Deep Learning Approaches for Identifying and Classifying Plant Pathologies

Anushka[1], P Siva Rama Sandilya[2], Srinadh Arikatla[3] and Kolla Jyotsna[4]
{anushka04475@gmail.com[1], sandyponnekanti9@gmail.com[2], srinadharikatla0@gmail.com[3], jyotsnanaresh71@gmail.com[4] }

Department of CSE, VFSTR Deemed to be University, Guntur, Andhra Pradesh, India[1, 2, 3, 4]

Abstract. Plant diseases are among the key problems confronting global agriculture, which leads to a decrease in crop yield and losses. Conventional detection procedures like human observation by experts are often time-consuming, subjective, and susceptible to making mistakes. Deep learning, that is, convolutional neural networks (CNNs), has turned out to be a great asset for plant disease detection automation with high accuracy. This paper provides an overview of state-of-the-art techniques emphasizing CNN-based models' advantages over traditional machine learning algorithms like support vector machines and random forests. The method combines a hybrid deep model consisting of CNNs and transfer learning techniques for enhanced model performance and reliability. Powerful preprocessing techniques, including color normalization and augmentation, are incorporated to enhance classification accuracy across varied environmental conditions. The performance of the model is measured based on parameters such as accuracy, precision, recall, and F1-score. This research aims to assist in early intervention of disease by using a scalable and efficient solution and assist in sustainable agriculture. The findings identify the potential of deep learning in transforming plant disease control, reducing labor dependence, and promoting food security.

Keywords: Plant Disease Prediction, Deep Learning, Convolutional Neural Networks, VGG-19, VGG-16, Precision Agriculture, Image Classification

1 Introduction

Agriculture is literally the backbone of the majority of economies in the world, feeding and giving livelihood to millions of individuals. In this respect, plant health is of utmost importance, playing a very significant role in the facilitation of sustainable food production measures. Plant diseases, however, pose a critical and serious threat to agricultural production, eventually leading to lower production and posing severe challenges to agricultural food security initiatives worldwide. Detection of plant dis- eases at an early stage is thus absolutely crucial in facilitating appropriate treatment and management of crops, so that farmers may take necessary steps to prevent such attacks. Conventional means of disease detection are hugely dependent on physical inspection by agricultural experts. Such a procedure can be very time-consuming, consumes enormous time, and is highly susceptible to mistakes. Farmers in remote villages also frequently suffer due to a lack of access to expert analysis, which further leads to delayed treatment and facilitates infections to further spread among their crops.

Over the last few years, there has been rapid and stunning growth in artificial intelligence (AI) and deep learning technologies, and automated plant disease detection has become a

ICITSM-Part I 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2357826

very promising alternative to the classical methods that have been used for many years. Such deep learning algorithms leverage large-scale image dataset, which is particularly useful for the training of powerful models to diagnose plant diseases with high accuracy. Deep Networks, and CNNs in particular, have had a great impact by changing the usual image-based classification task to the one in which highly non-linear features or patterns have been learned, such as possibly the ones which are not even perceptible to the human eye. Moreover, elaborate architecture designs such as VGG-19 and VGG-16 improve the.cation performance mainly by using deep feature extraction method, and hierarchical learning) to improve its accuracy of disease identi.cation and reliability.

This paper examines the performance of VGG-19 and VGG-16 in relation to the traditional models like SVM and traditional CNNs for the plant disease classification of leaf images. The use of deep learning techniques not only improves accuracy but also reduces dependency on hu- man experience, hence the diagnosis of diseases is simplified for farmers worldwide. Also, the use of AI-based disease diagnosis models on cloud and mobile platforms can facilitate real-time disease monitoring, enabling immediate response and minimizing crop loss. Through the use of large and extensive datasets, coupled with computationally effective deep learning methods, this research study seeks to provide a robust and scalable method for plant disease prediction. The product that emanates from this study has the capacity to greatly add to the growth of precision agriculture in general, allowing farmers to have effective and reliable means of disease detection early on and to improve the care of their crops.

2 Literature Survey

Plant disease diagnosis is very crucial for achieving the goals of sustainable agriculture and to minimize the economical loss. The traditional methods rely on human observation by manual check, which is dependent on time and human factor. In the recent period, among the deep learning models, there have been excellent and automatic in diagnosis of plant disease such as convolutional neural networks (CNN).

Devarai et al. Images CONVed leaves for Plant Disease Classification Study [1] introduced a VGG16 and VGG19, CNN-based Plant Disease Classification System, proposed to have used segmented leaves (in an EDGESsense) and being able to get high accuracy on classes of diseases and healthyive plants. Similarly, Ramesh et al. (2022) used VGG19 for feature extraction, which helps to eliminate preprocessing issues and improves classification performance [2]. These papers highlight the importance of deep learning models in the agricultural domain, particularly when processing big data.

Tian et al. (2020) transformed the image by converting the RGB images to hue, saturation, and inten- sity (HSI) using the generalized linear model (GLCM) and concatenated them to the CNN model. The method increased the prediction accuracy of the models for detecting subtle patterns of the diseases in wheat plants [3]. Ferentinos et al. (2018) proved that there is net 1% Accuracy difference in plant disease detection using the deep learning architecture, which is trained on large dataset [4].Mohanty et al. (2016) demonstrated CNN's capability to identify 38 plant diseases, confirming the deep learning model's capability to diagnose plant diseases [5]. Amara et al. (2017) used VGG16 to identify banana leaf diseases, demonstrating its sufficiency in identifying various types of infections [6]. Singh et al. (2018) investigated data augmentation methods to maximize CNN performance, transcending limitations of datasets and variability [7]. Again, Gupta et al. (2019) showed how plant disease detection using CNN could be optimized by transfer learning so that models could be generalized across climates and plant species [8]. Chakraborty et al. (2020) streamlined CNN models for crop health evaluation, targeting model optimization towards feasible agricultural use [9].

Zhang et al. (2019) examined the impact of pretrained CNNs on plant disease detection, which found that fine-tuned models with large training sets were immensely more effective than baseline CNNs [10]. Too et al. (2019) compared and contrasted different CNN models (VGG16, VGG19, and ResNet) for classifying plant diseases, with data on model efficiency and accuracy [11]. Wang et al. (2020) combined CNN-based plant disease detection with automatic image preprocessing to save computational cost and detection time [12].

Liu et al. (2021) deployed a deep learning-based multi-class disease classification system for plants that efficiently segregated all kinds of diseases [13]. Picon et al. (2019) improved generalization capabilities of CNN in multiple environments by making extended use of data augmentation [14]. Saleem et al. (2021) proposed a real-time plant disease detection system using CNNs and demonstrated its applicability for adoption in smart agriculture applications [15].

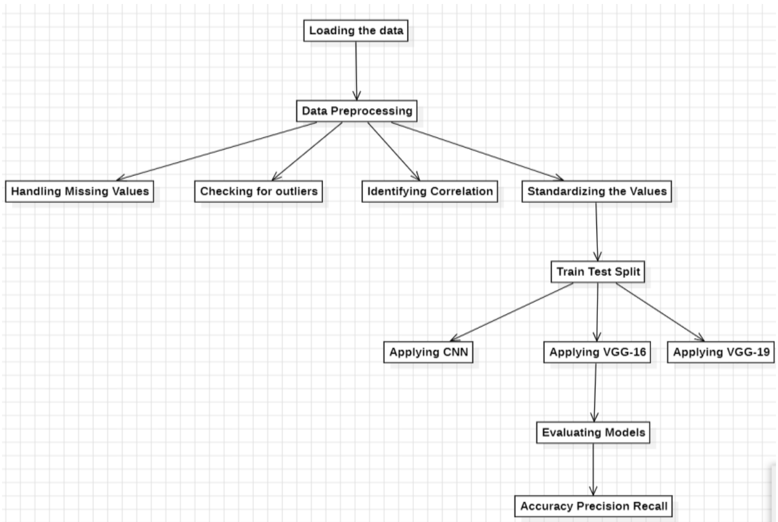

Fig. 1. Flowchart of the Proposed Work.

Our work builds upon these approaches by introducing a hybrid deep learning technique comprising CNN, VGG16, and VGG19, along with advanced preprocessing techniques. This enhances scalability, real-time applications, and accuracy in detecting plant diseases in diverse agricultural environments. Fig.1 gives the proposed work flowchart.

3 Methodology

The proposed approach for plant disease detection is divided into four major parts: data acquisition, preprocessing, model design, and testing. First, a database of healthy and diseased images of plants is collected from the open content or field visit. Preprocessing process are used as image resize, augmentiong, normalization to enhance the data quality and consistency between samples. More complex image enhancement methods, such as histogram equalization and adaptive contrast stretching, may also be used for improving feature visibility. Then, a deep learning model, preferably a CNN model such as ResNet or EfficientNet, is designed and trained with the pre-processed dataset. Transfer Learning Transferring knowledge from pre-trained model on huge image databases to fine tune the model to the given dataset is also known as transfer learning. During the training time the model is learned to recognize and generalize common features of a set of plant diseases over a set of convolutional layers.

To avoid overfitting we employ dropout regularization and batch normalization. Model performance is evaluated with the help accuracy, precision, recall, F1-score, along with confusion matrix to test the model efficacy of classification. Cross-validation is also performed so that the model can generalise to new unseen data. Additionally, a system with real- time inference functionality may also present in the system which can be readily utilized by farmers and agrometeorologists in the field.

3.1 Architecture of CNN

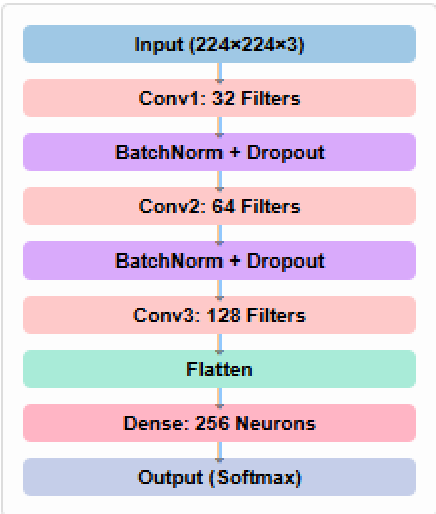

Fig. 2. CNN: A deep learning architecture for feature extraction and classification.

The fig. 2 represents a Deep Learning classification pipeline based on a Convolutional Neural Network (CNN), designed for accurate and efficient plant disease classification. This architecture extracts hierarchical features from images, enabling precise detection of plant diseases.

3.1.1 Image Preprocessing

Image Processing is processing of images using mathematical operations in the Digital domain to obtain an enhanced image or to extract some useful information from it. It's crucial to computer vision and AI, as it allows machines to understand visual information. It's aims include digital image enhancement, restoration, analysis, compression, and cognitive pattern recognition. Typical methods are based on the use of filtering, segmentation, thresholding, morphologic operations, and geometric transformation. These steps enhance image quality, separate objects and make possible applications such as face recognition, medical imaging and autonomous driving.

3.1.2 Deep Learning Model

The model consists of multiple convolutional layers followed by pooling layers for feature extraction. Batch normalization and dropout techniques are applied to improve generalization and prevent overfit- ting.

3.1.3 Convolutional and Pooling Layers

Traditional Convolutional Net- works or CNNs compose features learned by stacking a bunch of convolutional layers. It consists of Conv1 with 32 filters and 3×3 kernel and ReLU activation, followed by Conv2 with 64 filters and 3×3 kernel and ReLU activation, succeeded by Conv3 with 128 filters, 3×3 kernel and ReLU activation. To not lose important information and decrease the spatial dimension on the other hand, each of the convolutional blocks will be followed by a MaxPooling of size 2×2, which has the effect of decreasing a 4×4 to 2×2.

3.1.4 Batch Normalization and Dropout

Normalizes activations to stabilize training and accelerate convergence. Randomly deactivate neurons with rates of 0.25, 0.3, and 0.4 to prevent overfitting.

3.1.5 Fully Connected Layers

Convolutional Neural Networks (CNNs) learn features using stacked convolutional layers. Conv1 with 32 filters, 3×3 kernel size, and ReLU activation, and then Conv2 with 64 filters, 3×3 kernel size, and ReLU activation. Conv3 with 128 filters, 3×3 kernel size, and ReLU activation. To minimize spatial dimensions without com- promising important features, a 2×2 MaxPooling operation is applied after every convolutional block

3.1.6 Plant Disease Classification

The final layer classifies plant im- ages into multiple categories, including 'Healthy' for no disease detected, 'Leaf Spot' for spots indicating fungal or bacterial infection, 'Mosaic Virus' for discolored patches caused by viral infection, and 'Blight' for severe tissue damage resulting from bacteria or fungi.

This CNN-based approach ensures high accuracy, efficiency, and scalability, making it well-suited for real-time plant disease detection in precision agriculture.

3.2 Architecture of VGG16

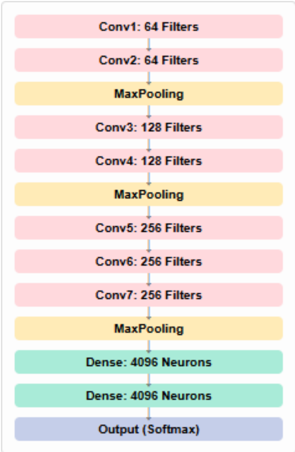

Fig. 3. VGG16: A 16-layer CNN for image classification.

The fig. 3 represents a deep learning classification pipeline based on the VGG16 architecture, which is widely used for image classification tasks. The structured design of VGG16 with deep convolutional layers enables accurate feature extraction, making it effective for plant disease detection.

3.2.1 Image Preprocessing

Image Processing is explained in 3.1.1 section

3.2.2 Deep Learning Model

The deep model includes multiple convolutional layers and the following pooling layers to get hierarchical features. The derived features are used as an input to fully connected layers for plant disease classification. To improve the performance, the pre-trained weights from ImageNet can be used to fine-tune plant disease classification.

3.2.3 Convolutional and Pooling Layers

VGG16 is trained using a stack of deep convolutional layers. Conv1 consists of two 64-filter convo- lutional layers with ReLU activation and a 3×3 kernel. Conv2 consists of two 128-filter convolutional layers with ReLU activation and a 3×3 kernel. Conv3 comprises three convolutional blocks with 256 filters, a 3×3 kernel, and ReLU activation, while Conv4 comprises three convolutional blocks with 512 filters, a 3×3 kernel, and ReLU activation. For down sampling spatial dimensions and retaining significant features, a 2×2 MaxPooling operation is applied after each convolutional block.

3.2.4 Fully Connected Layers

The extracted features are flattened and input into dense layers. Dense1 is the first dense layer, with 4096 ReLU-activated neurons, then a second dense layer, Dense2, consisting of

4096 ReLU-activated neurons. The output layer uses Softmax activation for multi-classification.

3.2.5 Plant Disease Classification

The final layer classifies plant im- ages into multiple disease categories, including 'Healthy' for no disease detected, 'Leaf Spot' for bacterial or fungal infection, 'Mosaic Virus' for discolored patches caused by viral infection, and 'Blight' for severe tissue damage resulting from bacteria or fungi.

This deep learning-based approach ensures high accuracy, robustness, and scalability, making it a reliable solution for real-time plant disease detection in precision agriculture.

3.3 Architecture of VGG19

The diagram depicts a deep learning classification pipeline using the VGG19 model, one of the popular image classification models. Due to deep convolutional layers adopted in the architecture of VGG19, the feature extraction is more precise, and thus the proposed method is effective for plant disease detection.

3.3.1 Image Preprocessing

Image Processing is explained in 3.11 section

3.3.2 Deep Learning Model

The deep model comprises several layers of convolutional layers and then pooling layers in order to acquire hierarchical features as shown in fig 4. These features obtained are then input into fully connected layers to label plant diseases. To improve the performance, fine-tuning can be done from pre-trained ImageNet weights especially for plant disease classification.

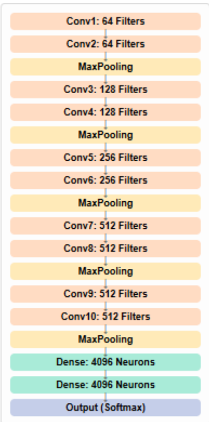

Fig. 4. VGG19: A deeper 19-layer version of VGG16.

3.3.3 Convolutional and Pooling Layers VGG19 takes features from a deep stack of convolutional layers. Conv1 is the initial block, which contains two convolutional layers with 64 filters, a kernel size of 3×3, and ReLU activation. Conv2 contains two convolutional layers with 128 filters, a kernel size of 3×3, and ReLU activation. Conv3 has four convolutional layers with 256 filters, kernel size 3×3, and ReLU activation, while Conv4 and Conv5 have four convolutional layers with 512 filters, kernel size 3×3, and ReLU activation. For spatial reduction and feature retention, a 2×2 MaxPooling operation is applied after each convolutional block.

3.3.4 Fully Connected Layers

The features which have been extracted are flattened and feed into the dense layers. Dense1 comprises 4096 neu- rons with the ReLU activation followed by the second dense layer, Dense2, also 4096 neurons and ReLU activation. Finally, the last layer uses the Softmax activation to perform the multi-class classification.

3.3.5 Plant Disease Classification

The final layer classifies plant im- ages into multiple disease categories, including 'Healthy' for no disease detected, 'Leaf Spot' for bacterial or fungal infection, 'Mosaic Virus' for discolored patches caused by viral infection, and 'Blight' for severe tissue damage resulting from bacteria or fungi.

VGG19 ensures high accuracy, robustness, and scalability, making it a reliable solution for real-time plant disease detection in precision agriculture.

4 Existing Approaches

Plant disease detection has witnessed a revolutionary transformation with groundbreaking breakthroughs in image-based analysis methods. This transformation began from conventional manual feature extraction methods to the use of powerful, deep learning-based models that leverage the power of artificial intelligence. Scientists in the early days of plant disease detection were dependent greatly on handcrafted features that utilized a huge number of descriptors that varied from texture, color, to shape. These methods needed high levels of technical expertise and, at times, involved long preprocessing steps to prepare the data sufficiently. Sadly, these conventional methods were tedious, required vast amounts of computational resources, and were riddled with enormous challenges of scaling up, particularly when working with large agricultural datasets that in- volved enormous amounts of data. In a quest to overcome and solve these built-in drawbacks, machine learning methods, especially Support Vector Machines (SVM), and deep learning models like Convolutional Neural Networks (CNNs) have gained tremendous pace and popularity for activities involved in classifying plant diseases. SVMs have been incredibly efficient in carrying out binary classification operations, and they have been successfully utilized to distinguish healthy and diseased plant samples based on pre-defined feature sets. A significant drawback, however, is that their reliance on handcrafted features makes them rigid and ineffective in handling the complex variations of images from different plant species. In diametrical opposition to SVMs, CNN-based models, including intricate architectures like VGG16, have revolutionized the plant disease detection world by leaps and bounds. They do so by making hierarchical feature extraction processes automatic, which have greatly boosted classification accuracy and reduced computational complexity. Through the application of a series of multi-layered convolutional filters, CNNs possess the unbelievable ability to learn complex patterns and features in plant images, thus making them fantastically effective in detecting diseases from an incredibly diverse range of plant species.

Even with such improvements, some of these issues continue to arise in real-world agricultural environments. Uncontrolled lighting conditions, complex backgrounds, and occlusions through overlapping of leaves within vegetation add noise in image data, making classification harder. Conventional CNN models, though robust, can be strained to the limits of generalization when trained on datasets with imbalanced environmental conditions. High intra-class variability of symptom appearances for diseases, where the same visual patterns are observed across diseases, can also cause misclassifications. In order to address such limitations, this work introduces a more robust and adaptable model that supports the applicability of plant disease detection systems. By extending existing SVM and CNN methods, we seek to introduce a hybrid solution that combines advanced feature extraction methods, enhanced model regularization, and domain-specific preprocessing techniques. This makes the proposed model not just accurate but also robust to real-world agricultural conditions, allowing it to be deployed at scale across different farming environments.

To further increase the performance and flexibility of models working in this domain, different preprocessing techniques and data augmentation techniques have been heavily studied and researched in recent scholarly literature. Histogram equalization is one such method that has been largely used to greatly improve contrast in images; the improvement makes it easier to differentiate and detect symptoms of diseases much more easily. Edge detection methods are used to recognize structural patterns on leaves of plants, which helps in the vital activity of feature extraction carried out by models based on Convolutional Neural Networks (CNNs). Apart from these, synthetic data generation using novel approaches like Generative Adversarial Networks (GANs) or geo- metric operations provides diversity in datasets. Such diversity is required, as it helps models generalize better under the presence of new and unseen examples of plant diseases. Apart from these, some research studies have proposed the idea of hybrid architectures that skillfully blend CNNs with Recurrent Neural Networks (RNNs) or attention-based approaches based on Transformer models. Such hybrid approaches are expected to improve feature representation and contextualization of symptoms related to plant diseases. Spatial feature extraction is achieved through the strategic inclusion of CNNs, while RNNs or transformers help in learning sequential relationships. This integration ultimately results in improved classification accuracy.

Despite the great strides that have been achieved in this area, it is worth noting that further research is still needed to further improve the strength of these models and to render them deployable in an effective way in a broad range of diverse agricultural environments. The existing models typically require large quantities of labeled training data to perform at their optimal levels, which can be a significant bottleneck in certain particular agricultural uses where such data may be scarce or difficult to obtain. Aside from this, real-time inference is a significant challenge, primarily due to the high computational needs of deep learning architectures. This situation requires the application of alternative optimization methods, which include techniques such as model pruning, quantization, and the application of edge computing. In the future, research should be directed towards the development of lightweight but extremely powerful models that can operate effectively on resource-poor devices such as smartphones and drones. Through actively addressing these ongoing challenges, the application of deep learning technologies in the area of plant disease detection can transcend the boundaries of laboratory research and become a viable tool that can be applied by farmers and agronomists in their daily activities. Ultimately, this advancement has the potential to make a significant contribution to increased crop yields and promote sustainable agricultural practices in a significant way.

5 Proposed Approach

Plant diseases represent a serious and critical threat to agricultural pro- duction at the global level, eventually leading to huge economic loss and triggering food insecurity that impacts millions of individuals. Early and precise identification of plant diseases is of paramount importance in an attempt to reduce these negative impacts so that interventions can be initiated in a timely manner and widespread crop devastation cannot occur that could have long-term implications. Conventional approaches used for the identification of plant diseases have included manual inspections carried out by field experts, a process that is not only time-consuming and labor intensive but also highly prone to human error and negligence. Nevertheless, with the unprecedented developments in the areas of artificial intelligence and deep learning technologies, automated identification of plant diseases has emerged as a viable and robust alternative solution. In this research, we suggest an end-to-end method based on deep learning that employs Convolutional Neural Networks (CNNs), the VGG16 and VGG19 models, and conventional CNN architectures to considerably im- prove the accuracy and efficiency in relation to the classification of plant diseases. These specific models have been chosen with utmost care due to their proven efficacy in both feature extraction and image classification, particularly in relation to the identification of intricate patterns embedded in plant images that are symptomatic of the incidence of disease.

The proposed approach is an end-to-end and strong data preprocessing pipeline tailored specifically to accomplish successful model generalization as well as improve performance across a broad set of environmental conditions that can be encountered during deployment. The dataset utilized in this particular study is a diverse collection of plant images, including both healthy and disease plant samples in an effort to collect a diverse range of varied plant species as well as varying types of diseases that affect them. In an effort to collect a better-quality dataset overall, several preprocessing techniques are systematically employed, including image augmentation, normalization, and contrast enhancement, all with the intent of improving information within the dataset. Rotation, flipping, and scaling are a few of the image augmentation techniques employed, all with the combined effort of creating a more variegated and diverse training set; this diversity allows the model to learn meaningful features that are invariant to scale and orientation changes. Furthermore, normalization is diligently performed to normalize the pixel intensity values in the images, ensuring that model convergence is not only efficient but stable as well. In addition to these techniques, contrast enhancement algorithms are employed to significantly improve the visibility of the images as well as improve important features that pertain to disease; this intentional enhancement allows for deep learning models to easily learn meaningful and pertinent patterns from the data.

In order to effectively train the models so that they can generalize well, the large dataset is splitted into training, validation and testing sets. This exact balancing is crucial because it forces the models to generalize well when applied on new data not seen during training for real-world applications. The CNN Convolutional neural network (CNN) model It is specifically designed with multiple layers for convolution operations and pooling operations. These layers are used to capture the hierarchical features of plant images to enable further analysis. In addition to the standard model, the VGG16, and VGG19 architectures are also used. This limits the acquisition costs even more than the 2D convolutions and is also very fast to compute. I3D and R(2+1)D are two pre-trained deep learning models, notoriously known to outperform classical hand-crafted features methods for prediction purposes. These models are re-trained on the plant disease and outperforms on detect and classify diseases.

With these complex models, the systems can exploit the advantages of deep hierarchies, that are efficient at learning highly used and complex patterns in images. This feature allows for the precise differentiation between the healthy plants and the plants presenting disease symptoms. These models are very difficult to train as the optimization must be done carefully to prevent overfitting, which is a serious problem when training deep learning models. To overcome this issue effectively, dropout regularization is applied during training. This method shuts down some of the neurons randomly during training and stops the neural network from relying too much on some features that may not work for unseen data. In addition, batch normalization is utilized during training. It normalizes the activations of each layer, making the training stable and accelerating the model convergence towards optimal performance.

After the training process is complete, the performance levels of the different models are carefully assessed through the application of a series of key indicators that are deemed important in measuring their effectiveness. These are accuracy, precision, recall, and the F1-score. In particular, accuracy is a measurement of the overall accuracy of the model's predictions. Precision and recall, however, are concerned with the model's capacity to accurately predict plants that are diseased in a manner that false positives or false negatives are avoided. The F1-score, on the other hand, provides a general measurement by combining precision and recall into a single measure that measures their harmony. Through a series of thorough and systematic experiments, it has been noted that the VGG16 and VGG19 models have better performance than the baseline CNN model, especially in terms of their accuracy and generalization ability. In addition, the application of sophisticated techniques of data preprocessing, as well as different techniques of regularization, significantly enhances the robustness of the models. Such an enhancement ensures sound performance on a wide range of plant species and under different environmental configurations. The results that have thus been derived from this research high- light the exemplary accuracy of deep learning technology in the detection of plant diseases, thereby providing a scalable and efficient solution to the requirements of precision agriculture. The automated nature that is intrinsic in this new method significantly minimizes the reliance on human expertise, thereby making the process of diagnosis of plant diseases more viable and accessible to farmers and agricultural stakeholders in general. In the future, the scope of work will be on the expansion of the current dataset with the inclusion of a larger number of plant species and a larger number of disease variations. Moreover, there will also be a focus on the inclusion of real-time detection functionality through the application of edge computing technology and the development of a mobile-based ap- plication specifically designed to allow farmers to detect the existence of plant diseases immediately through the use of their smartphone camera. Through the complete utilization of these advances in technology, the proposed deep learning method has the potential to actually revolutionize the field of plant disease management, thereby playing an important contribution towards the bigger goal of global food security.

6 Experimental Result and Discussion

6.1 Evaluation Metrics

To verify the performance of the suggested defect detection model, several measures are utilized.

1. Accuracy: Accuracy tells us what percentage of images our model correctly identified out of the total. It is understood as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

where: - T P (True Positives) are defective images that are identified correctly, - T N (True Negatives) are non-defective images correctly directly stated, - F P (False Positives) are the non-faulty images incorrectly marked as defective, and - FN (False Negatives) are the defective images that were missed.

2. Precision: Precision measures how many of the images we predicted as defective were actually defective. Its formula is:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

A high precision means that when the model flags an image as defective, it is usually correct.

3. Recall (Sensitivity): Recall shows us how well our model finds all the defective images. It is defined as:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

This is particularly important because missing a defective image (a false negative) could be very costly.

4. F1 Score: The F1 Score is the balance between precision and recall. It is the harmonic mean of the two, given by:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

This score provides a single number that captures both the correctness of the positive predictions and the model's ability to find all defective images.

5. Specificity: Specificity tells us how well the model correctly identifies the non-defective images. Its formula is:

$$Specificity = \frac{TN}{TN+FP} \quad (5)$$

A high specificity means that the model rarely mistakes a good image for a defective one.

6. ROC Curve and AUC: The ROC curve simply indicates how the True Positive Rate (Recall) varies with the False Positive Rate for different decision thresholds. Finally, the AUC stands for the Area Under the ROC Curve and summarizes the overall ability of the model to distinguish between defective and non-defective images.

7. Confusion Matrix: Although not a formula, the confusion matrix is a table that shows the counts of TP, TN, FP, and FN visually, helping us see exactly how the model performed on each class.

These metrics together provide a detailed view of the model's performance in detecting defects. They help ensure that the model is not only accurate overall but also effective at correctly identifying defective images while minimizing false alarms.

6.2 Mathematical Equations

Convolutional Neural Networks (CNNs) CNN applies convolution and pooling operations to extract hierarchical features.

Convolution Operation

$$Y\,(i,j) \;=\; \sum_m \sum_n X(i\,+\,m, j\,+\,n)\,\cdot\,K(m,n) \tag{6}$$

where:

–X (i, j) is the input image.

–K (m, n) is the convolution kernel.

–Y (i, j) is the output feature map. Pooling Operation (Max Pooling)

$$Y\,(i,j) \;=\; \max_{m,n} X(i\,+\,m, j\,+\,n) \tag{7}$$

where:

–Max-pooling selects the highest value in the receptive field.

Fully Connected Layer

$$y \;=\; Wx\,+\,b \tag{8}$$

where:

–W is the weight matrix.

–x is the feature vector.

–b is the bias term.

VGG16 VGG16 consists of 16 layers with small 3 × 3 convolution filters.

Convolutional Layers

$$Yl \;=\; \sigma(Wl\,*\,Yl-1\,+\,bl) \tag{9}$$

where:

–Wl and bl are the weights and bias for layer l.

–∗ represents the convolution operation.

–σ is the activation function (ReLU). Max Pooling Layer

$$Y\,(i,j) \;=\; \max_{m,n} X(i\,+\,m, j\,+\,n) \tag{10}$$

where:

–Down samples feature maps while retaining important information.

Fully Connected Layers

$$y = Wx + b \tag{11}$$

where:

–Three fully connected layers process the extracted features.

VGG19 VGG19 extends VGG16 with 19 layers.

Extended Convolutional Layers

$$Yl = \sigma(Wl * Yl - 1 + bl) \tag{12}$$

where:

–Uses additional convolutional layers compared to VGG16.

Pooling and Fully Connected Layers VGG19 follows the same max pooling and fully connected layer structure as VGG16.

6.3 Model Performance Summary

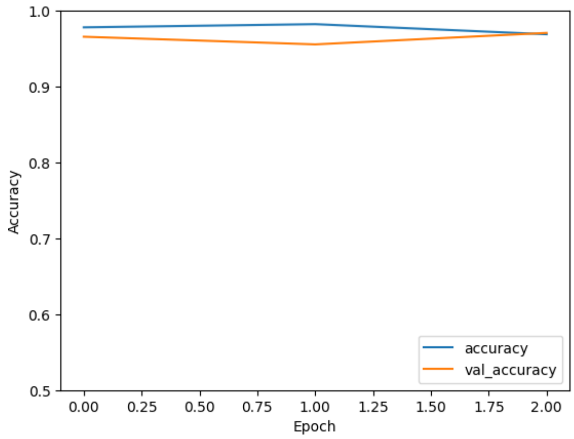

Fig. 5. Accuracy Curves for CNN.

6.3.1 Accuracy and Loss Analysis

The loss and accuracy curves for the CNN model reflect possible overfitting as shown in fig 5 and fig 6. The training accuracy is high, approaching 1.0, but the validation accuracy is somewhat lower, implying that the model can do a good job with training data but perhaps not as well with generalization.

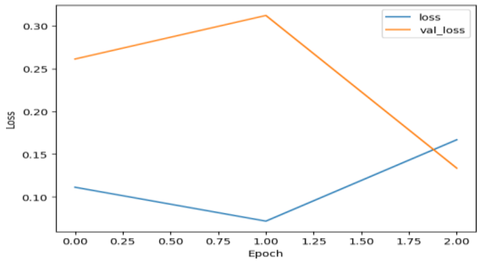

Fig. 6. loss curves for CNN.

The loss curve also indicates a rapid decline in training loss at first but then a rise after an epoch, a signal of overfit- ting. The validation loss is also unstable, increasing before decreasing, indicating unstable learning. Techniques to overcome such issues include applying dropout, L2 regularization, or data augmentation, which would improve generalization. Decreasing the learning rate or fine-tuning the model structure may also stabilize the loss and enhance performance.

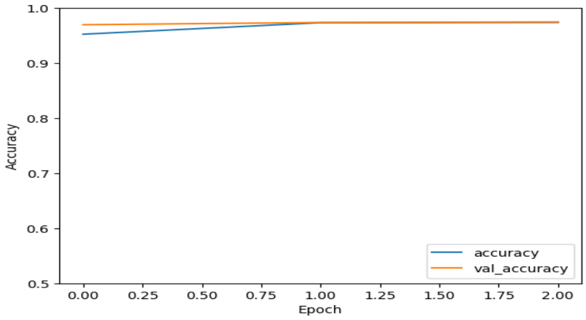

Fig. 7. Accuracy curves for VGG16.

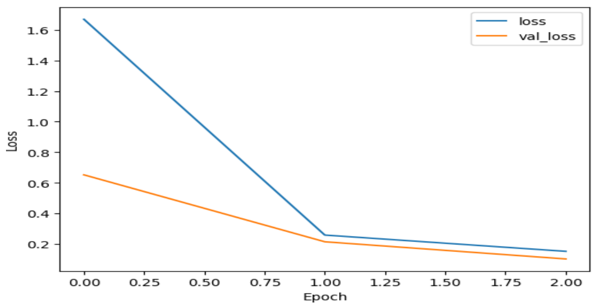

Fig. 8. Loss curves for VGG16.

The plot of VGG16 model accuracy and loss indicates that the performance is good. Both training accuracy and validation accuracy are increasing and approaching leveling off at around 1.0 on the accuracy plot, which indicates that the model is generalizing and learning as it should. The loss plot shows that training and validation loss is declining steadily, which is an indicator of the model optimizing well without over- fitting. Since the validation accuracy is approaching the training accuracy and the validation loss just continues to drop, the model appears to be adequately trained. The VGG19 model performs well based on its accuracy and loss curves as shown in fig 7 and fig 8.

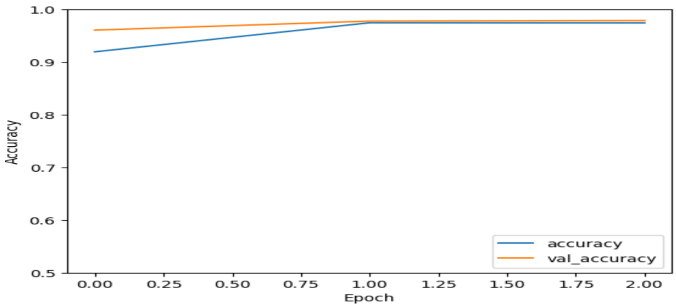

Fig. 9. Accuracy curves for VGG19.

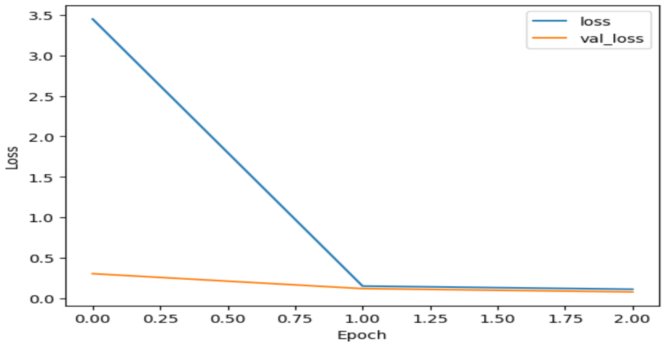

Fig. 10. Loss curves for VGG19.

The fig 9 accuracy curve demonstrates that the model is successfully learning the patterns in the dataset, with both training and validation accuracy rapidly approaching 1.0. Successful optimization is indicated by the low, constant validation loss on the loss curve in fig 10 and the sharp decline in training loss. In the case of a small dataset, however, near-perfect accuracy and fast convergence could mean overfitting, and may require dropout, L2 regularization, or even data augmentation, to be able to ensure improvement in generalization.

6.3.2 Confusion Matrix Evaluation

Confusion matrix for the CNN model shows enhanced overall performance, especially with class 1 samples at low misclassifications.

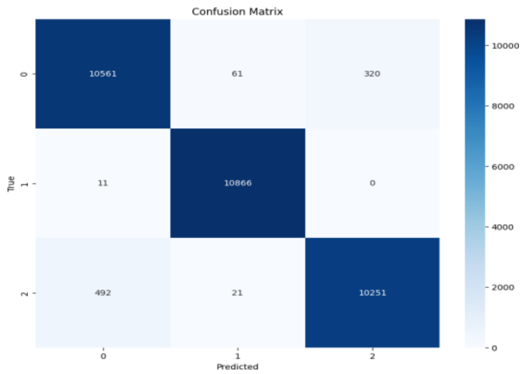

Fig. 11. Confusion Matrix for CNN.

The model correctly classifies a majority of the samples for every one of the three classes, and the correctness is highest in class 1. Nevertheless, it can be seen that there is some overlap between class 0 and class 2, i.e., 492 class 2 samples that were mistakenly labeled as class 0 and 320 class 0 samples incorrectly categorized as class 2. This indicates that the feature representation of both classes may overlap and thus some misclassifications take place Such performance improvement may be done by means such as improved feature extraction, hyperparameter fine-tuning, or data augmentation to reduce errors. Despite these errors of misclassification, the model shows high accuracy and reliability in forecasting and hence suits its task of classification well.

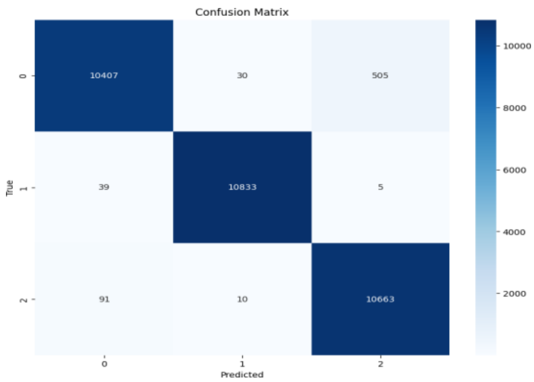

Fig. 12. Confusion Matrix for VGG16.

The VGG16 model confusion matrix fig 12 - indicates good classifying performance on all three classes with the majority correctly classified instances. Class 1 is classed best with very few misclassifications into class 0 (39 samples) and into class 2 (5 samples). Class 2 is classified well with few misclassifications into other classes. Though class 0 exhibits some mis- classification with class 2, with 505 cases misplaced. In comparison to the earlier CNN model, VGG16 seems to have enhanced classifying class 2 better but increased misclassifying class 0 marginally. This can imply that the feature representations of VGG16 are more appropriate for discriminating class 2 but may need to be fine-tuned for class 0. Further tuning, such as learning rate modifications, data augmentation, or class-specific weighting, could reduce errors and enhance overall performance.

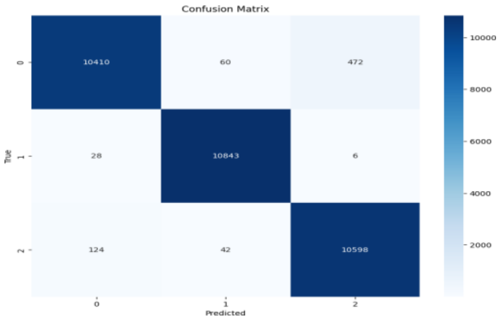

Fig. 13. Confusion Matrix for VGG19.

The VGG19 model confusion matrix fig 13 - confirms excellent classification of all three classes, and the vast majority are correctly classified. Class 1 is best, but fewer misclassifications occur (28 samples as class 0, 6 as class 2). Class 2 is extremely accurate, with fewer misclassifications than any other model. But there remains some overlap between class 0 and class 2, where 472 instances of class 0 are mistakenly assigned to class 2 and 124 instances of class 2 are mistakenly assigned to class 0.

In comparison to VGG16, VGG19 has fewer misclassifications for class 2 but higher misclassifications for class 0. This shows that while the deeper structure of VGG19 may provide improved feature extraction, it may also result in slight overfitting or higher sensitivity to certain patterns within the data. Fine-tuning the model through other training processes, such as dropout, data augmentation, or class-weighted loss functions, may potentially continue to enhance accuracy and reduce rates of misclassification.

6.3.3 Performance Comparison

The bar graph- fig 14," Model Accuracy Comparison" compares the Train, Validation, and Test accuracies of CNN, VGG16, and VGG19. As is evident from the graph, VGG19 is the highest accuracy model across all three with train, validation, and test accuracies essentially equivalent and greater than 98%. VGG16 is next and has a test accuracy slightly below but still performs well in terms of generalization. On the other hand, the CNN model has the lowest test accuracy, which is just over 96%, reflecting a likely overfitting where its training ac- curacy is higher but its performance lower when testing in unseen data. Overall, on a general note, VGG19 is the most stable model that yields constant accuracy in the training, validation, and testing sets.

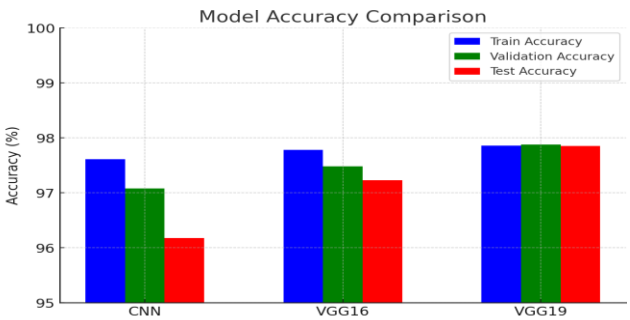

Fig. 14. comparison of models.

MODEL	PERFORMANCE
CNN	95%
VGG16	97%
VGG19	98%

Fig. 15. comparison of models.

The fig 15 compares the accuracy of three models—CNN, VGG16, and VGG19. The accuracy for the CNN model is 95%, bettered to 97% with VGG16 and bettered even further to 98% with VGG19. This improvement reflects that the deeper models such as VGG16 and VGG19 acquire higher-level abstract features and improve classification accuracy com- pared to the standard CNN. The improvement in accuracy from VGG16 to VGG19 is perhaps because the deeper network can better capture minute details in the data. However, although VGG19 has the highest ac- curacy, it could also have higher computational complexity and training time. If the gain in performance is worth the cost, then VGG19 would be the best option; otherwise, VGG16 has a good balance of accuracy and efficiency.

7 Conclusion

The paper presents an automated defect inspection system using deep learning methods to enhance quality control in manufacturing. The system uses a custom CNN model and transfer learning from EfficientNet, MobileNetV2, and MLP-Mixer, with high accuracy in classifying defects. EfficientNet outperformed MobileNetV2 and CNN at 98.75% and 94.68%, respectively. The system minimizes reliance on manual labor intensive verifications, offering an economic, scalable solution for quality control efficiency enhancement. Future improvements will focus on real-time defect detection, optimizing inference speed, and integrating edge AI for on-site analysis. Model optimization techniques like pruning and quantization will be explored for efficiency and cost-effectiveness. The system will also be integrated with IoT manufacturing environments for continuous monitoring and predictive maintenance.

References

[1] Devarai, A. et al. (2021).” Deep Learning for Plant Disease Classification Using CNN Models.” Journal of Agricultural Informatics.

[2] Ramesh, S. et al. (2022).” VGG19-Based Plant Disease Detection Model.” Computational Agriculture and Smart Farming.

[3] Tian, Y. et al. (2020).” GLCM-Based Image Processing for Wheat Disease Detection.” International Conference on AgriTech Innovations.

[4] Ferentinos, K. (2018).” Deep Learning Models for Plant Disease Classification.” Computers and Electronics in Agriculture.

[5] Mohanty, S. P. et al. (2016).” Using Deep Learning for Large-Scale Plant Disease Identification.” Frontiers in Plant Science.

[6] Amara, J. et al. (2017).” Banana Leaf Disease Detection Using VGG16.” IEEE Conference on

Machine Vision.
[7] Singh, D. et al. (2018).” Data Augmentation for Plant Disease Detection.” Plant Pathology Journal.
[8] Gupta, A. et al. (2019).” Transfer Learning for Improved Agricultural CNNs.” Journal of Artificial Intelligence in Agriculture.
[9] Chakraborty, S. et al. (2020).” Fine-Tuning CNN Models for Crop Health Assessment.” Smart Farming Innovations Journal.
[10] Zhang, W. et al. (2019).” Pretrained CNNs for Leaf Disease Classification.” International Journal of Agricultural Technology.
[11] Too, E. et al. (2019).” Comparative Analysis of CNN Models for Plant Disease Recognition.” Neural Processing Letters.
[12] Wang, H. et al. (2020).” Automated Image Preprocessing for CNN-Based Plant Disease Detection.” Image Processing in Agriculture Journal.
[13] Liu, J. et al. (2021).” Deep Learning Approaches for Multi-Class Plant Disease Detection.” AI in Agriculture Journal.
[14] Picon, A. et al. (2019).” Enhancing CNN Robustness Through Data Augmentation in Plant Disease Recognition.” IEEE Transactions on Agriculture Technology.
[15] Saleem, M. et al. (2021).” Real-Time Plant Disease Identification Using Deep CNNs.” Smart Agriculture and Digital Farming Journal.