MAFH Based Maize Leaf Disease Detection and Classification

K. Amrutha
Vellore Institute of Technology University
P. Mithun Raj
Vellore Institute of Technology University
S Ananda Kumar
s.anandakumar@vit.ac.in

Vellore Institute of Technology University

Article

Keywords: Deep learning, Disease detection, Machine learning, MobileNetV3, Attention, Fusion, Accuracy.

Posted Date: September 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10497938/v1

Additional Declarations: No competing interests reported.

MAFH Based Maize Leaf Disease Detection and Classification

K. Amrutha [1], P. Mithun Raj [1], S. Ananda Kumar*
[1]School of Computer Science and Engineering (SCOPE),
Vellore Institute of Technology,
Vellore, Tamil Nadu, India
karnadevi.amrutha2022@vitstudent.ac.in mithunraj.p2022@vitstudent.ac.in
Corresponding Author: s.anandakumar@vit.ac.in

Abstract

Maize plant plays crucial role not only in the field of agriculture but also in global economy, since it is the third most cultivated crop across the globe. However, these plants are usually affected by various types of diseases such as blight, common rust, gray leaf spot, etc., protecting the plants from these disease is very important. This research proposes a new deep learning model for disease detection to perform well than other models. The hybrid model uses MobileNetV3 as the backbone architecture integrated with attention, fusion and head. The framework includes data preprocessing, feature extraction, classification and interpretation. We will compare the performance of this model with other models such as VGG16, ResNet50, DenseNet121, ALEXNET, etc. criteria for the final evaluation includes Accuracy, Precision, Recall, F1score, Specificity, Logloss, AUC-ROC curve. Through this proposed model we have achieved an accuracy of 98% which is high than the other models compared. The lightweight nature of MobileNetV3 enables us to implement the model in the mobile and IoT devices also. The present study contributes to the development of deep learning model in the field of agriculture, offering a efficient solution for early maize leaf disease detection.

Key words :
Deep learning, Disease detection, Machine learning, MobileNetV3, Attention, Fusion, Accuracy.

I. Introduction

Maize is one of the major and most cultivated crop across the world. It is the third highest cultivated crop after wheat and rice. Not only in animal husbandry, but it also plays a major role in the raw material industry for development of industrial products. However, maize usually suffer from the Northern Leaf Blight disease (NLB). Decrease in the yield of maize by northern leaf blight is steadily increases. Therefore, it is extremely important to find the disease. However, identifying NLB at early stage is not an easy task. Traditional method disease detection is by hand picking by agricultural experts and expertise in plant pathology. Misjudgment of disease may lead to loss of plant and inaccurate use of pesticides, which leads to the increase in the environmental pollution. Convolution Neural Networks is a popular method of target detection, especially in the field of crop detection. It is a Machine Learning(ML) algorithm which can achieve high accuracy by training large number of images. In this they use a NLB dataset which contains 3 parts. The first part is handset, which contains images taken by hand, the second part contains a boom set which is taken by mounting the camera on a 5-meter boom, the last part is unmanned drone dataset, which means taking image by mounting DJI drone. In this they have divided the dataset into 3 parts, training, validation and testing in the ratio of 5:4:1. The data used in this paper is taken from the field of high intensity light causing the reflection phenomenon. In this they have used single stage target detection algorithm (SSD). The final criteria they have chosen are Mean Average Precision and Frame per seconds. They have used data preprocessing method, feature fusion techniques, feature selection and disease detection techniques. The main use of data preprocessing is to reduce the effect of high light intensity on the leaf. They have proposed a new model with combination of 4 algorithms and the mAP of this new model is higher than the regular SSD at 91.83%. The new model is useful for the identification for the maize leaf disease, with high accuracy and efficiency, which could reduce the on-site identification by the human experts [1].

Diseases are the main and primary disaster for maize plant production, with annual loss at 6.10%. According to the statistics there are more than 80 types of maize disease. For this disease rapid accurate detection is very crucial to improve yields. However using the machine vision, it is complicated, because appearance of maize leaf, size, color, texture etc. vary between maize varieties and stages of growth. In this paper they have collected dataset from Science Part of China Agriculture University and Technical College of Inner Mongolia Agricultural University. They have used nearly 4428 images. They have used Gaussian based sampling method to generate the missing data. In this paper they have chosen PSNR and SSIM to evaluate the quality of processed images. Using this method they have generated nearly 89000 images. They have chosen PyTorch Framework to run the algorithm. Multi Activation Function(MAF) is fused to every algorithm like VGCC19, ResNet50, DenseNet161, GoogLeNet etc. The images are given as input to these algorithms in the 2 sets as training and testing dataset. Above all the algorithm MAF-ResNet50 was visualized to high accuracy than the other algorithms. They have reached an accuracy of about 97.4%. From the success of the above algorithm, they have developed mobile application for the iOS platform [2]. This paper present deep learning as a feasible approach in classification of disease in maize plant. They focus on underlying the specific function of all ANNs and there uses in the dataset. This paper contains multi-scale topology network of CNN. They created a multidisciplinary dependence connection with the third dimensional map of corn plant. They have taken the images from plant village dataset. To decrease the training time of dataset, it was scaled automatically with the help of OpenCV. Total of 7316 images are used for classification. They have done two experiments; first experiment uses many deep learning algorithms. They have used different deep learning algorithm "Inception-V3, Dense-Net-121, ResNet-101-V2, Xception". Among that Xception have validation accuracy of 95.08% and validation loss of 0.31%. In the second experiment they use CNN as the background model with different epochs. This model shows an accuracy of 0.93% at epoch 7 [3].

There are many techniques available for recognition using deep learning models, they have some problems among them such as large memory consumption, unsuitable for mobile deployment. In this experimental study they deploy it in mobile phones to identify the maize disease. They use Mosaic Data Enhancement, which takes an random images and slice the image at random position. It increases the data diversity in this they use YOLOv5 algorithm. They add coordinate attention mechanism to improve the accuracy. They have added this CA in the backbone of YOLOv5 keeping the parameters unchanged. In model training stage, the total number of training round set to 300 and the batch size is of 16. The YOLOv5 have high accuracy of 0.92% at 0.01 learning rate. They have compared the mAP of different mechanism. CA_YOLOv5 has high accuracy than others at 95.2%. this method has high accuracy and smaller size [4]. We have chosen the hybrid algorithm for the disease detection which gives an accuracy of ~98 % for 30 epochs. Instead of using a single algorithm we have used a combination four algorithms to obtain high accuracy. With the help of this algorithm, we can quickly identify the maize leaf disease.

II. Literature Review

There are two types of CNN based model algorithm: two-stage detection and one-stage detection using regression. R-NN comes under the two-staged detection which have high accuracy and it is far from the real-time in terms of speed. One stage algorithm is represented by the SSD and YOLO are good in real-time in terms of speed but far from the accuracy [5]. Agronomic Teacher is a semi-supervised leaf disease detection algorithm. This framework is based on the mutual relationship between teacher and student model. After several experiments

YOLOv5s is chosen as the baseline model. To enhance accuracy of this experiment they have proposed Agro-YOLO model. It consists of two parts Agro-Backbone network and Agro-Neck network. Both are integrated together to increase the overall performance [6]. First we took the datasets from the open source platform. When training the leaf disease identification algorithm we used Stochastic Gradient Descent (SDG) algorithm. It is most suitable for non-convex optimization problem, and the learning rate was set to 0.001 and the number of epochs are set to 800 [7]. Using Convolution Neural Networks Yang assessed maize seedlings by analyzing in the near visible IR region, over 400 samplings were used for this process. Zhen used RDNet with VGG16 by replacing global pooling layer with fully connected layer. Two authors used CNN which was applied to the different studies based on the supervised learning and achieved an accuracy of 82.19% & 94%. Other studies employed RDNet using unsupervised learning and achieved an accuracy of 84.04% [8]. All the papers that are working on the maize leaf disease detection have accuracy at most 94% – 96%. We are planning to have an accuracy greater than 96% along with the best suitable algorithm to use.

The datasets are collected from 3 different regions which include Arusha, Kilimanjaro, Manyara. Another CNN methods were introduced to analyze the images of healthy & unhealthy plants and a total of 87,000 images were used for this. Torch71 machine learning framework were used. The study aims to develop a combined deep learning model for detection of MSV and MLN based on the collection of images from the fields [9]. Image processing applications captures the minute features of crops helps in the growth stage detection and the yield estimation activities. Along with the image preprocessing techniques several ML algorithms such as SVM, fuzzy logic, Neural Networks (NN) and decision trees are used to address the various problems. Automation is highly possible with the integrated field, climate and environment related data available to intelligent crop management system [10]. The growth of maize crop is not only affected by the weed but also by the pathogens and by various pests. Pests include leaf beetle, red spider mite after damage caused by the beetle on the leaves, which appears to be small, irregular white spots [11]. Since the neural networks is suspectable for the small spots in the maize disease classification and the introduction of the attention is necessary for this. However current attention mechanism cannot understand the relationship and dependency on the data and hence resulting in the poor generalisation [12].

The transformer model is introduced in 2017 and has achieved remarkable success in the field of NLP and computer vision [13]. In disease detection task advantages of transformer are prominent. By introducing the transformer we can be able to handle the complex structures with different environments. YOLOv11 is one of the lightweight detection in YOLO series, with a good balance of detection speed and performance lays a solid foundation for rapid model development. In this paper they have used BCS_YOLOv11 which is good in feature extraction, feature fusion and overall output [14]. Wang et al. proposed a convolution neural networks model that integrates inception and residual along with Convolution Block Attention Module(CBAM), and the experimental results showed an accuracy of 95% disease for the plant village dataset. Performance of these models will decrease when used on the real time environment. To overcome this real time crop images are added to the plant village dataset taken from the internet [15]. Application on remote sensing technology are implemented in existing areas field-wide and the sub-fields remain limited. It is important to focus on sub-fields because most of them are small farmers. The study aims to develop the model to identify the leaf disease based on the RGB images at the sub-fields [16].

A recognition tool developed for Android from pre-trained data on convolution neural networks for recognizing crop disease. They have trained 3 deep CNN architecture for maize leaf disease detection and these architecture are VGG16, ResNet50 and AlexNet [17]. Attention mechanism are introduced in the CNN architecture making the model to focus more on the most important regions of the images and skipping the irrelevant data. This is more useful in the crop disease detection, where distinguish between the healthy leaf and diseased leaf. Vision Transformers have been emerged an effective solution for the model in both local and global feature through self- attention [18]. Existing attention based and residual CNN works well on general image classification and shows some limitations to the domain-specific problems. To overcome this we have propose the MaizeNet a light weight CNN model that incorporates channel wise-squeeze and excitation mechanism [19]. To address the poor performance of multi-scale detection faced by the RT-DETR model a new improved model is proposed. To reduce the parameter size and computational complexity the backbone network introduces RepVGG to improve feature extraction capabilities. This reduces the redundant computations and improves model performance [20].

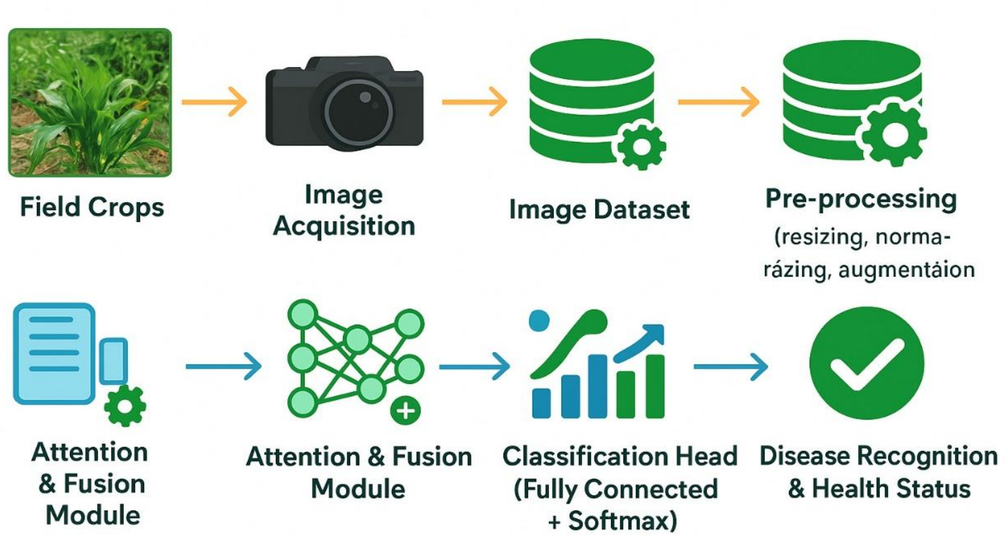


Fig 1. Proposed framework to recognition of the disease and health of maize leaf

III. Summary

YOLOv5 is the widely used YOLO algorithm model which has the advantage of fast detection and the strong generalization ability. YOLOv5 creates a lightweight model and MobileNetv3 is used as the backbone for this model. FRAFM employees CARAFE to up-sample the feature map and to preserve the details in the deep features. Maize leaf blight disease occurs in the small and dense area, so most critical information are lost during the image processing. To overcome this issue CIPAM is proposed. Frames Per Second (FPS) is introduced to evaluate the impact of multiple methods on detection speed. On comparing to other models the proposed one has least FPS [5]. YOLOv5 is chosen as a baseline model. Agronomic Teacher is a semi supervised leaf disease detection framework. They proposed a new model called AgroYOLO detector by combining the two, enhance the accuracy in the complex environments. It consists of two parts Agro-Backbone network and Agro-Neck network. They compared the results using labeled dataset at different portions in terms of mAP [6]. LFMNet is a deep learning network that identifies the maize leaf disease on the real time environment. The structure has two main modules PMFFM and Mattion. PMFFM extracts the multi-scale feature of maize leaf disease. Mattion has a point wise CNN layer that down samples the input map and 2 attention modules in the parallel. Proposed model is evaluated based on six evaluation metrics [7]. A deep learning model is taken for training in which VGG16 is taken and XAI are implemented to visualize it. XAI uses LRP, which uses back propagations algorithm. This XAI model show fair thinking beyond prediction. They have compared the results based on the evaluation metrics [8]. CNN is basic deep learning algorithm.

This model was developed with nearly 27000 images, and a Rectified Linear unit known as ReLU was used as an activation function. Vision Transformers was developed with a dataset consists of 6000 images. ViT is a new image recognition technology that makes use of transformer architecture. The evaluation metric choosed is accuracy and ViT achieved an accuracy of 93.10% and CNN with 90.96% [9]. We developed a multistage leaf disease classification model using ResNet50. This algorithm uses regularization parameters that do not lead to gradient decent loss, hence reducing overfitting. ReLU is used to handle nonlinear relationship among different features. They have taken 100 epochs with the batch size of 64. From this algorithm they have able to achieve an accuracy of 95% [10]. LFMNet is proposed based on the ResNex50 to achieve high accuracy. LFMNet combines LSMB, FLB and MLFFA. These helps to increase the accuracy and ensure superior identification. This model has multiple layers and the input traverse through all the layers. The number of iterations is 120 with batch size of 16. To evaluate the performance of the algorithm they have choose accuracy and loss values [11]. Integrated Multidimensional Attention is used to make more interactions and dependency between the elements. It contains two parts Feature extraction and Feature weight update. They have used the training with 150 epochs of batch size of 24. To evaluate the performance the three models they have chosen accuracy, and they have achieved 84% [12]. Constructed a multistage feature extraction and fusion model based on state-space tree. The model begins with the initial input pass through the feature extraction module and extract the features, later these features pass through the different layers in the state-space modules. For evaluation they have considered accuracy, precision and recall. For the maize northern leaf blight, they have achieved recall score of 0.99 [13]. Based on YOLOv11, we designed an image detection model to improve the performance in small target detection and complex backgrounds is BCS_YOLO. It divides the input into low frequency and high frequency, and self-attentions. To evaluate the model, we have used precision-recall curve to different models. The BCS_YOLO curve is outermost, stating that it has high precision levels across full recall region [14]. The framework of proposed model is Dy-Tri-ResNet50 for maize leaf disease classification. It is developed by integrating dynamic convolution and triple attention mechanism, into residual blocks. This allows the model to capture local features by preserving the global information. The models performance is evaluated using Precision, Recall and F1-Score. Comparing to the other models the proposed model has high precision, recall and F1score [15]. In this paper they have used both CNN and ResNet50 for the classification of leaf, both the models were trained over 100 epochs with the batch size of 64. Accuracy is used to evaluate the model’s performance. From both the models ResNet50 have high accuracy when compared to CNN [16]. Three pre trained architectures were taken AlexNet, VGG16, ResNet50. These models were combined and developed a mobile application. They evaluated the mobile application using accuracy. The application has high accuracy the other models [17]. MSCPNet model is composed of three essential components MobileNetv2, Multiscale feature aggregation and Pool Former. The overall approach of model encompasses data preparation, augmentation, normalization, training and testing. The overall performance is evaluated based on evaluation metrics and the accuracy they have obtained ~96% [18]. This paper covers the experiment done on the MaizeNet. They did experiment for five folds and taken their average and final evaluation metric they have taken is precision. From this model they have gained precision of 95% [19]. Introduced a lighter Quantization Aware RepVGG by redesigning the RT-DETR model. This model not only reduces the redundant computation, but also increases the models performance. To evaluate the model, they have taken mAP at threshold 0.5 and achieved 92% [20].

Advantages and Disadvantages

Model	Advantages	Disadvantages
CIPAM(MobileNetV3+ FRAM + CARAFE) [5]	Lightweight and fast for mobile devices.	Limited depth restricts performance on complex backgrounds.
AgroYOLO Detector (Agro Backbone + Agro Neck) [6]	High speed object detection for real time field applications.	Requires high-quality labeled bounding boxes.
LFMNet (PMFFM + Mattion) [7]	High accuracy for multi-disease classification	Longer training time due to multiple fusion blocks.
VGG16 + XAI [8]	Suitable for educational and diagnostic transparency.	Slower inference and lower accuracy than modern CNNs
ReLU + Vision Transformer [9]	Performs well on datasets with diverse patterns.	Computationally expensive for real-time task
ResNet50 + ReLu [10]	High accuracy on most leaf datasets.	High model complexity and memory use.

LFMNet (LSMB + FLB + MLFFA) [11]	Lightweight yet accurate for mobile applications.	Slightly slower training due to multiple attention stages
Intelligent Multi-Attention Model [12]	Suitable for mulit-disease classification.	Attention redundancy can cause overfitting.
BCS_YOLO [14]	Fast real-time detection with efficient gradient flow.	Detection quality decreases on low-resolution inputs.
Dy-Tri-ResNet50 [15]	Robust and high accuracy on complex datasets.	High computational Cost
MSCPNet (MobileNetV2+Multiscale feature aggregation + Pool Former) [18]	Excellent balance between speed and accuracy	Moderate accuracy drop on extremely fine-grained datasets.
MaizeNet [19]	Optimized specifically for maize leaf texture and color variation	May struggle under severe illumination changes.
Quantization-Aware RepVG [20]	Combines quantization for lightweight deployment	Requires fine-tuning quantization parameters.

IV. Materials and Methods

A. Data Source

The target dataset utilized in this experiment is maize leaf disease. As shown in the Fig 3. Among the various types of maize leaf diseases, the most common ones are Blight, Gray leaf spot and the Common Rust disease, the images of all the three leaf diseases and the normal leaf images are used in the dataset.

a) Healthy leaf

A healthy maize leaf is green in color, smooth, and free from any spots or discoloration. Leaf blades are flattened and broad. It shows uniform color and texture, indicating good plant growth and no other infections.

b) Blight leaf

Leaf Blight is a fungal disease. It causes lesions which are large, irregular brown or grayish spots on the leaf. The affected areas may dry out and spread quickly, reducing the leaf's ability to photosynthesize.

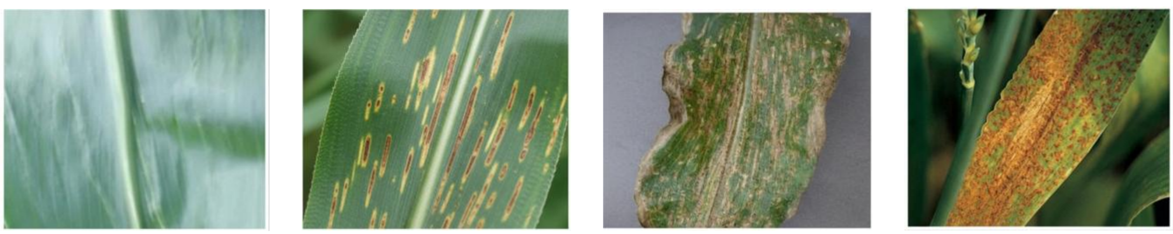
c) Gray-leaf spot

It is destructive fungal disease of maize caused by the Cercospora zeae-maydis and Cercospora Zeina that appears as small, rectangular gray or tan lesions between the leaf veins. As it spreads, the spots merge and can make the leaf dry and brittle.

d) Common Rust

It is also a type of fungal disease which is caused by the Puccinia sorghi, characterized by small, reddish-brown, oval pustules on both the upper and lower leaf surfaces. Over time, these pustules can turn black and reduce the plant's strength.

The dataset of maize leaf diseased images which are selected and that are used in this paper are taken from the open source website called Kaggle and accessed on 31st July 2025 (https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset). The total number of images in the dataset are 4188. The three diseased leaves and healthy leaf are shown in the below Fig 3. The number of maize leaf datasets that are used in the model are shown in the Table 1.



(a) (b) (c) (d)

Fig 2. Examples of maize leaf diseases. (a) Healthy leaf; (b) Blight leaf; (c) Gray-leaf spot; (d) Common Rust disease.

Table 1. Number of maize leaf dataset

Leaf Types	Number of Images
Healthy Leaf	1162
Blight Leaf	1146
Gray- leaf spot	574
Common Rust	1306

B. Dataset splitting

The dataset used in our study contains 4188 images of four different categories. These images are to be used for both training and testing the model. The dataset was organized into 2 folders, training data and testing data. Training data is used for model learning and testing data is used for evaluating the model performance. Additionally , a WeightedRandomSampler is applied to the training dataset to balance class frequencies.

C. Image preprocessing

To improve the model's robustness, an extensive data augmentation techniques were applied using the torchvision.transforms library. They prevent overfitting and improve recognition under various conditions. Only resizing and normalization are applied to valid data to ensure unbiased evaluation.

D. Proposed Model

We have proposed a hybrid deep learning framework named MAFH. It is the combination of four models(MobileNetV3,Attention, Fusion, Head). MobileNetV3 is used as the backbone of our model, and it is pre-trained on ImageNet. Attention Module enhances important spatial and channel-wise features while suppressing irrelevant information. Fusion Layer combines multi-scale features from the different layers to capture fine and coarse details of leaf diseases. Classification Head is a global average pooling layer followed by dropout and a fully connected layer for the final disease prediction. We take the batch size of 32 and epochs of size 30.

a) MobileNetV3

It is a lightweight and efficient deep CNN architecture designed for the mobile and vision applications, build upon MobileNetV2. It was developed by Google research as an improvement over MobilNetV1 and MobileNetV2. It integrates several advanced techniques such as depth-wise separable convolutions and the squeeze-and-excitation (SE) modules**,** hard swish activation functions. The improvements include reduced computational complexity, increased accuracy, and lower latency. It contains two versions MobileNetV3-Large and MobileNetV3-Small. Due to its lightweight and low computational requirement, it is well suited for real time environment, and on mobile devices because it doesn't require high processing power.

b) Attention Module

It is a powerful mechanism incorporated in the deep learning models to enhance their ability to focus on the most important information within the image. In our scenario it helps the network to focus more on the diseased regions on the leaf – such as lesions, spots etc. and minimizing the noises such as soil, healthy part of leaf, and lighting. It is inspired from the human vision perception, where our eyes focus more on the important parts of the scene. This model learns to assign higher weights to most significant features and lower weights to low significant features. It has different types of mechanisms such as channel mechanism and spatial mechanism; we combined these two and form Convolution Block Attention Modules (CBAM) or Self-Attention Mechanism. The module becomes more robust to various lighting, noise, and complex leaf texture resulting in better performance.

c) Fusion Layer

A fusion layer in deep learning is a component that combines the two or more inputs. The inputs are combined using the common methods such as concatenation or addition. Its main purpose is to integrate complementary information from various layers such as low-level, mid-level, high-level into a single more informative feature map. Convolution layer captures the fine details, and the fusion layer combines these layers simultaneously. Fusion can be combined in different ways such as feature concatenation, element-wise-addition and weighted average according to the architecture. By merging the features effectively fusion can increase the robustness and accuracy of the disease classification model.

d) Classification Head

It is the final stage of deep learning model responsible for interpreting the extracted feature. After MobileNetV3, Attention and Fusion processed the input images, and these are passed into the classification head. Classification Head ensures that the complex visual patterns identified by the earlier network components are effectively translated into clear, interpretable diagnostic outcomes. A well-designed classification head enhances model accuracy, confidence, and stability

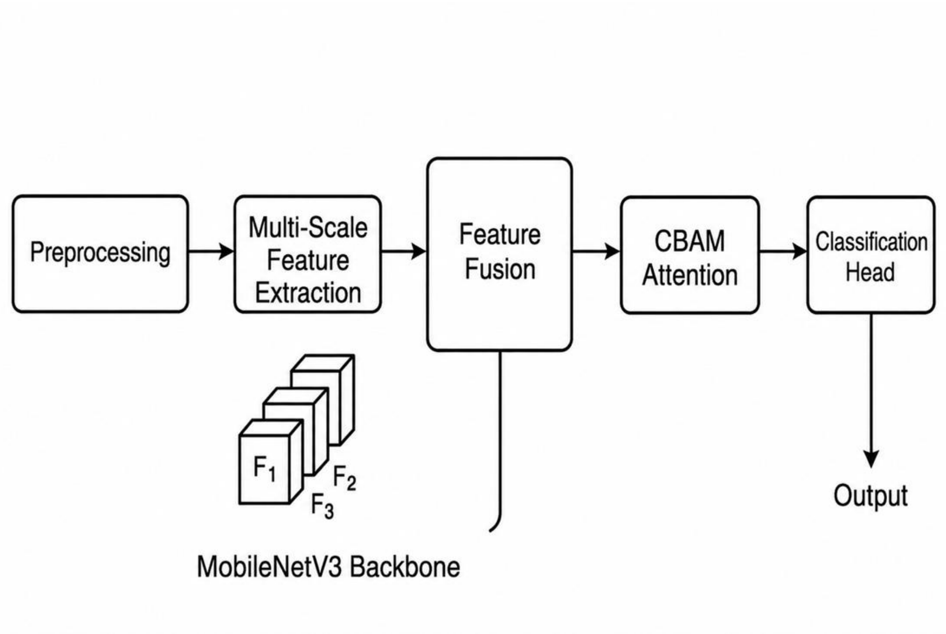


Fig 3. Architecture of our hybrid model

V. Survey Table

Name of the firsts author	List Methods/Protocols/models	Parameter 1	Parameter2	Parameter 3	QoS	Simulator	Year
[5]	YOLO + FRAFM + CIPAM + MobileBit	Precision	mAP	FPS	Yes	PyTorch	2023
[6]	AgroYOLO(Agro backbone + Agro Neck)	Precision	mAP	FLOPs	Yes	PyTorch	2024

[7]	LFMNet (PMFFM + Mattion)	Accuracy	Precision	Recall	Yes	Python 3.8, PyTorch	2024
[8]	VGG16 + XAI	Accuracy	Misclassification rate	No	Yes	TensorFlow	2024
[9]	CNN + ViT	Accuracy	Loss	No	No	TensorFlow	2024
[10]	ResNet50 + ReLU	Precision	Recall	F1score	Yes	TensorFlow, keras	2024
[11]	LFMNet (LMSB + FLB + MLFFA)	Accuracy	Precision	Recall	Yes	PyTorch	2024
[12]	ICPNet (IMA + CDC)	Accuracy	Precision	Recall	Yes	PyTorch	2024
[13]	Multi-stage Feature Extraction	Precision	Recall	Accuracy	Yes	TensorFlow, PyTorch	2024
[14]	BCS_YOLO	Precision	Recall	F1score	Yes	PyTorch	2025
[15]	Dy-Tri-ResNet50	Precision	Recall	F1score	Yes	PyTorch	2025
[16]	ResNet50, CNN	Accuracy	Precision	Recall	Yes	No	2025
[17]	AlexNet, VGG16, ResNet50	Accuracy	Precision	Recall	Yes	Google Colab	2025
[18]	MSCPNet (MobileNetV2 +Multi scale Convolution + Pool Former)	Accuracy	Precision	Recall	Yes	PyTorch, Torchvision, Seaborn, OpenCV.	2025
[19]	MaizeNet	Precision	Recall	F1score	Yes	Python's keras library	2025
[20]	RT-DETR + QARepVGG	Precision	mAP	FPS	Yes	Not Given	2025

Problem Definition

Maize is one of the most cultivated crops across the globe and it serves as the food, food and industrial raw materials. However, the crop production is mostly affected by various diseases such as Northern Leaf Blight, Gray Leaf spot and Common Rust. These disease will significantly reduce the crop production and if not identified and treated in time it will cause a lot of damage to the production. Traditional crop disease identification mostly depends on the farmers, agriculture experts and experts in plant pathologists which is mostly time consuming, subjective and error-prone to large fields or at early-stage infection. In many developing areas, farmers are not having access to the plant pathologist and modern diagnostic tools, leads to the late or misclassification of diseases. Further most environmental factors like lighting, weather conditions, and leaf alignment make the classification of the disease more challenging to detect. Hence there is a requirement for automated, accurate and efficient maize leaf disease detection that assist farmers in real time.

There are a lot of methodologies or models proposed to identify the maize leaf disease detection. The most used one is CNN, ResNet50, YOLO algorithms. CNN are the most basic and commonly used image processing algorithm; it is used to handle complex and diverse maize leaf images, it requires a large dataset, and the computational cost is high. ResNet50 is widely used pretrained CNN algorithm; the 50 indicates the no of layers in the network and the images bypasses through all the layers and enable the model to deeper representation. It requires a long training time, and the performance is low in noisy and uneven lighting. YOLO series are able to cope with the meeting various conditions in industrial aspects which has led to widespread attention. It is widely used in recent years, and it has many variants with adjusting the width and depth of various models. It struggles to find the small lesions on the leaf and requires a well classified dataset. These models have been deployed by many people to various datasets and have achieved an accuracy among the range 90%-96%.

We have proposed a new hybrid model with a combination of four models to overcome limitation of the above models. Our model comprises of MobileNet, Attention, Fusion and Head. This model achieves high validation accuracy, demonstrating strong generalization across all disease classification. It successfully distinguished between visually similar disease such as gray leaf spot and blight. It achieves high accuracy with lower FLOPs and ideal for mobile edge devices, performs well on complex backgrounds, make sure it uses all the images in the dataset equally. Due to its efficient and small size, it can be deployed in mobile devices, drones and IoT devices as well.

Example

Let us consider a farmer in the remote village, who finds a brown spots on the leaf and he is not sure whether it is Gray Leaf Spot or Common Rust. The farmer uses his smart phone and takes the photo of the infected leaf and uploads it on the maize disease detection system. The system pre-processes the images and passes it through all the models and classifies the disease as Gray Leaf Spot with 98% accuracy. The system will also give a small description about the disease and the potential control methods. This helps farmers to identify the type of disease that affected the leaf and appropriate preventive steps to minimize the crop loss and to increase the yield production.

IV Procedures

Data acquisition

We collect the maize leaf images containing four classes :

Healthy leaf, Blight leaf, Common Rust, Gray Leaf Spot

The dataset is divided into two parts of training set and testing sets in the ratio of 80:20.

The divided images are stored in the train/ and val/ classes.

Image preprocessing

Images are resized to 300 X 300 for uniform input.

Applying data augmentation such as Random rotation, cropping , brightness adjustment.

This helps the model to learn under various conditions and perform well.

Dataset loading

The dataset is loaded using PyTorch framework

A WeightedRandomSampler is used to balance class frequency ensuring equal learning across all the classes.

Model architecture

MobileNetV3 : Extracts initial features efficiently using depth wise separable convolutions.

Feature Fusion : Combines multi-scale feature maps from different layers to capture both local and global information.

Attention Module : Enhances important spatial and channel features while suppressing irrelevant background details.

Classification Head : Applies global average pooling, dropout, and a connected layer to predict the disease class.

Model Training

The model is trained using AdamW Optimizer.

The cosine learning rate adjusts the learning dynamically.

The training continues for up to 30 epochs.

Validation and performance evaluation

The model is evaluated on validation set after each epoch.

The metrics includes:

- Accuracy
- Precision
- Recall
- F1-Score
- Specificity
- MSE
- LogLoss
- AUC-ROC

Visualization

After training confusion matrix and graphs are generated.

Accuracy and Loss curves are plotted to visualize the performance trends.

Algorithm

Algorithm 1: Maize Leaf Disease Recognition Using MAFH (MobileNefV3 + Attention + Fusion + Head)

---

1. **Image Preprocessing**
   Input: maize leaf imagesdataset $D = \{I_1, I_2, ., ., , I_n\}$
   Output: Predicted class label C (e. Healthy, Blight, Rht, Leaf Spot, and confidence score
2. **Step 1 — Image Preprocessing**
   Resize each image I_i to fixe d dimensions (e: 300)
   Normalize pixel intensities using ImageNet mean ad sid.
   Apply data augmentation (rotation, flip, brightness, contrast.
   Split dateset into training and validation subsets
3. **Step 2 — Feature Extraction** (MobileNetV3 Backont)
   Input: preprocessed maize leaf image I_i
   Extract feature maps from low, rniddle, ana aeep ıayers
   Concatenate all features — $F_{fused} = [F_1, F_2, F_3]$
4. **Step 3 — Feature Fusion and Attention Module**
   Input: Attention-enhanced features F_{attn}
   Optimize with AdamW optimizer and cosıncheduer.
   MixUp and Label Smoothing for robust training
   Apply early stopping based on validation accuracy
5. **Step 5 — Model Training**
   Define loss $L = \text{CrossEntropy}(y_{true}, y_{pred})$
   Optimize with AdamW optimizer and cosine scheau.
   Use MixUp and Label Smoothing for robust training
   Apply early stopping based on validation accuracy
6. **Step 6 — Evaluation**
   Compute Accuracy, Precessn, Precision Recall, and F1–score
   Plot confusion matrix and performance graphs
7. **Prediction** For a new maize leaf image I_{test} . perfoı
   Output: Disease class C and health status (Healthy / Diseased

The above image is the algorithm for the proposed model. It consists of eight steps. Maize leaf is taken as input from the dataset. The input image is resized to required dimensions, preprocessing techniques are applied to the image, AdamW optimizer is used to optimize the image. After applying the preprocessing and feature extraction modules, the model is trained based on training dataset. After successful training of data, the model has been evaluated based on testing dataset. The model's performance is evaluated using the standard evaluation metrics. Finally, we predict output of the image either the leaf is diseased or healthy.

Flow diagram – proposed method

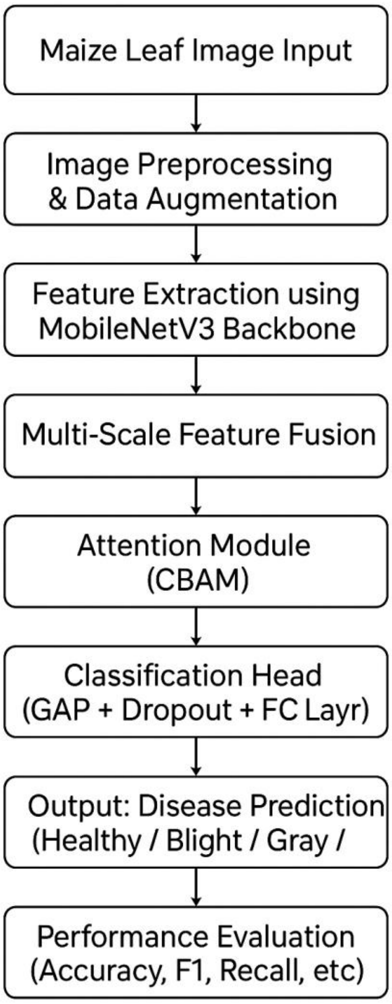


Fig 4. Flowchart of our proposed model

V Implementation

The proposed model MAFH for the maize leaf disease detection is implemented using PyTorch. The model was trained and tested for the datasets containing healthy leaf, blight leaf, common rust leaf and gray leaf spot. We have divided the dataset into training and testing folders. The training folder is used for the training of the model and testing folder is used for testing and evaluation of the model based on the performance and they are divided in the ratio of 80:20. Each image is resized into 300 X 300 pixels. Figure 5 talks about the data preprocessing, datasets.ImageFolder reads the images randomly from the folders. TRAIN_DIR and TEST_DIR is used to store the data. It scans the folder, loads the images and their classes and applies the transformation to each image.

```
train_ds = datasets.ImageFolder(TRAIN_DIR, transform=train_tf)
val_ds   = datasets.ImageFolder(VAL_DIR,   transform=val_tf)
NUM_CLASSES = len(train_ds.classes)
```

Fig 5. Code for Data Preprocessing

After loading the dataset, we transform it, because the images in the dataset may not be same. The images may be in different sizes, shapes, texture and color. Figure 6. talks about the data transform, it first resizes the images into same size and changes the color contrast, brightness and normalize the data.

```
train_tf = T.Compose([
    T.RandomResizedCrop(IMG_SIZE, scale=(0.6, 1.0)),
    T.RandomHorizontalFlip(),
    T.RandomVerticalFlip(p=0.1),
    T.RandomRotation(20),
    T.ColorJitter(brightness=0.25, contrast=0.2, saturation=0.2, hue=0.02),
    T.RandomAutocontrast(p=0.2),
    T.ToTensor(),
    T.Normalize(mean, std),
])
```

Fig 6. Dataset transformation

MAFH is proposed. It is a combination of four models, and it uses MobileNetV3 as the backbone for our model. It is one of the basic deep learning models.

```
class MAFH_MobileNetV3(nn.Module):
    def __init__(self, num_classes, tap_ids=None, fuse_dim=128):
        super().__init__()
        m =
models.mobilenet_v3_large(weights=models.MobileNet_V3_Large_Weights.IMAGENE
T1K_V1)
        self.stem = nn.Sequential(m.features)   # nn.Sequential of blocks

        # choose tap points (indices) if not provided
        L = len(self.stem)
        if tap_ids is None:
            # take three progressively deeper layers
            tap_ids = [L//3, 2*L//3, L-1]
        self.tap_ids = sorted(set([max(0, min(L-1, t)) for t in tap_ids]))
```

Fig 7. MobileNetV3 model

After loading the model, we split the dataset and apply the training dataset to the model. We divide the dataset into two halves training part into 80% and the testing part into 20%. After training phase we deploy testing phase and the model’s performance is evaluated based on the testing phase, and we chose common evaluation metrics to evaluate the performance.

```
acc = accuracy_score(y_true, y_pred)
prec = precision_score(y_true, y_pred, average="weighted")
rec  = recall_score(y_true, y_pred, average="weighted")
f1   = f1_score(y_true, y_pred, average="weighted")
spec, cm = specificity_multiclass(y_true, y_pred, NUM_CLASSES)
mse  = mean_squared_error(y_true, y_pred)
ll   = log_loss(y_true, y_prob)
try:
    auc = roc_auc_score(y_true, y_prob, multi_class="ovr")
except:
    auc = None
```

Fig 8. Evaluation metrics to calculate the performance of model

VI Result and Discussion

Table 2. Experimental parameters information

Specific Parameters	Values
No. of Training Images	3029
No. of Testing Images	757
Image Size	300 X 300
Learning rate	$3 \text{ X } 10^{-4}$
Optimizer Selection	AdamW Optimizer
Weight Decay factor	$1 \text{ X } 10^{-4}$
Batch size	32
Training Epochs	30
Loss Function	0.05
Learning rate adjustment strategy	Cosine annealing with warm-up
Attention Mechanism	CBAM
Framework used	PyTorch
Compiler	Google Colab

Experimental configuration

To ensure the consistency in results, the whole experiment is conducted in the same hardware and software configuration. The device used for this experiment is HP Pavilion Ryzen includes an AMD Tyzen 5 5500U (Hexa-core/Octa-core, up to 4.0 GHz) and a 7 Cores integrated AMD Radeon Graphics, with 8 GB DDR4 and Windows 11. The versions of the python do not affect the result of the experiment. We deployed Torchvision, NumPy, scikit-learn, Matplotlib, Seaborn on PyTorch 1.12 and Google Colab with specific parameters settings shown in the Table 2.

Model evaluation indicators

The key performance indicators such as Precision, Accuracy, Recall, F1-Score, Specificity, MSE, LogLoss, AUC-ROC curve are used to evaluate the performance of the proposed model.

Accuracy tells the ratio of correctly predicted samples to the total number of samples taken. It tells how many images the model classifies correctly. It helps to find the overall effectiveness of the model. Precision tells how many samples that are predicted positive are correct.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

In the above equation True Positive(TP) indicates the no of diseased leaves correctly found as diseased. True Negative(TN) indicates the no of healthy leaves correctly found as healthy. False Positive(FP) indicates the no of healthy leaves incorrectly found as diseased leaves. False Negative(FN) indicates the diseased leaves incorrectly found as healthy.

Recall can measure how well the model can find the actual positive samples. The ratio of true positive to actual positives. F1-Score is the combination of Precision and Recall. It gives a balanced measure of model's performance in the unbalanced data.

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = 2\ X\ \frac{Precision\ X\ Recall}{Precision\ +\ Recall}$$

Specificity measures how well the model can correctly measure the data. It tells how many leaves are correctly classified as healthy by the model. Mean Squared Error (MSE) is a metric that measures the average squared difference between predicted and actual value in the dataset. It is mainly used when the model outputs the continuous values.

$$Specificity = \frac{TN}{TN + FP}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}i)^2$$

Log Loss measures how uncertain or confident our model predictions are It evaluates the probability scores output by a classification model. The value ranges between 0 and 1. 1 for the diseased and 0 for healthy. AUC-ROC stands for Receiver Operating Characteristic-Area Under Curve. It measures the model's ability to distinguish between the classes. ROC plots true positive and false negative, AUC summarizes the curve.

$$Log\ Loss = -\frac{1}{N}\sum_{i=1}^{N}\left[yi\log(pi) + (1 - yi)\log(1 - pi)\right]$$

$$AUC = \int_{0}^{1} TPR(FPR)d(FPR)$$

Table 3. Comparison of performance of different models

Model	Accuracy	Precision	Recall	F1-Score	Specificity	MSE	LogLoss	AUC-ROC
VGG16	0.95	0.95	0.95	0.95	0.98	0.01	0.13	0.99
DENSENET121	0.93	0.94	0.93	0.04	0.98	0.19	0.15	0.99
ALEX NET	0.94	0.94	0.91	0.92	0.98	0.17	0.15	0.99
RESNET-50	0.96	0.95	0.93	0.94	0.98	0.14	0.13	0.99
EFFICIENT NET-B3	0.96	0.96	0.96	0.96	0.94	0.12	0.13	0.99
PROPOSED MODEL	0.98	0.98	0.98	0.98	0.99	0.07	0.09	0.99

Tabel 3 shows that our proposed model performs well for Blight leaf, Common Rust leaf and Gray Leaf spot. This improvement is due to the MobileNetV3 model which is used as the backbone for the proposed model. Feature fusion combines multi-scale feature maps from different layers, and attention module enhances the important features while suppressing irrelevant background details. In addition to it classification head is applied to predict the disease class. Figure 10 shows the Confusion Matrix for the proposed model.

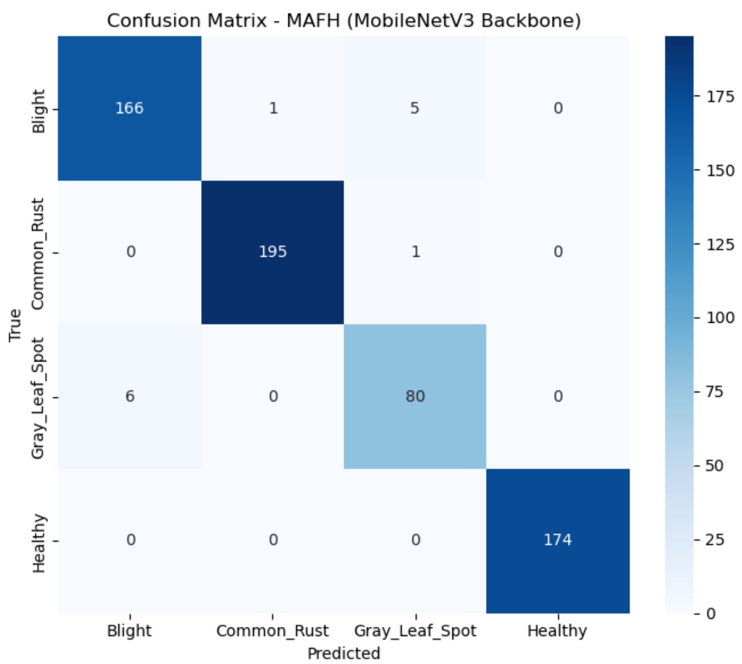


Fig 9. Confusion matrix.

We plot the graph for accuracy and loss curve. In the accuracy curve the training value goes from 0.2 to 0.6 and validation accuracy increases from 0.83 to 0.98. in loss curve the training curve decreases from 1.2 to 0.2 and validation loss decreases from 0.76 to 0.1.

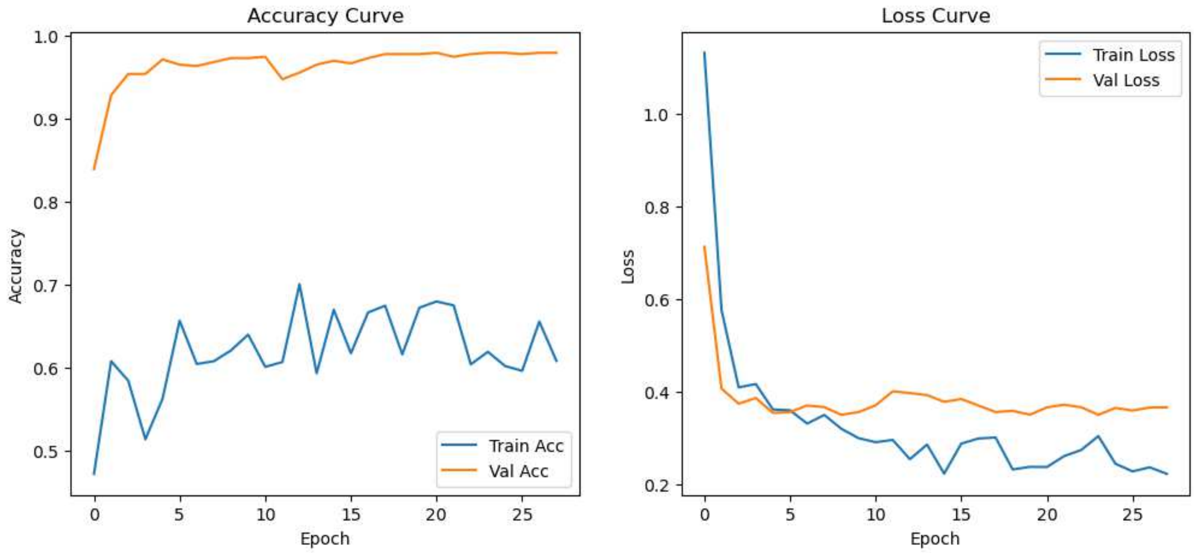


Fig 10. Accuracy and Loss Curve for proposed model

VII Conclusion and Future work

In this work an automatic maize leaf disease detection system is using deep learning has been proposed. The model successfully detects several maize leaf diseases at high accuracy through MobileNetV3 architecture with attention and fusion layers for enhancing features. The suggested method shows higher performance than typical VGG16, DENSENET121, ALEX NET, RESNET-50 and EFFICIENT NET models regarding Accuracy, Precision, Recall, F1-score, Specificity, and LogLoss AUC-ROC. MobileNetV3 reduces computational complexity, increase the accuracy. Due to its lightweight and low computational requirement, it is well suited for real time environment, and on mobile devices. Attention module is used to enhance the ability of focusing on the most important parts of the image for the processing. The module will be more robust in various lighting, noise and complex data environments. Fusion layer is used to combine two or more layers in the model. Proposed model will have a lot of layers and the image processing will happen in each of the layers and finally all the output in each layer will combine to produce the desired output. Classification head is the final stage in which it ensures that the complex visual patterns identified by the earlier network components are effectively translated into clear, interpretable diagnostic outcomes. A well-designed classification head enhances model accuracy, confidence, and stability The accuracy of the system in classifying diseased and healthy leaves helps in early diagnosis and better crop management. This research can be extended in the future by adding larger and more extensive datasets, implementing the model in mobile or IoT-based systems, and utilizing real-time field images to aid farmers with real-time disease detection and decision-making support. These findings confirm that deep learning techniques have the potential to be of utmost significance in precision agriculture via rapid and precise disease identification of plants. Secondly, the system could be easily deployed mobile applications, or IoT-based monitoring tools for farmers to diagnose diseases in real-time from field images. This can reduce crop loss, prevent pesticide misuse, and increase quality of yield.

In the future, the suggested maize leaf disease detection framework can be improved in some ways to enhance its efficiency, robustness, and real-world viability. An improvement in one such direction is an extension of the dataset by including real-time images of different maize varieties under diverse environmental conditions, lighting, and phases of disease. The multispectral or hyperspectral imaging can be investigated for the capture of spectral data outside the visible spectrum, allowing early detection of infections prior to the manifestation of visible symptoms. Further, deep learning architectures like Vision Transformers, EfficientNet, or hybrid networks involving ensemble-based networks can be utilized to maximize feature extraction and classification performance. The system can also be optimized for real-time deployment on mobile phones and IoT-based platforms, enabling farmers to conduct field-level disease detection and receive instant suggestions. We can deploy this model in drones, which are the future in agriculture field. Using them farmers can be able to find the type of disease at early stage with this farmers can save their time and money. Usage of drones in the classification of disease will decrease the dependency of humans and the error made by them. Edge computing deployment in the future can also be targeted for delivering reduced latency and enhanced efficiency in remote locations with less internet connectivity. Finally, applying this model to detect diseases in other crops and adding decision-support modules for fertilizer and pesticide advice can turn the system into an end-to-end smart agriculture solution promoting sustainable and data-based farming techniques. it works efficiently in big and real-time datasets and in all the environmental situations.

Funding: This research received no external funding.
Disclosure statement: The authors declare no conflict of interest.

Data Availability

The datasets generated and/or analyzed during the current study are available in the CORN OR MAIZE LEAF DATASET repository,
https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset

Reference

[1] Sun, J., Yang, Y., He, X., & Wu, X. (2020). Northern Maize Leaf Blight Detection under Complex Field Environment Based on Deep Learning. *IEEE Access*, *8*, 33679–33688. https://doi.org/10.1109/ACCESS.2020.2973658.

[2] Zhang, Y., Wa, S., Liu, Y., Zhou, X., Sun, P., & Ma, Q. (2021). High-accuracy detection of maize leaf diseases cnn based on multi-pathway activation function module. *Remote Sensing*, *13*(21). https://doi.org/10.3390/rs13214218

[3] Yasin, A., & Fatima, R. (2023). *On the Image-Based Detection of Tomato and Corn leaves Diseases : An in-depth comparative experiments*. http://arxiv.org/abs/2312.08659

[4] Yasin, A., & Fatima, R. (2023). *On the Image-Based Detection of Tomato and Corn leaves Diseases : An in-depth comparative experiments*. http://arxiv.org/abs/2312.08659

[5] Leng, S., Musha, Y., Yang, Y., & Feng, G. (2023). CEMLB-YOLO: Efficient Detection Model of Maize Leaf Blight in Complex Field Environments. *Applied Sciences (Switzerland)*, *13*(16). https://doi.org/10.3390/app13169285

[6] Liu, J., Hu, Y., Su, Q., Guo, J., Chen, Z., & Liu, G. (2024). Semi-Supervised One-Stage Object Detection for Maize Leaf Disease. *Agriculture (Switzerland)*, *14*(7). https://doi.org/10.3390/agriculture14071140

[7] Liu, J., Hu, Y., Su, Q., Guo, J., Chen, Z., & Liu, G. (2024). Semi-Supervised One-Stage Object Detection for Maize Leaf Disease. *Agriculture (Switzerland)*, *14*(7). https://doi.org/10.3390/agriculture14071140

[8] Tariq, M., Ali, U., Abbas, S., Hassan, S., Naqvi, R. A., Khan, M. A., & Jeong, D. (2024). Corn leaf disease: insightful diagnosis using VGG16 empowered by explainable AI. *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/fpls.2024.1402835

[9] Mayo, F., Maina, C., Mgala, M., & Mduma, N. (2024). Deep learning models for the early detection of maize streak virus and maize lethal necrosis diseases in Tanzania. *Frontiers in Artificial Intelligence*, *7*. https://doi.org/10.3389/frai.2024.1384709

[10] Vishnoi, S., & Persis, J. (2024). Intelligent crop management system for improving yield in maize production: evidence from India. *International Journal of Productivity and Performance Management*, *73*(10), 3319–3334. https://doi.org/10.1108/IJPPM-11-2023-0620

[11] Liu, J., He, C., Jiang, Y., Wang, M., Ye, Z., & He, M. (2024). A High-Precision Identification Method for Maize Leaf Diseases and Pests Based on LFMNet under Complex Backgrounds. *Plants*, *13*(13). https://doi.org/10.3390/plants13131827

[12] Yang, J., Zhu, W., Liu, G., Dai, W., Xu, Z., Wan, L., & Zhou, G. (2024). ICPNet: Advanced Maize Leaf Disease Detection with Multidimensional Attention and Coordinate Depthwise Convolution. *Plants*, *13*(16). https://doi.org/10.3390/plants13162277

[13] Yang, J., Zhu, W., Liu, G., Dai, W., Xu, Z., Wan, L., & Zhou, G. (2024). ICPNet: Advanced Maize Leaf Disease Detection with Multidimensional Attention and Coordinate Depthwise Convolution. *Plants*, *13*(16). https://doi.org/10.3390/plants13162277

[14] Hao, S., Gao, E., Ji, Z., & Ganchev, I. (2025). BCS_YOLO: Research on Corn Leaf Disease and Pest Detection Based on YOLOv11n. *Applied Sciences (Switzerland)*, *15*(15). https://doi.org/10.3390/app15158231

[15] Tang, F., Porle, R. R., Tung Yew, H., & Wong, F. (2025). Identification of Maize Diseases Based on Dynamic Convolution and Tri-Attention Mechanism. *IEEE Access*, *13*, 6834–6844. https://doi.org/10.1109/ACCESS.2025.3525661

[16] Nkuna, B. L., Abutaleb, K., Chirima, J. G., Newete, S. W., van der Walt, A. J., & Nyamugama, A. (2025). Identification of maize leaf diseases using red, green, blue-based images with convolutional neural network (CNN) and residual network (ResNet50) models. *Smart Agricultural Technology*, *12*. https://doi.org/10.1016/j.atech.2025.101226

[17] Askale, G. T., Yibel, A. B., Taye, B. M., & Wubneh, G. D. (2025). Mobile based deep CNN model for maize leaf disease detection and classification. In *Plant Methods* (Vol. 21, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13007-025-01386-5

[18] Al-Gaashani, M. S. A. M., Alkanhel, R., Ali, M. A. S., Muthanna, M. S. A., Aziz, A., & Muthanna, A. (2025). MSCPNet: A Multi-Scale Convolutional Pooling Network for Maize Disease Classification. *IEEE Access*, *13*, 11423–11446. https://doi.org/10.1109/ACCESS.2024.3524729

[19] Parashar, N., Johri, P., Elbeltagi, A., Salem, A., Choudhary, P., Kumar, V., & Agrawal, T. (2025). Enhanced residual-attention deep neural network for disease classification in maize leaf images. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-14726-1

[20] Yao, J., Zhu, Z., Yuan, M., Li, L., & Wang, M. (2025). The Detection of Maize Leaf Disease Based on an Improved Real-Time Detection Transformer Model. *Symmetry*, *17*(6). https://doi.org/10.3390/sym17060808

Authors Profile

	**Mithun Raj Palani** earned his Integrated M.Tech degree in Computer Science and Engineering, specialization in Data Science from the School of Computer Science and Engineering (SCOPE), at Vellore Institute of Technology (VIT). His interest includes data analytics, data visualization, and space technology. Apart from his academic pursuits, he is passionate about social service and actively engages in community development. **Corresponding Author:** **mithunraj.p2022@vitstudent.ac.in** **Personal Mail:** **mithunrajpalani@gmail.com**
	**KARNA DEVI AMRUTHA** received the integrated M.Tech. degree in computer science and engineering with a specialization in data science from the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India. She is currently pursuing her studies in the field of data science and computer engineering. Her research interests include machine learning, data analytics, artificial intelligence, and data visualization. **Corresponding Author:** **karnadevi.amrutha2022@vitstudent.ac.in** **Personal Mail:**

	deviamrutha.k@gmail.com