International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A HYBRID VGGNET–VISION TRANSFORMER APPROACH FOR LEAF DISEASE CLASSIFICATION USING PGGAN-AUGMENTATION

Rina Mahakud[1*], Mihir Narayan Mohanty[2], Dayal Kumar Behera[3], Binod Kumar Pattanayak [4], Pravat Kumar Rautaray[5], Bibhuprasad Mohanty[6]

[1*]Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

[2]Department of Electronics and Communication Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

[3]Department of Computer Science and Engineering, KIIT Deemed to be University, Bhubaneswar, India

[4]Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

[5]Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

[6]Department of Electronics and Communication Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

ABSTRACT

- Plant diseases represent a significant issue to agriculture field globally, causing substantial reductions in crop yield and considerable economic losses. Automated plant disease detection from leaf images has gained momentum with advances in deep learning, yet its effectiveness is often hindered by limited and imbalanced datasets. This work presents a hybrid deep learning framework as the integration of the Visual Geometry Group (VGG16) convolutional network with a Vision Transformer (ViT) to capture both local texture information and global contextual dependencies. To overcome data scarcity, Progressive Growing of GANs (PGGAN) is employed to synthesize realistic plant leaf images for augmentation. Experimental results on tomato and potato disease datasets demonstrate that the proposed VGG16–ViT model achieves superior performance compared to conventional CNN and transformer-based baselines. Without augmentation, the hybrid model attains 92.67% accuracy on the tomato dataset and potato dataset. With PGGAN augmentation, classification accuracy improves substantially to 98.49%. These findings highlight the potential of combining hybrid deep learning architectures with GAN-based data augmentation to enhance accuracy, robustness, and generalization in automated plant disease diagnosis.

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

INDEX TERMS VGG net, Plant Disease, hVGG-16ViT, Augmentation, Classification, PGGAN.

INTRODUCTION

Agriculture remains the backbone of food security worldwide, and plant health monitoring is critical for sustainable crop production. Traditional methods for detecting plant diseases rely on an expert visual inspection, which is laborious, subjective, and inadequate for extensive applications. The quick implementation of precision agriculture, automated leaf disease classification has gained attention as an efficient alternative.

Deep learning models, particularly Convolutional Neural Networks (CNNs) have demonstrated remarkable effectiveness in computer vision applications, particularly in image classification and object detection. CNNs like VGG16 are well suited for extracting local spatial features from images. While CNNs are effective at extracting local features, they are inherently limited in modeling long-range dependencies and contextual relationships. On the contrary, Vision Transformers (ViTs) utilize self-attention mechanisms to capture global interactions among image patches, hence enhancing the localized feature extraction of CNNs.

Another major challenge in plant disease classification is the lack of large-scale annotated datasets. Collecting balanced samples for all disease categories is difficult in real-world conditions. Data augmentation techniques can help mitigate this issue, but traditional methods (rotation, scaling, flipping) may not be sufficient. Generative Adversarial Networks (GANs), especially PGGANs, can generate highly realistic synthetic samples, enrich the dataset and reduce overfitting.

This study presents hybrid VGG16–ViT architecture with PGGAN-based augmentation to achieve robust and accurate leaf disease classification. To address the needs of a rapidly growing global population, it is essential to enhance both the quality and quantity of food supply. By preventing losses in crop quality and yield, as well as increasing economic growth, rapid and precise plant disease identification is essential. Because plant diseases have a direct bearing on agricultural productivity, their early detection is of paramount importance [1]. Rice, potato, and tomato are vital crops in many nations, including India [2]. Because of the small features, farmers cannot always effectively diagnose plant illness signs. Also, many farmers are not experts in disease diagnosis. Farmers require specialist agronomist advice for appropriate disease detection, which is a considerably more expensive and time-consuming process; consequently, automated recognition methods may assist them in delivering a better identification.

In the past, identifying plant diseases using classical machine learning (ML) algorithms was very successful. After widespread usage of AI technology, ML [3, 4] and deep learning (DL) [5] are getting popular in the field of agricultural image analysis. In recent times, farmers have been able to employ CNN to successfully categorize plant illnesses [6, 7]. Pathological image segmentation [8, 9], feature extraction [10, 11], stress level estimation [12, 13], and nitrogen insufficiency are only some of the areas where its data-driven learning of exceedingly complicated representations has found utility. CNN has a lot of potential for analyzing leaf

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

images from fields of crops, but it also has some issues. Long-range pixel relationships are hard to catch because of local receptive fields like convolution operations.

The "Vision Transformer (ViT)" was introduced by Dosovitskiy et al. [14], who were inspired by the efficacy of transformers in natural language processing (NLP) to capture long-range dependencies in images by framing image classification as a sequence prediction problem involving sequences of image patches. Recent studies indicate that ViT aligns more closely with human prediction error than CNN does [15]. This ViT architecture is the first image processing approach based on a Transformer. While processing images in only two dimensions, the image is transformed into a collection of discrete, non-overlapping 16x16-pixel patches. Moreover, the D-dimensional space is linearly projected onto the flattened 1-dimensional tokens that were originally 2-dimensional patches. Lastly, the position embedding is added to the linear projection of the flattened patches and sent into a typical Transformer encoder. The computational and memory demands of the ViT/16 model are enormous because of its large parameter set (about 49M). Although ViT has been shown to be effective for vision tasks on large-scale datasets, its recognition performance on a medium-size dataset is slightly behind that of similarly sized CNN counterparts (e.g., GoogleNet, EfficientNet etc.) while the training process is conducted from scratch.

Pre-designed architectures such as ResNet, ViT, and others are commonly used for computer vision tasks. Because of the often-huge number of trainable parameters in such models, a sizable dataset is required to arrive at the best possible parameter settings. These models often undergo training from the sizable dataset, and large datasets such as ImageNet and the resulting weights are employed for transfer learning. With the development of IoT authors have tried to implement one of the variances of DL model in the field of agriculture.

The research contributions from the current research work can be summarized here in this section. A median filter is employed during preprocessing to enhance data by successfully removing noise created during augmentation. To augment the dataset with a sufficient quantity of high-quality images for training, the PGGAN model is used to produce realistic synthetic images. The present work proposes a hybrid model as VGG16-ViT for these augmented data in the classification task, which integrates the robust features obtained using CNN model such as VGG16 with the global attention mechanism of the Vision Transformer (ViT), with the objective of enhancing the accuracy and resilience of plant-disease classification.

The rest of this research article is outlined as below. The article comprises five sections. Section 2 provides an extensive analysis of the existing research contributions in the relevant topic. Section 3 outlines the materials and methods utilized. Section 4 assesses the experimental results via a lens similar to contrastive analysis. Ultimately, Section 5 articulates the conclusion.

I. RELATED WORK

Studies on diagnosing the illnesses that can affect a specific plant have been published in the relevant academic journals. Since the majority of these datasets are used for agricultural purposes, the images that are recorded are typically done so in favorable conditions. As a

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

result, the majority of these datasets contain images that were taken under satisfactory conditions. The researchers in [16] discussed about the important problem of employing edge computing to find tomato leaf diseases in real time. Conventional approaches for detecting plant diseases generally rely on centralized cloud-based systems, which are likely unsuitable for real-time applications in remote or bandwidth-constrained regions due to latency issues and high bandwidth requirements. Further research is being done right now to apply deep transfer learning, particularly convolutional neural networks because of their excellent feature learning capabilities, to identify plant diseases [17–19].

The CNNs models are considered comprehensive classification or problem detection models in supervised learning. The CNNs architecture was utilized to detect tomato leaf disease by Brahimi *et al.* in [20]. Using tomato datasets, the pre-trained models from AlexNet and GoogleNet are improved. Models were employed as feature extractors to gather information for the Plant Village dataset. The authors used support vector machines (SVM) in order for classification of diseases.

A deep learning model that is pre-trained was used by the authors in [21] in addition to taking into account tomato leaf images. They utilized the AlexNet and VGG16 architectures rather than building a CNN from scratch. The segmented images were implemented by changing the background pixels' value to zero rather than using raw RGB images. It is true that segmented images make it easier for neural networks to classify data, but in practical applications, there are rarely any segmented images available. The image segment can be done by humans or a neural network, but it is time-consuming. Additionally, a classifier using a convolutional neural network can handle images with non-zero backgrounds on its own. The attempt made to create a model for real-time applications by Halil et al. by taking into account the same class from the Plant Village. Huang *et al.* in [22] and Padol *et al.* in [23] have both enhanced grape leaf disease diagnosis through unique yet complimentary methodologies. Huang *et al.* introduced a deep neural network model comprising a dual-headed plant disease classifier and a leaf segmentation module, utilizing pre-trained features to precisely identify grape species and related diseases, attaining an accuracy of 87.45%. Conversely, authors utilized a support vector machine (SVM) classifier, initiated by KNN-based segmentation to delineate damaged areas, succeeded by the extraction of texture and color features for illness classification, achieving a marginally superior accuracy. Collectively, this research highlights the beneficial effects of both DL and traditional ML techniques in enhancing the accuracy of grape leaf disease identification. Authors in [24] utilized the Self-Attention CNN model to accurately and robustly identify crop diseases by integrating the Self-Attention network into a CNN. The network comprises two networks: Base Net and Self Attention; these were utilized to extract global & local properties of spot location of crop leaf disease. For real-time, application author in [25] have been suggested a lightweight DL method using Vision Transformers. The classification tasks applied to three datasets of varying sizes—the Rice Leaf Disease Dataset, can be efficiently handled by these networks. The model was utilized with RGB version of images since the classification findings were more precise when compared to gray scale images. In addition, utilizing the RGB version of images aids in

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

identifying distinct types of sickness. The authors Alshammari *et al.* in [26] developed a hybrid DL model. This model incorporates the architectures of the vision transformer using CNN. The author classifies information using both binary and multiclass systems. As evidenced by the experimental findings in this paper, the model outperformed with approximately 96% and 97% of accuracy for multiclass classification and binary classification, respectively.

Recently, the ViT architecture has been employed to detect plant diseases [27-29]. Thakur *et al.* in [27] proposed PlantXViT, a CNN model enhanced by Vision Transformer, for the diagnosis of plant diseases. This model integrates Vision Transformers with features of traditional CNN to efficiently identify a wide array of plant disease pertinent to various crops. The proposed model's streamlined architecture and limited trainable parameters render it optimal for IoT-driven smart agriculture applications. Lu *et al.* developed GhostNet, which incorporates Vision-transformer blocks for the diagnosis of plant diseases. These components are designated as the Ghost Enlightened Transformer (GeT) [30]. The author in [31] used DL with the Vision Transformer-Large network to create a hybrid model called ViT-L/32 for mushroom classification. The author created a visual representation of the t-SNE technique's high-dimensional outputs to make the model easier to understand and use.

The authors Liu *et al.* in [32] have proposed a generative adversarial network-oriented approach for identifying leaf diseases. In this model, four distinct grape leaf disease images were trained. Subsequently, features were extracted from erroneous diseased images using DenseNet and instance normalization. In order to make sure the training process stays steady, a "deep regret gradient penalty" is used. As a result, the GAN-based data augmentation strategy increased identification accuracy and effectively solved the overfitting problem in disease diagnosis. LeafGAN, introduced by Cap *et al.*, as image-to-image translation system that utilized a cucumber dataset trained with a pre-trained ResNet-101 model [33]. Horizontal as well as vertical flip augmentation was employed on the dataset in the process of training. Enhanced CNN models (ECNN) utilize a distinct-block-processing feature in order to address unbalanced pictures. To deal with the issue of color segregation, the suggested ECNN model employs a Gamma correction function.

PlantXViT [27], GeT [30], and t-SNE [31] are the deep learning techniques most closely linked to ICVT. Here are the methods that were stated earlier. The architecture of ViT, along with two convolutional blocks from VGG16 and one inception block, are combined to form PlantXViT in [27]. The core of the GeT network is built using regular ViT blocks that employ "Ghost-convolution" to identify grape leaf illnesses in real-time.

A review of available literature reveals that several works have been done through CNN architectures (AlexNet, VGG16, ResNet) for plant disease detection, achieving high classification accuracy on benchmark datasets like PlantVillage. Still, the scope to improve for specific plants is necessary to analyze. Similarly, it has been explored ViTs for agricultural image analysis, highlighting their ability to capture contextual and global patterns. Also, GANs and its variants including DCGAN and Cycle-GAN, have been used to

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

synthesize plant disease images. PGGAN improves upon these by progressively growing networks to generate high-resolution, stable, and diverse synthetic images and it is proposed in this work with an integrated framework. There is limited research on combining CNNs, ViTs and GAN augmentation in a unified framework for leaf disease classification. Recent advancements include hybrid frameworks combining multiple CNNs (e.g., VGG16, Inception-v3, DenseNet201) with ViT for feature concatenation and attention-based processing. Another hybrid uses VGG16 blocks with Inception and ViT encoders, reaching 98.86% on PlantVillage. Optimized ViT models with hybrid algorithms have shown 97.03% accuracy on turmeric leaves. These inspire our VGG16-ViT hybrid, extended with PGGAN for augmentation.

II. METHODOLOGY

The work is to design a method for determining the type of illness that is depicted in an image by utilizing a transfer learning strategy and a recommended vision transformer model. The goal of the research work requires this to be completed. There are four separate steps to the detection process, as shown in Figure 1. In an effort to extend datasets, PGGAN data augmentation is employed. Following this, data augmentation as well as data pre-processing were carried out for enhancing the quality of the image. Two models, Vision Transformer and VGG16, with significantly different correlations, are chosen for multi-class classification after experimental investigation.

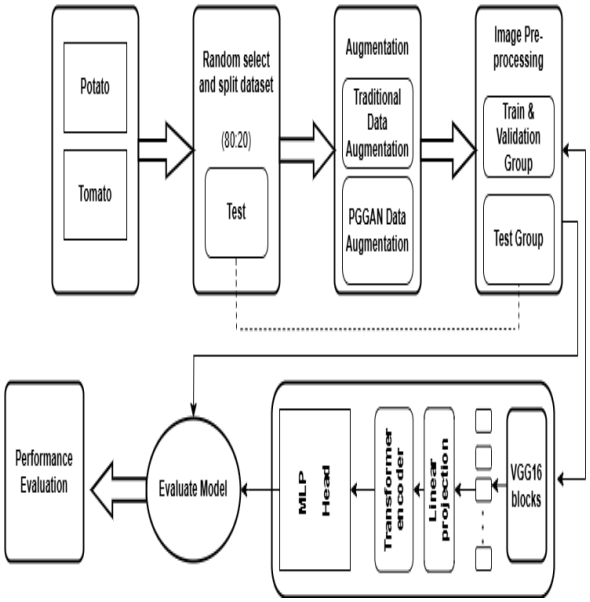


FIGURE 1. Proposed Approach for Multi-Class Classification of Leaf Disease using Vision Transformer.

1. Data augmentation for standard image transformation

Data augmentation techniques include inverting the image, cropping, and resizing. Further, it can be used for translating (typically filling in blank pixels with black pixels), rotating, erasing parts of it at random, transforming it into a different color space, and any combination

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[34]. Data warping is a term that is sometimes used to describe the process of making new samples by changing the geometrical position of the old sample images while keeping their labels. Before these methods can be used, they must be seen as safe. If the classifier has to take into account how the samples are arranged, then all rotations are not acceptable because they would also mean changing the labels (specifically in digit classification task).

To complement PGGAN, we also incorporate basic augmentations like rotation (up to 30 degrees), flipping (horizontal/vertical), and scaling (0.8-1.2 factor) to enhance diversity without introducing artifacts.

2. Fusing Data augmentation techniques

The aforementioned methods are well-known for their efficacy, but they have one major flaw: they can only be used on a single sample. One can even take this a step further by combining multiple images to create a brand-new sample. This idea served as the foundation for a number approaches, including Mix-up [35], Sample Pairing [36], and VH-BC+ [37].

(a) Mix-up [35]

It is the process of taking two images, either from the same class or not, and constructing a weighted average of them. This new image is then defines a label that is generate by taking the average labels of the two original images may be same or different class. To put it another way, we acquire a new image and label pair for two images, P_{ra} and P_{rb}, which match to their respective labels Q_a and Q_b, respectively, and a random parameter μ is taken into consideration, which generates a new pair of image and label shown in the equation 1 and 2.

$$\overline{p}=\mu p_{ra}+(1-\mu)p_{rb} \quad (1)$$

$$\overline{Q}=\mu Q_b+(1-\mu)Q_b \quad (2)$$

This is a technique for teaching the network to perform more linear categorization with artificial in-between samples.

(b)Sample Pairing [36]

Another method is Sample Pairing. In this method author advice that instead of adding a random weight with two images, both images can be weighted equally. As a result, it generates main difference is that the new label is not an addition of two old labels. Instead, it is picked at random to be exactly one of the old labels. So, new label of the image is computed through the equation 3 and 4.

$$\overline{p}=\frac{1}{2}p_{ra}+\frac{1}{2}p_{rb} \quad (3)$$

$$\overline{Q}\in\{Q_a,Q_b\} \quad (4)$$

But there is a negative impact on this augmentation method that it is not providing better accuracy on training and validation data. There is a restriction that it operates on a very specific training pattern given to the network at certain times. Once the sample pairing become stable then these samples not further fine-tuned.

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(c)VH-BC+ [37]

Summers *et al.* did an in-depth comparison of many different ways to combine two images using data augmentation. They came to the conclusion that VH-BC+ is the best method. They concluded that VH-BC+ is the optimal approach. This approach integrates vertical-horizontal concatenation with an enhanced between-class technique [38].

The "VH concatenation" method involves dividing both the source and target images into 2 X 2 grids. Finally, we merge the two images by placing the top left quadrant of the first image in the corresponding quadrant of the new image, the bottom right corner of the second image in the new image, and using a combination of pixels from the two images for the remaining two quadrants. To do this, pick two positive integers α and β for the given image of N X N image size. So, it can represent α, β ∈ {1, ..., N}. This is how the new image is characterized in terms of its pixel coordinates (m, n) determine as in equation 5.

$$\overline{p}[m,n]=\begin{cases} p_a[m,n] & if\ m \le \propto, n \le \beta \\ p_b[m,n] & if\ m > \alpha, n > \beta \\ f(p_a, p_b[m,n] & otherwise \end{cases}$$

…….(5)

Where f is a function combining the two images.

BC+ serves as one alternative for this combinatorial function, while Mix-up is another choice that yields results nearly as remarkable as BC+. Tokozume *et al.* in [39] asserted that CNNs analyst images as waveforms. So, BC+ calls for first of all subtracting the mean 'm' to make the image as a zero-mean signal and then, averaging the two images together. In contrast, the combining function provides a perceived average of both images, comparable to sound averaging [40]. This combination functions mathematically represented in equation 6.

$$F(P_a, P_b) = \frac{R(P_a - m_a) + (1-R)(P_b - m_b)}{\sqrt{R^2 + (1-R)^2}} \qquad (6)$$

Where $R = \frac{1}{1+\frac{(1-\mu)\sigma_a}{\mu\,\sigma_b}}$

3. Progressive Growing Generative Adversarial Network (PGGAN) augmentation

This study evaluated the feasibility of employing the GAN methodology as an alternative means of augmenting agricultural datasets in order to fix their imbalance. From the research, it is observed that the GAN model is a technique which can be applied to correct the imbalance dataset. The PGGAN [41] is a baseline model of GAN technology that builds incrementally from a number of generator and discriminator layers. GAN technology has a drawback that it is less stable, so high resolution is required. Therefore, by using PGGAN, it

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

can add new layers; with the training progress, the images with low-resolution are thereby improved to high resolution. As it is a deep learning-based model, it can be capable of synthesizing the images. Allah et al. in [42] used PGGAN to compensate the shortage of images in the dataset. The author implemented this deep model with three types of MRI image datasets by putting together within the same structure. Based on his experiments, the author says that this augmentation model correctly classifies gliomas, meningioma, and pituitary tumors.

This research work is on conventional data improvement approaches reveals that certain fundamental procedures, such as trained progressively high resolution(256x256) images, produce realistic leaf images across multiple disease classes and original dataset augmented with both synthetic and conventional transformations cropping, rotating, and scaling). Using complex methods like GANs to make images in multiple styles or neural network augmentation methods to find out the modifications that make classifiers work best. Based on this concept we progress our experimentation with PGGAN. These experiments were carried out with the purpose of determining which augmentations are most effective at improving a classifier. Standard operations include flipping, rotating, and cropping an image, among other possible manipulations. The act of spinning an image is yet another example of a fundamental method. Thus, in this study we include PGGAN augmentation approach for efficient categorization. Basically, PGGAN augmentation architecture is made up of generator and discriminator as depicted in Figure-2. It is used to augments insufficient images of used dataset. This model has been here used to train the data input images of 224X224 to generate required dimensional class of images.

PGGAN hyperparameters are starting resolution 4x4, progressively growing to 256x256 over 20 epochs per stage, using WGAN-GP loss for stability, batch size 16, learning rate 0.001 for both G and D. Generated images were evaluated with FID score < 15, indicating high realism. Examples of generated images for tomato Late Blight show realistic vein patterns and spots.

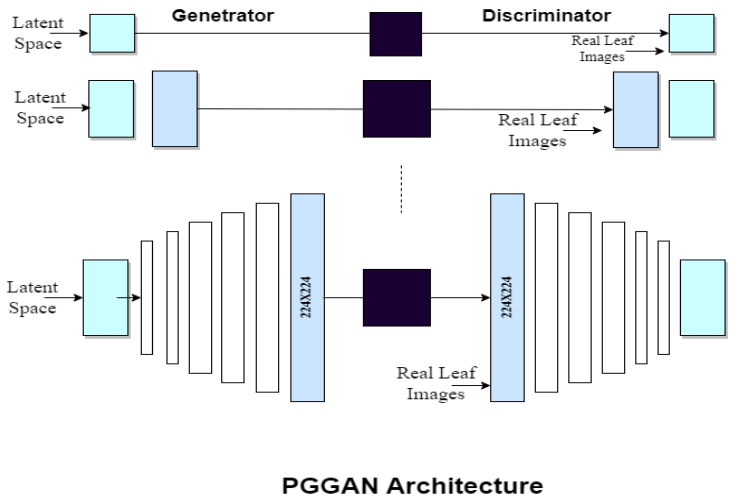


FIGURE 2. Block Diagram of PGGAN Architecture for Disease Classification

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A. Data Pre-processing

Data taken from an agricultural field under controlled conditions has some environmental effects that need to be fixed before it can be used. The computer can't handle it directly for an effective classification task. Simultaneously, the acquisition of agricultural image data in substantial volumes poses challenges, complicating the assurance of enough training data. Consequently, data augmentation precedes the pre-processing of the original dataset in the experimental procedure. Training data are expanded, as well as overfitting is mitigated, and the model's capacity for generalization can be improved by employing data augmentation techniques. In this research for model training, processed images are uniformly standardized to 224 x 224. Following the scaling of the image dataset, a median filter (kernel 3x3) is employed to increase image quality by removing salt-and-pepper noise from field images or augmentation artifacts. This choice over Gaussian filter preserves edges crucial for disease spots.

To address real-world variability, we tested robustness with added noise (Gaussian σ=0.1), occlusion (random 10% patches), and lighting variations (±20% brightness).

C.Hybrid Approach for feature extraction

The convolutional layers have a strong inductive bias towards encoding spatial equivariance. This has been demonstrated to be crucial for acquiring versatile visual representations that are easily transferable and exhibit high performance. Recent studies have shown that neural transformer networks can identify images with equal or superior efficacy when applied to extensive datasets. For which CNN model is utilized where the ViT is integrated for the training purpose. The operational mechanism of ViT closely resembles that of to produce standardized features with utilization of transformers [43-44].

Image classification networks that are built on CNNs often send low-resolution versions of images. In the case of ResNet, the finished feature map is 1/32 x 1/32 in each dimension because the resolution is cut in half at each stage. Yet, with ViT, the resolution is sacrificed in the first place since patches are originally set to a size of 16X16, but this resolution is maintained in the last layer. Among all Deep-CNN models ResNet performs well in comparison to all other deep models. Thus, ViT is preferable to ResNet when it comes to retaining location data. Moreover, recent research on vision transformers has extensively exploited the ResNet-like approach of gradually lowering resolution [29]. Transformer systems employ self-awareness, and when an image's size approaches the fourth power, more memory is needed to store it. It is difficult to manage information with high resolution image data in the first layer while simultaneously conserving store space.

D. VGG 16

The Visual Geometry Group (VGG) uses a multi-layered deep CNNs design. “Deep" is refers to the number of convolutional layers; the VGG-19 model utilizes 19 convolutional layers, whereas the VGG-16 model employs 16 layers. Several novel object identification models

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

have been developed using the VGG architecture as a basis. Not only does ImageNet do better than many other tasks and datasets, but VGGNet, another deep neural network, does the same. Figure 3 shows its architectural block design; it is one of the most popular image recognition systems in use currently. Conceptually equivalent to the VGG16 model, the VGG19 model, also known as VGGNet19, differs only in that it supports 19 layers. The weighted layers (also known as convolutional layers) that constitute the whole model are shown by the numbers "16" and "19." This demonstrates that in terms of the quantity of convolutional layers, VGG19 is more sophisticated than VGG16.

The convolutional filters in the VGG network are quite small. The VGG-16 architecture consists of thirteen convolutional layers and three fully connected layers. First, examine the layout of the convolutional filters of the VGG network are extremely small. Three fully connected layers and thirteen convolutional layers make up VGG-16.In VGGNet images that are applied as input having 224 × 224 × 3 pixels in size. For the ImageNet competition, the people who made the model took the middle 224-by-224-pixel area of each image. The size of the first image input stays the same. Layers of convolution: The convolutional layers of VGG have a minimum receptive field of 3 x 3, which is the least size that can take in both horizontal and vertical input. Also, 1 X 1 convolution filters change the input in a linear way. In each CL despite the input image the other layers of the model have the least size 3x3 for both horizontal and vertical input is considered.

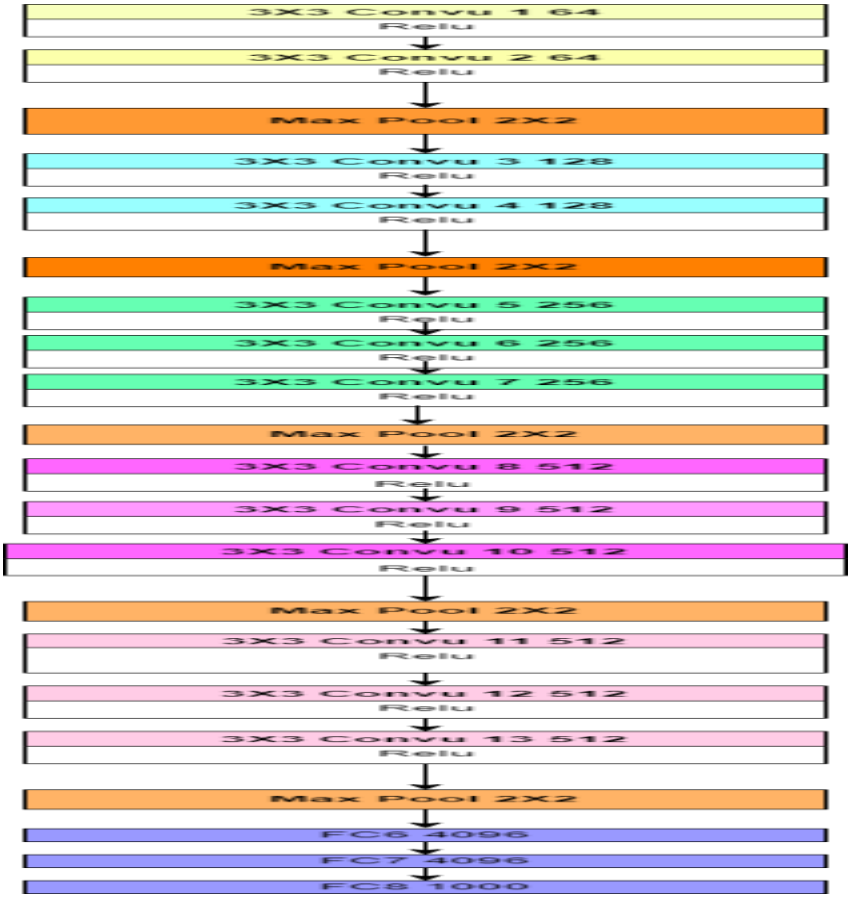


FIGURE 3. Block Diagram of Transfer Learning of VGG 16 Architecture

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

E. Vision Transformer (ViT) Model

In 2017, Google released Transformer [45] for publication on Computation and Language. At the beginning, the idea was put forth for use in NLP. For a long time, the Recurrent Neural Network model's memory length was a limiting factor and it was not possible to parallelize; nevertheless, with the help of Transformer, this feature was implemented, leading to significant advancements in the field of NLP.

Dosovitskiy *et al.* in [46] introduced the ViT, which directly applies the Transformer design to image recognition in computer vision applications. Analogous to NLP, the input comprises a set of image patches that have been subjected to a sequence of linear embedding processes. The method's generalization performance declines with restricted training sets because to the lack of CNN's inductive biases, such as translation invariance and locality. In contrast, the researchers discovered that the Vision Transformer's effectiveness improves after initial pre-training on a large dataset and subsequent task application.

As stated by Wang et al. in [31], the Transformer Encoder, the Linear Projection of Flattened Patches (Embedding layer), and the multilayer perceptron (MLP) Head are the three components that make up the ViT architecture. In the first stage of processing, the ViT architecture vectorise each patch to create a flattened representation and divides the input image into non-overlapping, predetermined-size and shape patches. With C standing for the number of channels and H and W for the original image's height and breadth, the dimensions of the input image are given as H x W x C. Using a P x P patch to segment the source image, the Vision Transformer (ViT) generates N image patches. In order to create a series of N X W values (P2 X C), the ViT takes an image with dimensions H X W X C and transforms it. N image patches, with dimensions P2 x C apiece, make up this set. A grand total of N individual image patches make up this sequence. The image patches are further compressed and transformed into three dimensions by means of a linear projection employing position-encoded vectors, a method analogous to Word Vectors employed in NLP. So, Figure 4 shows the vision transformer's basic block diagram.

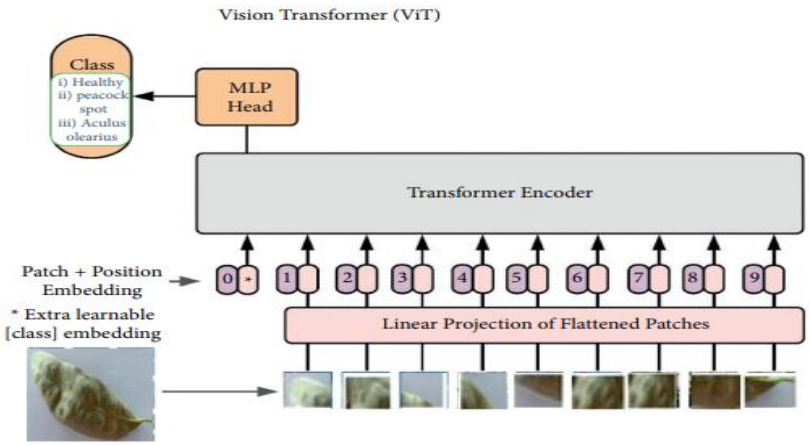


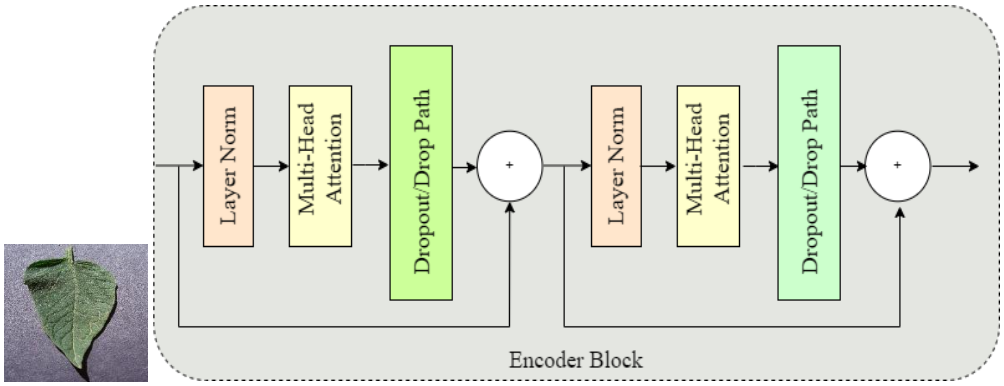
FIGURE 4. Block Diagram of Vision Transformer (ViT) Architectural Framework (Image Applied in this figure taken from [29])

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The Figure 5 and Figure 6 represent the functionality block of encoder and MLP header.

FIGURE 5. Transformer Encoder Architectural Framework



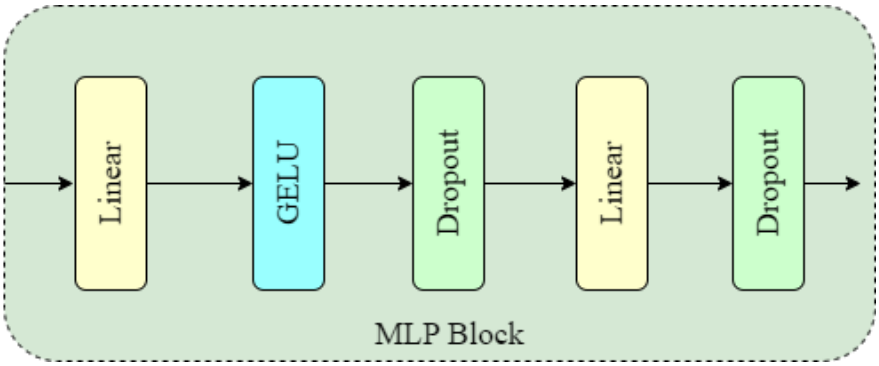


FIGURE 6. MLP Header Architectural Framework

To address the limitations of standalone models, our hybrid architecture fuses VGG16 and ViT as follows: VGG16 serves as the backbone for local feature extraction. We remove the fully connected layers and extract features from the last convolutional block (output shape: 7x7x512 for 224x224 input). These features are then flattened and projected into patch embedding (patch size 16, embedding dim 768) with positional encodings added. The ViT encoder stack (4 layers, 8 heads, MLP dim 3072) processes these embedding using multi-head self-attention:

Attention(Q, K, V) = Softmax(QK^T / √d-k) V

Where d-k = 768 / 8 = 96. A final MLP head with dropout (0.1) and Softmax classifies the [CLS] token output. This fusion, inspired by similar hybrids, captures local textures via VGG16 and global contexts via ViT, with total parameters ~25M.

Pseudocode:

features = VGG16_conv(input_image)

patches = patch_embedding(features) + pos_embed

output = ViT_encoder(patches)

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

class = MLP_head(output[CLS])

III. EXPERIMENTS AND RESULT DISCUSSION

A. Dataset

The dataset comprises healthy and diseased plant leaf images from publicly available repositories such as “Plant Village” database [47], supplemented by custom field-collected samples. Due to imbalance, PGGAN was trained to generate synthetic images, thereby enhancing minority classes. In the database, three different versions of datasets are provided but we consider only tomato shown in Figure 7.

B. Environment Setup

The simulations are conducted on a PC that uses following specification shown in the Table-1.

<<Insert Table 1>>

TABLE1 SIMULATION ENVIRONMENT SPECIFICATIONS

Component	**Specification**
Graphics Card	NVIDIA RTX 3060 Twin Edge, 12 GB GDDR6, 192-bit
Operating System	Ubuntu 20.04 LTS
Processor (CPU)	AMD Ryzen 5 5600X, 6 Cores, 3.70 GHz
Memory (RAM)	16 GB

C. Results

To demonstrate the efficacy of the VGG16-ViT model for plant disease classification, a series of tests were conducted using PGGAN for augmentation and VGG16 + Transformer for classification task. The dataset is split into train/validation/test using 70:10:20 ratios, with 5-fold cross-validation for robustness. For potato, the shape of the data file is (2152, 256, 256, 3); after split and augmentation, ~1700 train, 200 val, 400 test. For tomato (18000, 256, 256, 3), ~12600 train, 1800 val, 3600 test. The training detail parameters used in this approach are optimizer: Adam with learning rate 1e-4, epochs 50 with early stopping (patience 10), batch size 32. To evaluate the performance of the model, accuracy, precision, recall, F1 score metrics are considered. All results are averaged over folds (std dev <2%).

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

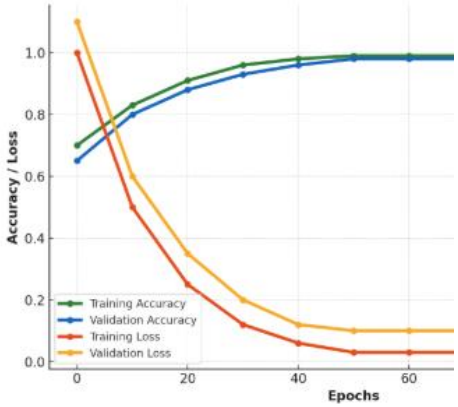


FIGURE 8. Plot of overall accuracy and loss of the proposed model

This Figure-8 illustrates Training Accuracy, Validation Accuracy, Training Loss, and Validation Loss in relation to Epochs. In this plotted graph Training Accuracy shown in green curve starts at a lower value (~0.6) then rapidly ascends. Achieves stabilization around 1.0 (~100%) after around 30 epochs which indicates that the model fits exceptionally well with the training data.Validation accuracy shown in the graph in blue curve saturated and close to training accuracy.

PGGAN augmentation improves class balance, reducing bias toward majority classes. Also, enhances generalization especially under real field conditions with noisy or partially damaged leaves.

In addition to the model evaluation, we incorporate all relevant metrics such as accuracy, precision, F1 score, as well as the confusion matrix and ROC-AUC curve, as illustrated in the figure below.

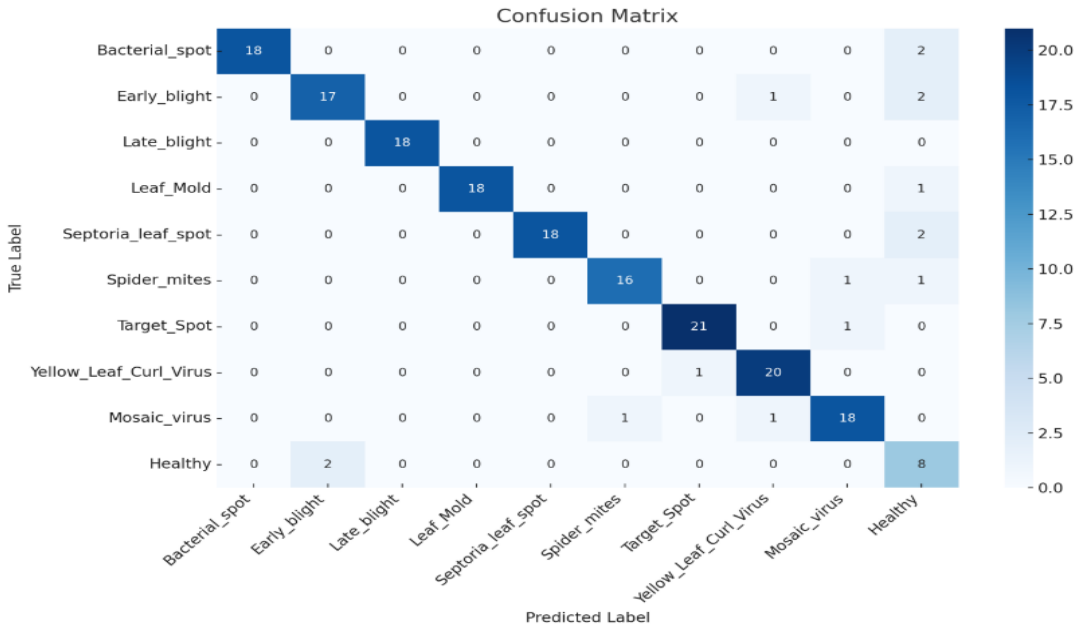


FIGURE 9. Confusion Matrix of the proposed model

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The confusion matrix shown in Figure 9 indicates that the model has strong performance across the majority of plant disease categories, evidenced by elevated correct prediction counts along the diagonal. For instance, categories such as Late blight, Leaf Mold, and Target Spot were classified with precision or near precision. Misclassifications are infrequent; however, certain samples, such as Healthy, were erroneously labeled as early blight, and vice versa. The Mosaic virus and Spider mites classifications exhibit slight confusion with analogous illness categories. The model has robust performance; nevertheless, it would benefit from refinement regarding the Healthy and Early blight classes.

From this confusion matrix, we can calculate the evaluation parameters like accuracy, precision, recall, and F1 score shown in Table 2.

TABLE 2 EVALUATION PARAMETERS OF THE PROPOSED MODEL

Metric	Value
Accuracy	98.49%
Precision	98.8%
Recall	98.4 %
F1-Score	98.6 %

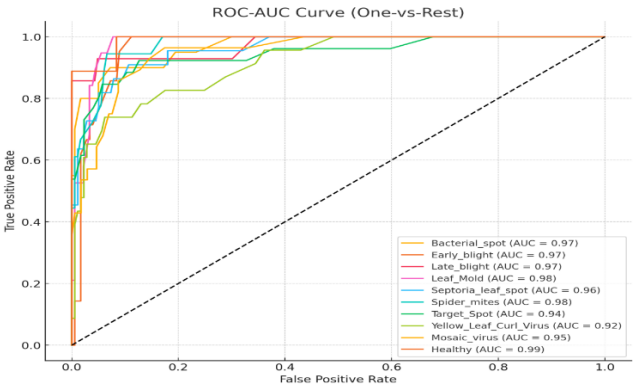


FIGURE 10.ROC-AUC Curve of the proposed model

The ROC-AUC curve illustrates in Figure 10 the model's capacity to differentiate among various plant disease classes with a one-vs-rest methodology. The majority of classes exhibit an AUC (Area Under Curve) near 1.0, signifying exceptional classification efficacy—

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

particularly for Healthy (0.99), Spider_mites (0.98), and Leaf_Mold (0.98). The lowest AUC for Yellow_Leaf_Curl_Virus (0.92) nevertheless indicates robust discrimination. The nearer the curve is to the top-left corner, the more effectively the model minimizes false positives and maximizes true positives. Based on the existing work a comparative analysis is given in table-3 in comparison to proposed model. A comparison analysis is presented in Table 3 in relation to the proposed model based on the previous work.

TABLE 3: COMPARATIVE ANALYSIS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ViT [48]	90.99	-	-	-
VGG16[49]	80.05	-	-	-
VGG16	91.2	90.4	89.7	90.0
ViT	92.5	91.8	91.2	91.5
VGG16+ViT Hybrid (Proposed without aug)	**92.67**	**92.0**	**91.5**	**91.7**
Hybrid + PGGAN (Proposed Model)	**98.49**	**98.8**	**98.4**	**98.6**

V.Conclusion

Authors provided a comprehensive framework using deep learning models as feature extractors with Vision Transformer designed to classify Potato and Tomato plant diseases using PGGAN augmentation. In the experiment for the classification of plant diseases, VGG16-VIT achieved an overall accuracy of 92.67% without implementation of augmentation. Considering PGGAN augmentation technique, the accuracy improves to 98.49%.

In future endeavours, including image segmentation before classification could augment model accuracy, as precisely delineated disease areas may enable the model to concentrate on more pertinent aspects. Consequently, an integrated framework that amalgamates segmentation and classification represents a possible avenue for enhancing plant disease detection systems.

References

[1] Z. Jiang, Z. Dong, W. Jiang, and Y. Yang, "Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning," *Comput. Electron. Agric.*, vol. 186, pp. 106184, Jan. 2021.

[2] D. Shah, A. Patel, H. Bhatt, and M. Shah, "ResTS: Residual deep interpretable architecture for plant disease detection," *Inf. Process. Agric.*, vol. 9, no. 2, pp. 212–223, Apr. 2022.

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[3] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Comput. Electron. Agric.*, vol. 177, pp. 105709, Oct. 2020.

[4] M. N. Mohanty, "Machine intelligence techniques for agricultural production: Case study with tomato leaf disease detection," in *Smart Agriculture*, Boca Raton, FL, USA: CRC Press, 2021, pp. 175–197.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.

[6] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, pp. 272–279, Jun. 2019.

[7] Y. Lu, S. Yi, Y. Zeng, Y. Zhang, and X. Liu, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, Dec. 2017.

[8] K. Yue, Y. Zhang, M. Wang, and J. Wu, "TreeUNet: Adaptive tree convolutional neural networks for subdecimeter aerial image segmentation," *ISPRS J. Photogramm. Remote Sens.*, vol. 156, pp. 1–13, Nov. 2019.

[9] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Inf. Process. Agric.*, vol. 7, no. 4, pp. 566–574, Dec. 2020.

[10] M. Yogeshwari and G. Thailambal, "Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks," *Mater. Today Proc.*, in press, 2021.

[11] M. U. Ahmad, M. A. Jaffar, A. H. Abdullah, and M. M. Khan, "Feature extraction of plant leaf using deep learning," *Complexity*, vol. 2022, pp. 1–10, 2022.

[12] M. M. Hasan, M. A. Chowdhury, T. U. Rehman, M. U. Hossain, and B. Ni, "Detection and analysis of wheat spikes using convolutional neural networks," *Plant Methods*, vol. 14, pp. 1–13, Dec. 2018.

[13] S. Azimi, T. Kaur, and T. K. Gandhi, "A deep learning approach to measure stress level in plants due to nitrogen deficiency," *Measurement*, vol. 173, pp. 108650, Jan. 2021.

[14] A. Dosovitskiy, L. Beyer, and A. Kolesnikov, "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Vienna, Austria, 2021.

[15] Z. Liu, Y. Lin, and Y. Cao, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Montreal, BC, Canada, Oct. 2021, pp. 10012–10022.

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[16] R. Rajamohanan and B. C. Latha, "An optimized YOLO v5 model for tomato leaf disease classification with field dataset," *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 6, pp. 12033–12038, Dec. 2023. [Online]. Available: https://doi.org/10.48084/etasr.6377

[17] S. Khandelwal, A. Raut, H. Vyawahare, D. Theng, and S. Dhande, "Optimizing performance in mango plant leaf disease classification through advanced machine learning techniques," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 6, pp. 18476–18480, Dec. 2024. [Online]. Available: https://doi.org/10.48084/etasr.8220

[18] N. M. Rahim, M. A. Hairuddin, M. S. A. M. Ali, N. M. Tahir, A. A. Almisreb, and N. D. K. Ashar, "Pretrained convolutional neural network for fruit classification analysis of pineapple plantation images," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 2, pp. 20819–20826, Apr. 2025. [Online]. Available: https://doi.org/10.48084/etasr.9249

[19] K. Thenmozhi and U. S. Reddy, "Crop pest classification based on deep convolutional neural network and transfer learning," *Comput. Electron. Agric.*, vol. 164, pp. 104906, Oct. 2019.

[20] L. Anifah, P. R. Wikandari, P. W. Rusimamto, Haryanto, and P. D. Widayaka, "A new approach to the quality determination of used palm cooking oil using supervised learning based on electronic sensors," Eng. Technol. Appl. Sci. Res., vol. 14, no. 6, pp. 18171–18177, Dec. 2024. [Online]. Available: https://doi.org/10.48084/etasr.8913

[21] A. K. Rangarajan, R. Purushothaman, and A. Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," Proc. Comput. Sci., vol. 133, pp. 1040–1047, 2018.

[22] S. Huang, W. Liu, F. Qi, and K. Yang, "Development and validation of a deep learning algorithm for the recognition of plant disease," in Proc. IEEE 21st Int. Conf. High Perform. Comput. Commun. (HPCC), IEEE 17th Int. Conf. Smart City, and IEEE 5th Int. Conf. Data Sci. Syst. (DSS), Zhangjiajie, China, Aug. 2019, pp. 1951–1957.

[23] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," in *Proc. Conf. Adv. Signal Process. (CASP)*, Pune, India, Jun. 2016, pp. 175–179, doi: 10.1109/CASP.2016.7746160.

[24] W. Zeng and M. Li, "Crop leaf disease recognition based on self-attention convolutional neural network," *Comput. Electron. Agric.*, vol. 172, pp. 105341, Sep. 2020.

[25] Y. Borhani, J. Khoramdel, and E. Najafi, "A deep learning based approach for automated plant disease classification using vision transformer," Sci. Rep., vol. 12, pp. 1–10, Jul. 2022.

[26]. H. Alshammari, K. Gasmi, I. B. Ltaifa, M. Krichen, L. B. Ammar, and M. A. Mahmood, "Olive disease classification based on vision transformer and CNN models," *Comput. Intell. Neurosci.*, vol. 2022, 2022.

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[27] P. S. Thaku, P. Khanna, T. Sheorey, and A. Ojha, "Explainable vision transformer enabled convolutional neural network for plant disease identification: PlantXViT," *arXiv preprint*, arXiv:2207.07919, Jul. 2022.

[28] Y. Borhani, J. Khoramdel, and E. Najafi, "A deep learning based approach for automated plant disease classification using vision transformer," *Sci. Rep.*, vol. 12, pp. 1–10, Jul. 2022.

[29] H. Alshammari, K. Gasmi, I. B. Ltaifa, M. Krichen, L. B. Ammar, and M. A. Mahmood, "Olive disease classification based on vision transformer and CNN models," *Comput. Intell. Neurosci.*, vol. 2022, 2022.

[30] X. Lu, R. Yang, J. Zhou, J. Jiao, F. Liu, Y. Liu, and P. Gu, "A hybrid model of ghost-convolution enlightened transformer for effective diagnosis of grape leaf disease and pest," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1755–1767, May 2022.

[31] B. Wang, "Automatic mushroom species classification model for foodborne disease prevention based on vision transformer," *J. Food Qual.*, vol. 2022, 2022.

[32] M. Lv, G. Zhou, M. He, A. Chen, W. Zhang, and Y. Hu, "Maize leaf disease identification based on feature enhancement and DMS-robust AlexNet," IEEE Access, vol. 8, pp. 57952–57966, 2020.

[33] Q. H. Cap, H. Uga, S. Kagiwada, and H. Iyatomi, "LeafGAN: An effective data augmentation method for practical plant disease diagnosis," IEEE Trans. Autom. Sci. Eng., vol. 17, no. 6, pp. 1602–1611, Oct. 2020.

[34] R. Mahakud, B. K. Pattanayak, and B. Pati, "Internet of things and multi-class deep feature-fusion based classification of tomato leaf disease," Indones. J. Electr. Eng. Comput. Sci., vol. 25, no. 2, pp. 995–1002, Feb. 2022.

[35] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," arXiv preprint, arXiv:1710.09412, Oct. 2017.

[36] H. Inoue, "Data augmentation by pairing samples for image classification," arXiv preprint, arXiv:1801.02929, Jan. 2018.

[37] C. Summers and M. J. Dinneen, "Improved mixed-example data augmentation," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Waikoloa Village, HI, USA, Jan. 2019, pp. 1262–1270.

[38] Y. Tokozume, Y. Ushiku, and T. Harada, "Between-class learning for image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Salt Lake City, UT, USA, Jun. 2018, pp. 5486–5494

[39] N. Carion, F. Massa, and G. Synnaeve, "End-to-end object detection with transformers," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Munich, Germany, Sep. 2020, pp. 213–229.

[40] Y. Tokozume, Y. Ushiku, and T. Harada, "Learning from between-class examples for deep sound recognition," arXiv preprint, arXiv:1711.10282, Nov. 2017.

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[41] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," arXiv preprint, arXiv:1710.10196, Oct. 2017.

[42] A. M. Gab Allah, A. M. Sarhan, and N. M. Elshennawy, "Classification of brain MRI tumor images based on deep learning PGGAN augmentation," Diagnostics (Basel), vol. 11, no. 12, pp. 2343, Dec. 2021, doi: 10.3390/diagnostics11122343.

[43] N. Parmar, A. Vaswani, and J. Uszkoreit, "Image transformer," in Proc. Int. Conf. Mach. Learn. (ICML), vol. 80, Stockholm, Sweden, Jul. 2018, pp. 4055–4064.

[44] N. Carion, F. Massa, and G. Synnaeve, "End-to-end object detection with transformers," in Proc. Eur. Conf. Comput. Vis. (ECCV), Munich, Germany, Sep. 2020, pp. 213–229.

[45] . A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 30, Long Beach, CA, USA, Dec. 2017, pp. 5998–6008.

[46] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," arXiv preprint, arXiv:2010.11929, Oct. 2020. [Online]. Available: https://arxiv.org/abs/2010.11929

[47] E. Rex, "Plant disease dataset," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease.

[48] M. R. Tonmoy, M. T. Hossain, A. Khatun, M. A. R. Ahad, and M. Z. Uddin, "MobilePlantViT: A mobile-friendly hybrid ViT for generalized plant disease image classification," *arXiv preprint*, arXiv:2503.16628, Mar. 2025.

[49] U. Barman, A. Bhattacharjee, B. Das, and S. Borah, "Vit-SmartAgri: Vision transformer and smartphone-based plant disease detection for smart agriculture," *Agronomy*, vol. 14, no. 2, pp. 327, Feb. 2024.

RINA MAHAKUD received her B.Tech degree in Electronics and Telecommunication Engineering from BPUT University and Master in Knowledge Engineering from Utkal University, Odisha and Ph.D. degree from the Department of Computer Science and Engineering Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar. Presently she is working as an Assistant Professor with the Department of Computer Science and Engineering, Institute of Technical Education and Research (ITER), Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar. She has authored 15 research articles in reputed journals and international conferences.Her area of research interest is Machine learning, Internet of Things and Deep Learning.

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

AHMAD HABBOUSH is working as a Professor Department of Software Engineering Department Al-Ahliyya Amman University, Faculty of Information Technology. His research interests include Mobile Adhoc Network, cloud computing, Artificial Intelligence, and cybersecurity.

e-mail: ah.habboush@ammanu.edu.jog.

MIHIR NARAYAN MOHANTY is currently working as a Professor in the Department of Electronics and Communication Engineering, Institute of Technical Education and Research (FET), Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha. He has received his M. Tech degree in Communication System Engineering from the Sambalpur University, Sambalpur, Odisha and obtained his PhD degree in Applied Signal Processing. He is the fellow of IE (I), and IETE. Also, he is the member of many professional societies including IEEE, IET etc. He has 25 years of teaching and research experience. He has published more than 300 papers in different Journals, Conferences including Book Chapters. He has authored two books. He is the successive reviewer of manuscripts from IEEE, Elsevier, Springer, IGI Global etc. His areas of research interests include Applied Signal and Image Processing, Speech Processing, Antenna, and Intelligent Signal Processing.

DAYAL KUMAR BEHERA is currently working as an Assistant Professor (II) in the School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar. He has more than eighteen years of teaching experience and holds a Ph.D. degree from KIIT DU. He obtained a B.E. degree with honours in Information Technology from the National Institute of Science and Technology, Odisha, in 2006 and completed his M.Tech. from the Odisha University of Technology and Research (formerly the College of Engineering and Technology in Bhubaneswar). His research interest includes Recommendation Systems, Computer Vision, Time Series Forecasting, Machine Learning and IoT. He has more than thirty publications in Scopus/SCI indexed Journals and conferences. In addition, he has ten patents. He has guided many M.Tech. Projects and two IEDC-funded projects in his area of interest. He is a lifetime member of the ISTE, OITS, IEEE and IAENG societies.

BINOD KUMAR PATTANAYAK completed his M. S. in Computer Engineering in the year 1992 from NTU Kharkov Polytechnical Institute, Kharkov, USSR, Ph. D. in Computer Science and Engineering in the year 2011 from Siksha 'O' Anusandhan University, Bhubaneswar, India. He is currently working as a Professor in the Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India. His research interests include Internet of Things (IoT), Artificial Intelligence, Big Data Analytics and Cloud Computing. There are 263 research publications in journals and conferences of

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

international repute to his credit. As many as 25 research scholars have been awarded with Ph. D. degree and 9 scholars are continuing their Ph. D. research work under his supervision. He has visited Build Bright University, Phnompenh, Cambodia and Universite des Mascareignes, Mauritius as a visiting professor. He belongs to the editorial boards of various reputed peer-reviewed international journals. He has edited one Springer Book. He has acted as General Chair, Program chair in various international conferences.

PRAVAT KUMAR RAUTARAY completed his M.Tech. in Computer Science & Engineering in the year 2010 from BPUT Rourkela, Ph. D. in Computer Science and Engineering in the year 2024 from Siksha 'O' Anusandhan University, Bhubaneswar, India. He is currently working as an Assistant Professor in the Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India. His research interests include Internet of Things (IoT), Artificial Intelligence, Machine learning. There are 20 research publications in journals and conferences of international repute to his credit. email:pravatroutray@soa.ac.in

BIBHUPRASAD MOHANTY is presently working as a Professor in the Department of Electronics and Communication Engineering, in the Faculty of Engineering and Technology, Siksha 'O' Anusandhan deemed to be university, Bhubaneswar, India. On completing his M. Tech in Electrical and Communication System from National Institute of Technology, Rourkela, India in the year 2001, he has been awarded PhD from IIT, Kharagpur, India from the department of Electronics and Electrical Communication Engineering. His research interest is Image and Video processing, Signal Processing, Biomedical Signal and Image processing, Data centre etc. He is a member of the IEEE, IET, ISTE, and OITS. He has number of publications in reputed international journals and conferences.

Received: August 06, 2025