImageNet and fine-tuned on agri-datasets like potato leaf diseases. VGG16 can potentially give good performance and generalization. But with due because of its much higher number of parameters and longer training duration, it may not be the ideal proper solution for real-time systems or for specific applications like edge devices in agriculture.

3.3. AlexNet:

AlexNet is among the earliest deep networks used in plant disease diagnosis and is still suitable for potato leaf classification tasks. Its very shallow design compared to more modern networks is made up of five convolutional and three fully connected layers. Its reduced processing demands and quicker training time mean it is an affordable option for less funded researchers or smaller data sets. Comparisons show AlexNet to have competitive performance when used on benchmark datasets such as PlantVillage [2,6] and has been compared with comparative testing against more deeper networks such as VGG16 [10]. While less able to describe more complex spatial relationships than deeper networks, AlexNet is still a good place to begin and well-suited to rapid prototyping, mobile deployment, and to applications where computational efficiency is an overriding concern.

4. Dataset Used

In the domain of deep learning-based potato leaf disease classification, dataset quality and selection play a critical role in determining model performance as well as generalizability. In the literature reviewed, two broad categories of datasets have been seen: controlled/laboratory datasets like PlantVillage [2].

PlantVillage is the most popularly utilized source for model training as well as testing of deep learning models. It has more than 54,000 images across 14 crop types and 26 diseases, taken in optimal conditions with standard backgrounds, constant lighting, and low noise. As shown in previous papers, such studies have heavily depended on this dataset due to its availability and clean labels. Such models as AlexNet, VGG16, and some custom CNNs have performed incredibly high levels of accuracy (e.g., ~99%) on this dataset [6]. However, its laboratory nature brings about a fundamental weakness: the inability to provide strong robustness when applied in the real world.

To overcome this, some studies have moved towards employing field-captured datasets, which present more realistic and challenging cases. Some of them utilized smartphone-captured images of potato leaves taken straight from farms in Ethiopia. Their dataset contained varying backgrounds, different lighting, and disease severity annotations. Also, a few of them used a field dataset to train tailored CNN models with enhanced model robustness and real-world performance. The datasets mimic real-world deployment conditions, where shadowing, occlusion, and non-uniform leaf orientation cannot be avoided.

Earlier research applied controlled and field images with both VGG16 being used to distinguish between early and late blight in the leaves of potatoes. By combining Grad-CAM with a fine-tuned VGG16 model, transparency in making decisions and verification of focus of the model in disease areas within noisy backgrounds were highlighted.

In conclusion, though PlantVillage continues to be a robust baseline for model performance [2]. The transition towards field-collected data reflects an emerging demand for tougher, real-world datasets. Future work should aim for dataset diversity and realism to facilitate effective deployment of deep learning systems in agriculture.

5. Performance Evaluation

In the comparison of models for potato leaf disease classification, we have explored how deep architectures such as Custom CNN, VGG16, and AlexNet perform. More straightforward in design, Custom CNN achieved a validation accuracy of 99.1% in some experiments and was a computationally feasible and suitable solution for real-time and low-resource deployment. VGG16, being more networked and capable of extracting robust features, proved to have consistently high accuracy after fine-tuning using transfer learning, particularly for datasets such as PlantVillage. VGG16, being memory-hungry and with longer training time, has its networked nature and performance on controlled data and field data, which earns it a place as a go-to for complex cases. AlexNet, even though it was among the first deep models, showed competitive performance in classification, especially in clean data. Its shallow architecture and low-cost training time make it deployable on edge and mobile platforms. These results indicate that very optimized light models like Custom CNN and AlexNet, or optimized models like VGG16, are

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

capable of achieving best-in-class performance with no expense of very computationally expensive networks. In general, the results indicate that better trained and less complex models are as good or even better than complex models in the case of potato leaf disease classification.

Table 1: Performance Comparison of Custom CNN, AlexNet, and VGG16 for Potato Leaf Disease Classification

Metric	Custom CNN	AlexNet	VGG16
Accuracy	96%	87%	97%
Macro Precision	0.97	0.60	0.95
Macro Recall	0.92	0.62	0.91
Macro F1-Score	0.94	0.61	0.93
Weighted F1-Score	0.96	0.84	0.96
Healthy Class Recall	0.84	0.00	0.78

The performance comparison table shows the merits and demerits of every model on various assessment parameters. VGG16 and Custom CNN show excellent and good-balanced performance on accuracy, macro precision, and weighted F1-score, but AlexNet performs extremely badly in identifying the Healthy class, as revealed by its zero recall as well as lower macro-averaged scores. These results are a perfect justification for choosing models not only based on overall accuracy but also based on their class-wise reliability and sensitivity towards imbalanced data sets.

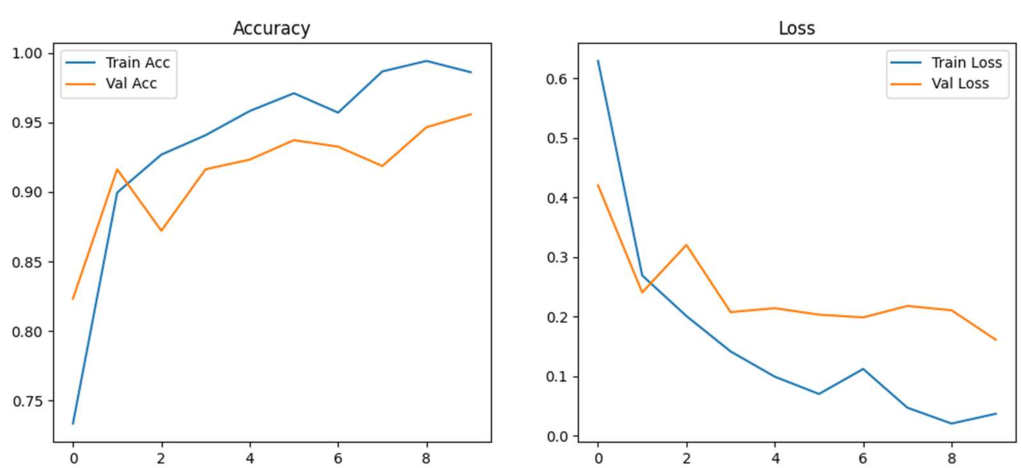

Fig 3: CNN Model Accuracy and Loss

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

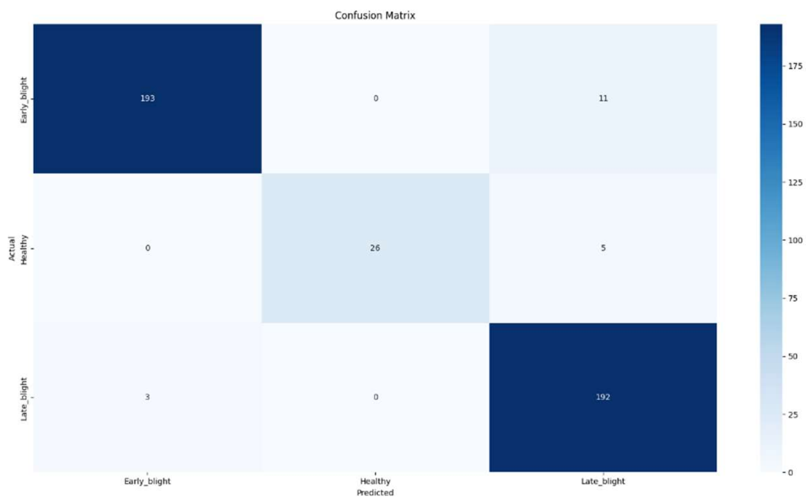

Fig 4: CNN Model Confusion Matrix

Fig 3 & 4 Show the accuracy and loss graphs of Custom CNN model that achieved high and stable performance with minimal overfitting also the confusion matrix shows strong classification of Early and late blight, with slight misclassification in the Healthy class.

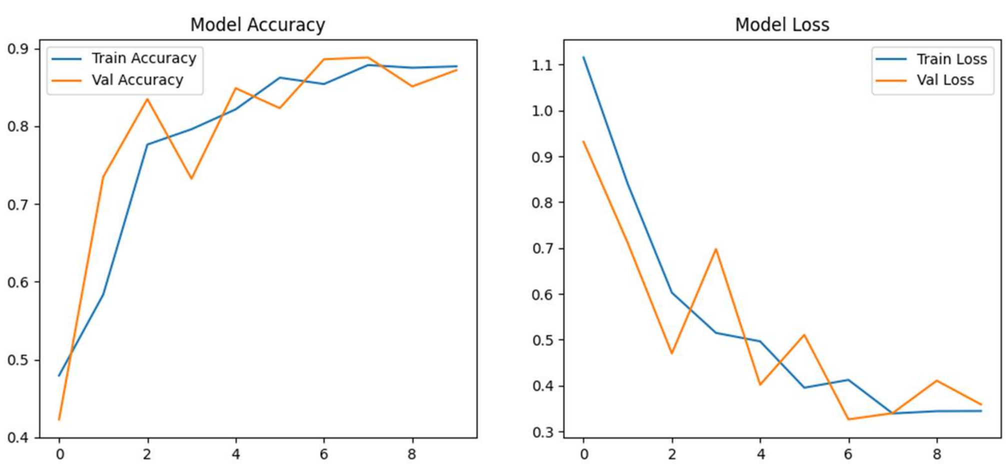

Fig 5: AlexNet Accuracy and Loss

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

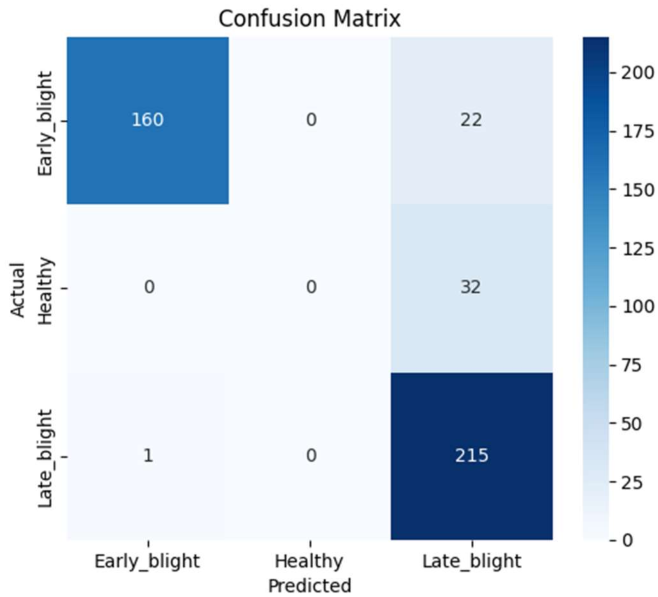

Fig 6: AlexNet Confusion Matrix

Fig 5 & 6 Show the training and validation curves of AlexNet show gradual improvement in accuracy and decreasing loss with more fluctuation than Custom CNN. The confusion matrix shows accurate classification of Late blight and Healthy, but notable misclassification in Early blight cases.

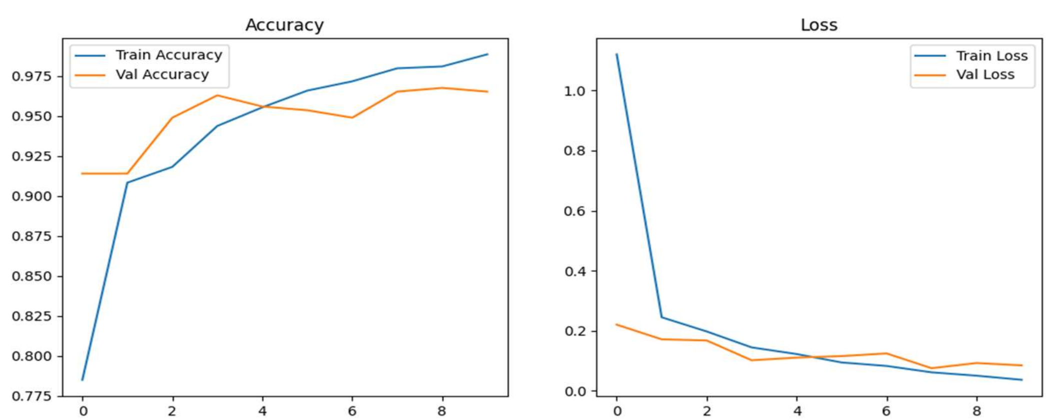

Fig 7: VGG16 Accuracy and Loss

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

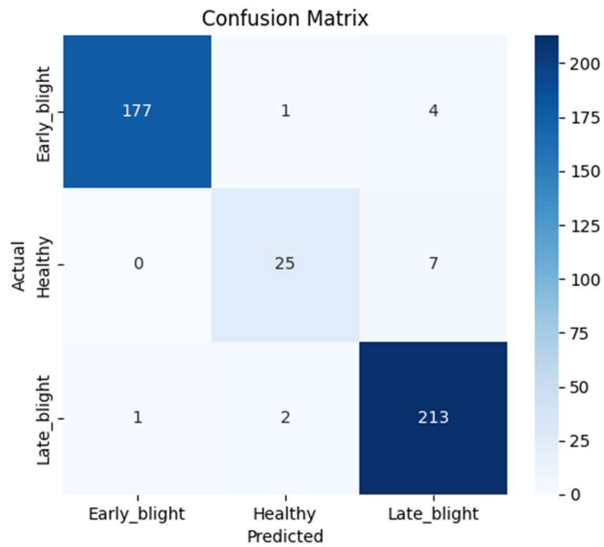

Fig 8: VGG16 Confusion Matrix

Fig 7 & 8 Show the training and validation curves show high and stable accuracy with steadily decreasing loss showing strong convergence and minimal overfitting. The confusion matrix confirms the excellent classification across all classes with only minor misclassifications.

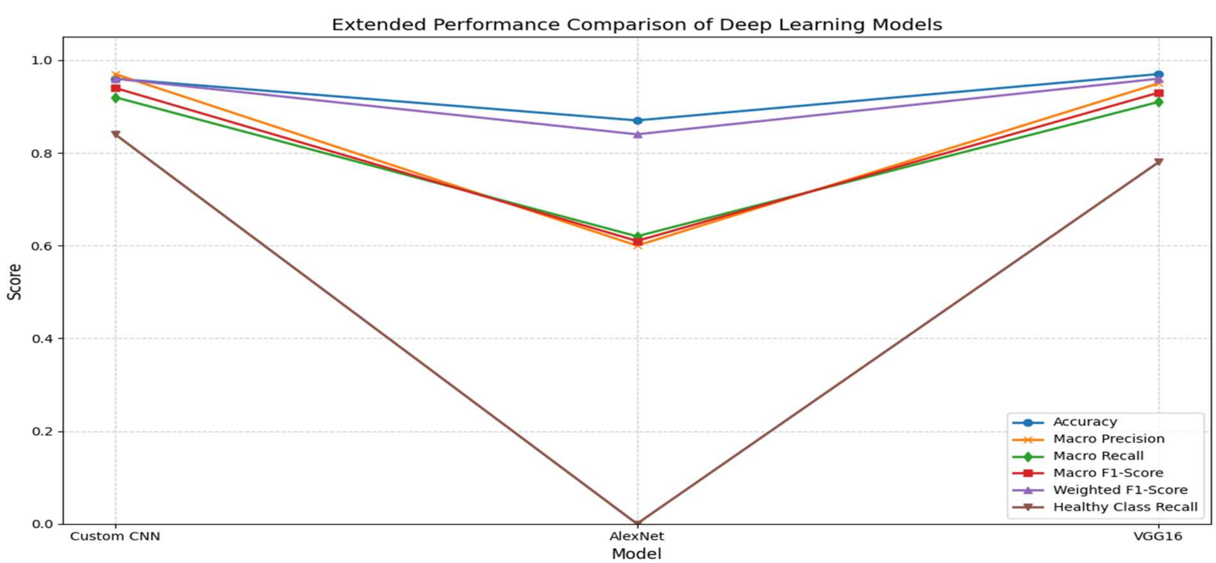

Fig 9: Performance Comparison of Deep Learning Models.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 50– 63

Fig 9 shows a multi-metric comparison of Custom CNN, AlexNet, and VGG16 on six of the most significant performance measures: accuracy, macro precision, macro recall, macro F1-score, weighted F1-score, and Healthy class recall. The plot only indicates that VGG16 and Custom CNN both do an excellent job on all of the metrics, but Custom CNN slightly does not have a better macro precision and Healthy class's recall. AlexNet performs admirably but suffers a huge loss of performance, particularly in the detection of the Healthy class, indicating its weakness in dealing with underrepresented classes. The visual imbalance strengthens the efficiency of light, but consistent models like Custom CNN and VGG16 in attaining balanced classification accuracy.

6. Performance Insights and Generalization Analysis

Literature surveyed documents a paradigm shift from traditional feature-based classifiers to deep learning models for potato leaf disease classification [2]. While some of the older machine learning techniques that are feature-based on GLCM classifiers like SVM and Random Forest proved helpful in practice in early systems for detection, they have largely been replaced by modern Convolutional Neural Networks (CNNs), which obviate handcrafted features because CNNs learn hierarchical representations directly from image information [3]. Among the deep learning models under consideration, Custom CNN, VGG16, and AlexNet are the most noteworthy on the basis of classifying accuracy, computational cost, and field condition adaptability. Custom CNNs have great validation performance with fewer architectures and lower memory consumption. They are then a very edge device and mobile deployment friendly in low-resource farm settings. Their applicability to wide image conditions, when appropriately augmented, suggests their field use adaptability. VGG16, although computationally expensive, always performs well when transfer learning methods fine-tune it. Its deep architecture and strong feature extraction make it produce strong results despite challenging conditions, i.e., noisy data, changing lighting conditions, or non-uniform leaf orientation. VGG16 has worked well on lab-dataset curated pictures and more challenging real-world pictures, with the application of interpretability tools such as Grad-CAM facilitating explanation of its decision-making process. AlexNet, being one of the pioneering CNN models, still finds application due to its light weight and speed of convergence. AlexNet performs well on clean, well-separated data but performs poorly in minority classes, especially when there is class imbalance or no data augmentation. The model could not identify the Healthy class in all the experiments, a shortcoming which highlights the importance of class balance and the model's sensitivity. The overall theme throughout the studies is the overreliance on the PlantVillage dataset, a well-controlled but curated dataset with uniform backgrounds and ideal illumination [6]. It's excellent for giving an easy comparison baseline to models, but it barely resembles the variation of real-world farm environments. Few studies did away with this by employing field-visited images, i.e., outside in real-world conditions with mobile platforms. These recorded performance degradations also indicated that models like VGG16 and Custom CNN are highly accurate if tested under real-world noise and background variation [5,9]. Together, all these results indicate potato leaf disease prediction model choice should not only be accurate, but also generalizable, interpretable, and deployable. Custom CNN and VGG16 are always the first choice for deployable and scalable solutions. Still, AlexNet would have to be improved or used in conjunction with pre-processing techniques for higher robustness [11]. All future work must be focused on heterogeneous datasets and efficient training pipelines to deliver accurate disease diagnosis in real agriculture.

Conclusion

This survey emphasizes the increased contribution of deep learning methodology in the classification of potato leaf disease over conventionally defined methods within the realm of handcrafted features and conventional machine learning. Although previous methods like visual observation, feature extraction using GLCM, and SVM and Random Forest classifiers were the foundation for computer-aided diagnosis, their non-scalability, vulnerability, and incapability to learn have been realized in actual agriculture environments. Deep models, or Convolutional Neural Networks (CNNs), have been identified as high-performance players which can learn rich, hierarchical features from supervised image data. Custom CNN and VGG16 were identified as the best of the architectures under test. Custom CNNs offer an optimal trade-off between performance and computational complexity and are therefore apt for resource-constrained environments such as mobile or edge-based applications. VGG16, while computationally costlier, always delivered high accuracy and robustness, especially with transfer learning augmentation. AlexNet, while historically significant and efficient, was poor with underrepresented categories and variation in real-world images, suggesting limitations in generalizability without other augmentations. Among the replicable findings reported in the literature is excessive dependence on the PlantVillage dataset, a well-controlled but sufficiently annotated image collection. Although perfectly suited to enable fair benchmarking, it is plagued by the lack of richness that comes with field conditions. Some attempts have gone a long way to complement images taken under field conditions, where there was considerable model performance degradation, and foregrounding the requirement for dataset diversity towards successful field deployment. Lastly, it is highlighted in this review that model selection needs to go beyond accuracy measures. Generalizability to noisy and imbalanced data, computability, and deployability are all significant in designing effective plant disease detection systems. Based on this, follow-up work efforts should then be on gathering field-representative datasets, including explainable AI tools for decision explainability, and exploring model compression and optimization techniques for real-time deployment. With continued innovation in these directions, deep learning will realize more of its centre-stage role in developing scalable, accessible, and smart agricultural technology.

References

1. Tambe, U. Y., Shobanadevi, A., Shanthini, A., & Hsu, H. C. (2023). Potato Leaf Disease Classification using Deep Learning: A Convolutional Neural Network Approach. *arXiv preprint arXiv:2311.02338*.
2. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, *7*, 215232.
3. Dame, T. A., Adera, G. B., & Girmaw, D. W. (2025). Deep learning-based potato leaf disease classification and severity assessment. *Discover Applied Sciences*, *7*(7), 1-22.
4. Khalifa, N. E. M., Taha, M. H. N., Abou El-Maged, L. M., & Hassanien, A. E. (2021). Artificial intelligence in potato leaf disease classification: a deep learning approach. *Machine learning and big data analytics paradigms: analysis, applications and challenges*, 63-79.
5. Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanovic, D. (2019). Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry*, *11*(7), 939.
6. Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, *161*, 272-279.
7. Oppenheim, D., Shani, G., Erlich, O., & Tsror, L. (2019). Using deep learning for image-based potato tuber disease detection. *Phytopathology*, *109*(6), 1083-1087.
8. Adluri, V. L., Reddy, P. S., Palthya, S., Jampani, S., & Gandhari, S. P. (2022, October). Potato Leaf Disease Detection and Classification Using VGG16. In *International Joint Conference on Advances in Computational Intelligence* (pp. 147-158). Singapore: Springer Nature Singapore.
9. Alhammad, S. M., Khafaga, D. S., El-Hady, W. M., Samy, F. M., & Hosny, K. M. (2025). Deep learning and explainable AI for classification of potato leaf diseases. *Frontiers in Artificial Intelligence*, *7*, 1449329.
10. Tomy, D. M., Jaison, A., Christopher, A., Tomy, A., Jacob, J., & Harsha, A. (2021, October). Comparison of ALEXNET and VGG16 for Analysis of Plant Leaf Disease. In *2021 International Conference on Advances in Computing and Communications (ICACC)* (pp. 1-6). IEEE.
11. Bhageerathi, T., Anagha, M., & Pushpa, T. S. (2024). Comparative Analysis of Deep Learning Models for Plant Disease Detection.