International Journal of Plant & Soil Science

Volume 38, Issue 3, Page 307-323, 2026; Article no.IJPSS.155117
ISSN: 2320-7035

Improved Soybean Leaf Disease Detection by Utilizing Separable-CNN and Global Average Pooling

Kalpesh Patel [a*] and Atul Patel [b]

[a] *College of Agriculture, Anand Agricultural University, Vaso, Kheda, Gujarat, India.*
[b] *Smt. C M Patel Institute of Computer Applications, Charotar University of Science and Technology, Changa, Anand, Gujarat, India.*

Authors' contributions

This work was carried out in collaboration between both authors. Author KP conceptualized the study, developed the methodology, contributed to software development, performed formal analysis, curated data and resources, carried out writing—review and editing of the manuscript, prepared visualizations, and managed the project. Author AP contributed to the conceptualization of the study, performed validation, formal analysis, and investigation, participated in writing—review and editing of the manuscript, and provided supervision throughout the study. Both authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/ijpss/2026/v38i36021

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here:https://pr.sdiarticle5.com/review-history/155117

Original Research Article

Received: 20/01/2026
Accepted: 25/03/2026
Published: 30/03/2026

Abstract

Soybean is one of the most important oilseed and protein-rich crops, and its productivity is seriously affected by leaf diseases that reduce both yield and quality. Both timely and prompt detection of the disease is thus a requirement in proper crop management especially in farming conditions of India where timely diagnosis at a field level can provide the opportunity to act fast. This paper proposes a lightweight deep-learning model, which is a Separable Convolutional Neural Network (Separable-CNN) with Global Average Pooling (GAP), to detect soybean leaf disease. The depthwise separable convolutions decrease the level of computation, and the GAP-based aggregation of features tends to reduce overfitting and enhance compactness of the model.

**Corresponding author: E-mail: kpatel@aau.in;*

Cite as: *Patel, K., & Patel, A. (2026). Improved Soybean Leaf Disease Detection by Utilizing Separable-CNN and Global Average Pooling. International Journal of Plant & Soil Science, 38(3), 307–323. https://doi.org/10.9734/ijpss/2026/v38i36021*

The suggested approach was trained and tested on a dataset of 40,000 leaf images of soybean cultivated in the field of eight classes, the seven classes comprising of disease classes, while the remaining one was a healthy one. It was observed in experiments that the model obtained a test accuracy of 95.70% with only a small architecture of about 0.78 million parameters and a model size of 9.04 MB. Relative analysis of the existing deep-learning architectures proved that the proposed architecture provides a good tradeoff between the classification performance and the computational efficiency. These results suggest that the suggested model can be effectively used in agricultural devices with limited resources to detect the soybean disease in real-time.

Keywords: Separable-CNN; Global Average Pooling; soybean leaf disease; Indian agriculture; deep learning.

1. Introduction

Soybean production is important to Indian economy and food security, because of its nutritional value and use in the kitchen. Farmers are struggling to identify diseases of soybean leaves, which result in the reduction of crop yield and quality. Traditional diagnosis is laborious and prone to mistakes due to manual examination by specialists, providing the need for efficient diagnostic techniques. Various pathogens and environmental conditions present in India complicate leaf diseases and rust, blight and mosaic viruses can cause huge damage to crop if not addressed on time (Elmore et al., 2022). Artificial intelligence and deep learning is a promising solution, especially Convolutional Neural Networks (CNNs). However, traditional CNN models have high computational costs and are prone to overfitting, making them unsuitable for resource-constrained Indian farms. In this paper we propose a new separable Convolutional Neural Networks (Separable-CNNs) model with Global Average Pooling (GAP) topology. Using depth-wise separable convolutions, conventional convolution is decomposed into depth-wise and point-wise steps and reduces computational complexity without degrading accuracy (Zhao et al., 2018). A GAP layer is used to improve the model by integrating feature maps into a global feature, which improves the extraction robustness and reduces the overfitting (Hu et al., 2023).

The main aim of this study is to provide an efficient model of soybean leaf diseases detection in various agronomical conditions of India. The objectives of the study are three:

- Design and Implementation: To design and implement a model of Separable Convolutional Neural Network (Separable-CNN) model with Global Average Pooling (GAP) layer for soybean leaf disease detection.
- Performance Evaluation To evaluate the performance of the model on a dataset of soybean leaf images, focusing on precision, recall, and accuracy metrics.
- Practical Applicability: To assess the feasibility of the model in terms of practical applicability to real-time disease detection in the field to ensure the efficient use of the model by farmers.

The use of a Global Average Pooling (GAP) layer for feature extraction is similar to that in Shaheed et al. in which it was used for classification in potato leaf disease detection (Shaheed et al., 2023). Interestingly, the proposed method is similar to other light weight models aiming at plant disease detection. For example, Wang et al. presents an ultra-lightweight efficient network (ULEN), which also makes use of depth-wise convolution and multi-scale features for plant disease classification (Wang et al., 2023). Similarly, Yuan et al. introduces an improved ResNet34 model (ESA-ResNet34), which substitutes the depthwise separable convolutions in place of standard convolution, which significantly decreases the parameter number and computational burden (Yuan et al., 2024).

The proposed Separable-CNN with GAP layer is conceptually related to recent lightweight CNN architectures such as MobileNetV3, EfficientNet(-Lite) and one stage object detectors such as YOLOv5 which have been optimised for mobile and embedded hardware. These models usually use a combination of depthwise separable convolution, neural architecture search and compound scaling to maximise ImageNet accuracy with latency constraints (Howard et al., 2019). In contrast, our design is intentionally task specific and shallow: it consists of five depthwise separable convolutional blocks followed by two GAP layers and no auxiliary heads or detector branches. This leads to a much shorter training time and a small memory footprint whilst maintaining 95.70% test accuracy on eight classes of soybean diseases with the 70/20/10 split. Rather than competing outright with large, general-purpose architectures, the proposed model thus shows that a well-tuned Separable-CNN + GAP topology which is compact can match or surpass the performance of heavier transfer learning baselines (e.g. Inception-ResNetV2, Xception, ResNetV2, VGG-16/VGG-19) reported in Table 3, and at the same time is easier to deploy on low-cost edge devices common in rural Indian farms.

Compared with the plant disease models based on MobileNetV3 and EfficientNet and trained on generic multi-crop plant datasets such as PlantVillage or composite benchmarks (Howard et al., 2019), our network is trained end-to-end on a field-acquired soybean-specific dataset with eight agronomically important classes. This crop and region specific focus, depth wise separable convolutions and GAP help to produce a model which is computationally efficient and is aligned with hardware and connectivity constraints faced by smallholder farmers in India.

Despite the fact that recent works have shown the utility of CNN-based and transfer-learning methods in the classification of plant diseases, there are still a number of practical challenges. Most of the reported models are based on large and computationally intensive backbones are tested mostly on curated or generic benchmark datasets and have little discussion of their deployment to resource constrained agricultural conditions. Moreover, few studies have been conducted specifically to detect soybean leaf disease based on a compact architecture on data collected in the field (regarding Indian weather conditions and Indian agriculture). Consequently, the compact but precise model requires the minimization of parameters and memory consumption and a high level of classification accuracy under the real conditions of agricultural variability on soybean diseases. In order to fill this gap, the current paper will present a Separable-CNN with Global Average Pooling to effectively detect soybean leaf disease and compare it with popular deep-learning baselines.

2. Related Work

The science of identifying disease in plants has made huge leaps forward with the aid of various forms of deep learning. Xu et al have proposed an improved ShuffleNetV2 model for detecting diseases in fungi (fruit bodies) such that the model is efficient and accurate of agriculture (Xu et al., 2023). Liu et al. presented a high-precision detection method suitable for a low-computing platform, based on a dynamic pruning gate to achieve an accurate balance between accuracy and computational efficiency (Liu et al., 2023). Vadivel and Suguna worked on the detection of tomato leaf disease by using improvements in the learning methods coupled with image processing to ensure reliable detection of the disease (Liu et al., 2023). Gong et al. developed a light-weight model for powdery mildew detection that can be used in-field using portable devices (Gong et al., 2022). Cheng et al. proposed an attentional feature fusion model, which focuses on the lightweight design for crop disease identification (Cheng et al., 2022). Haque et al. applied deep learning to identify diseases in maize crops, highlighting the model's potential in practical agricultural applications (Cheng et al., 2022). Thakur et al. involved a vision transformer integrated with a convolutional neural network to enhance interpretability and performance for plant disease diagnosis (Thakur et al., 2022). To identify illnesses in tomato plants, Abbas et al. used a combination of transfer learning and synthetic pictures produced by C-GAN, showing that the two methods may work together to improve accuracy (Abbas et al., 2021). To identify maize diseases efficiently and accurately, Chen et al. suggested a lightweight network with attention incorporated (Chen et al., 2021). For plant leaf disease identification, both Hasan et al. and Maji et al. used CNN-based methods; however, Maji et al. used shallow networks to strike a compromise between complexity and performance (Hassan et al., 2021; Hassan et al., 2021).

By using deep learning models, Chao et al. were able to identify illnesses in apple tree leaves with a much higher degree of accuracy (Chao et al., 2020). To highlight the importance of early diagnosis and treatment of plant diseases, Chakravarthy and Raman used deep learning to identify early blight in tomato leaves (Chakravarthy & Raman, 2020). In order to identify plant diseases, Mohameth et al. (2020)] used deep learning and feature extraction, and they validated their model using the Plant Village dataset. Wspanialy and Moussa demonstrated a method for detecting and quantifying the difficulty of generic illnesses in tomato pots grown in greenhouses, drawing attention to its usefulness in controlled settings (Wspanialy & Moussa, 2020). With the use of adversarial networks for unsupervised image translation, Nazki et al. significantly improved plant disease identification accuracy (Nazki et al., 2020). By using few-shot learning to categorize plant diseases from field pictures, Argüeso et al. (2020) shown that the model could generalize from sparse data. A model for plant disease detection was developed by Darwish et al. using convolutional neural networks and orthogonal learning particle swarm optimization. This study showcases unique optimization methodologies (Darwish et al., 2020).

By combining deep learning with an improved lightweight network, Chen et al. were able to identify plant illnesses more accurately and efficiently (Chen et al., 2020). In their presentation of ToLeD, Agarwal et al. showcased a convolutional neural network–based system for disease identification in tomato leaves, which has real-world implications for agricultural management (Agarwal et al., 2020). To determine which transfer learning algorithms, work best for plant disease diagnosis, Toae et al. compared deep learning models that have been fine-tuned (Too et al., 2019).

Barbedo demonstrated the model's accuracy in disease localization by identifying plant illnesses from specific lesions and spots using deep learning (Barbedo, 2019). Demonstrating the efficacy of lightweight models, KC et al. investigated depthwise separable convolution topologies for plant disease classification KC et al., 2019). As an example of the usefulness of hybrid models, Sardogan et al. (2018) used CNN with the LVQ algorithm to classify and identify diseases in plant leaves. Several typical constraints are still present in the literature, even though deep learning has made great strides in plant disease identification. When applying models to varied agricultural contexts, there is a risk of bias and generalization problems due to the dependence on big, annotated datasets, which may be tedious and time-consuming to create (Xu et al., 2023; Vadivel & Suguna, 2022; Cheng et al., 2022). Even models that are tuned for low-computing platforms or are lightweight still need a lot of computational power to train, which might prevent them from being widely used in environments with restricted resources (Haque et al., 2022; Chen et al., 2021; Chen et al., 2021). Further, unlike in controlled laboratory settings, real-world variances like as changes in illumination, occlusions, and disease stages may severely impair the performance of these models (Hassan et al., 2021; Chao et al., 2020; Chakravarthy & Raman, 2020; Wu & Xu, 2021). Another drawback is that deep learning models are sometimes not interpretable, which makes it hard for users to comprehend the decision-making process and have faith in the predictions. This is especially problematic in important agricultural applications (Mohameth et al., 2020; Karthik et al., 2020). Furthermore, approaches for synthetic data creation and transfer learning have great potential, but they are still in their early phases and need to be validated further to make sure they work in different agricultural settings (Darwish et al., 2020; Barbedo, 2019; Sardogan et al., 2018; Duong et al., 2020).

3. Materials and Methods

This section outlines the process of data set acquisition, pre-processing, augmentation process, experimental design, baseline models being compared, evaluation measures, and the proposed Separable-CNN with GAP architecture. In order to create more clarity, the procedures that were directly used in this study are described here, and the general background concepts of deep learning evaluation are also briefly discussed where needed.

3.1 Dataset Images

To assure the authenticity and applicability of the dataset to the real world, the data on soybean leaf deficiencies used in this study have been selected carefully after real field photographs have been accurately taken as shown in Fig. 1. The curation process is comprised of some main stages. To understand the soybean leaf diseases better, images have been taken directly on the fields of the soybeans trying the conditions of natural farming. The data was recorded in various areas related to the University and the ad-jacent areas, at different stages of the growth of the crops, with the natural variations in the leaf orientation, severities of the disease as well as clutter of the backgrounds. The original field photographs were annotated with agricultural experts and manually blacked in the background, to remove non-informative background (soil, weeds and sky) in order to increase label reliability and improve the photos to be more suitable for deep learning.

The last dataset will consist of 40,000 photos in eight categories with seven types of diseases and one healthy class. All the images were scaled to 224 x 224 pixels in order to standardize the learning process. This preprocessing was added to minimize visual noise to stimulate the model to pay attention to disease-specific symptoms of leaves like lesions, discoloration, mildew pattern, and rust pustule. But since performance in natural background deployment may also depend upon natural background clutter, this preprocessing option is also recognized as a weakness and it must be additionally tested on raw field images in future studies.

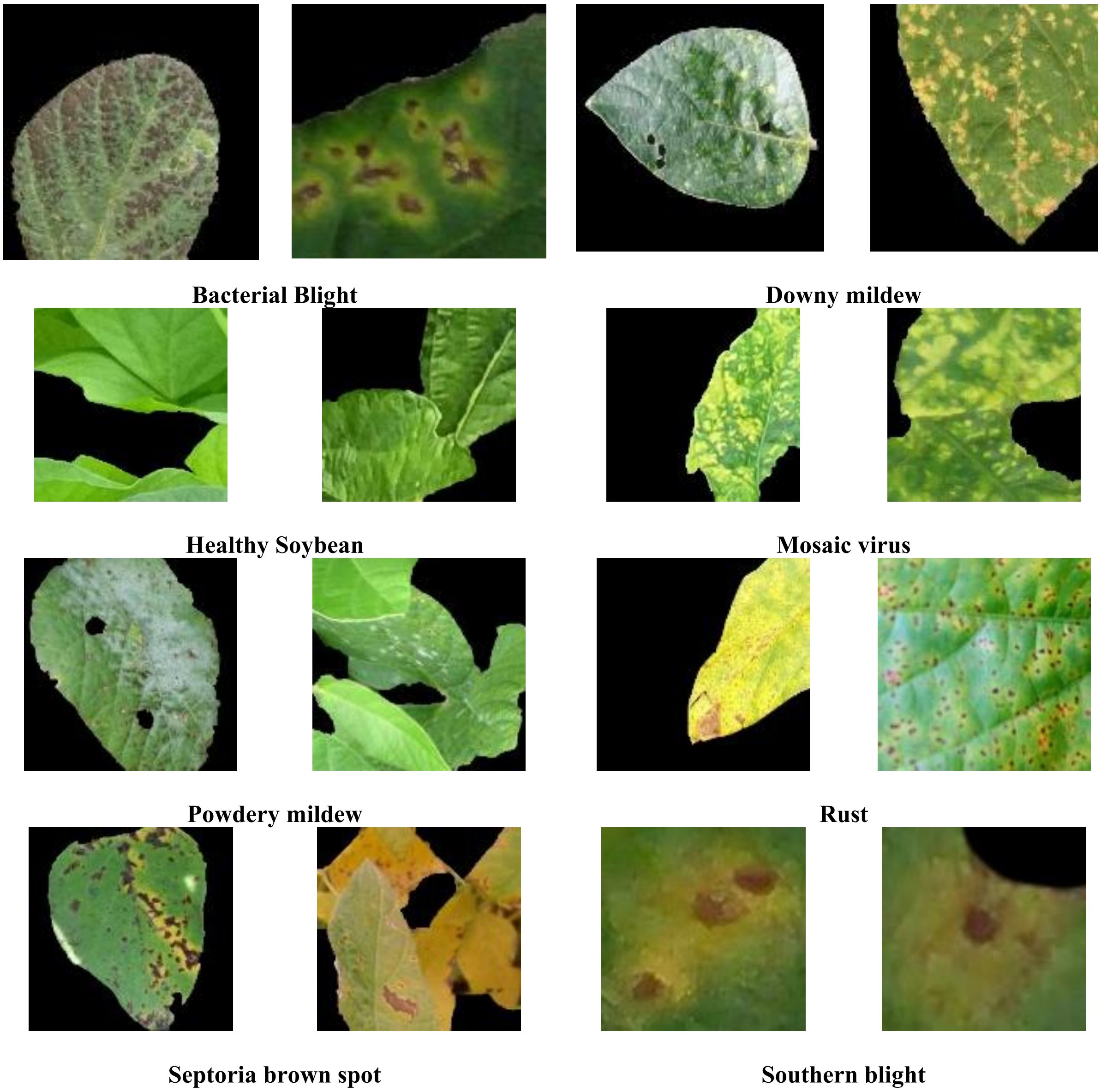


Fig. 1. Dataset Images of Soybean Leaf

Table 1. No of Dataset Images

No.	Class / Disease Name	No. of Images	Share (%)
1	Downy mildew	5,000	12.5
2	Bacterial-Blight	5,000	12.5
3	Powdery mildew	5,000	12.5
4	Mosaic-virus	5,000	12.5
5	Septoria brown spot	5,000	12.5
6	Southern blight	5,000	12.5
7	Healthy Soybean	5,000	12.5
8	Rust	5,000	12.5
	Total	**40,000**	**100**

3.2 Transfer Learning Models

The convolutional neural network AlexNet played an important part in the ImageNet competition. They demonstrated that a large RNN can perform well on a difficult dataset using supervised learning methods, proving that RNNs are useful for classification applications (Vadivel & Suguna, 2022; Gong et al., 2022). A crosscut connection, known as an avoided connection, is a feature of ResNet that enables unaltered data to flow. Input x is converted into output F(x) using a succession of activation layers in the network. A key aspect of the residual unit architecture is the handling of input-reference signal discrepancies x and F(x). This design guarantees that signals are sent across the network with minimal deterioration after the network approximates the output function of a certain layer. The optimizer often reduces the total weight of the remaining blocks at each level to almost zero, owing to this design (Cheng et al., 2022; Thakur et al., 2022; Abbas et al., 2021). The (VGG) network consists of several convolutional layers. Comparing the two models, we can see that VGG-19 is a model that has 19 layers and the VGG-16 model has 16 layers. These models showed the importance of deep convolution layers to improve performance and represent the backbone of modern object identification systems (Hassan et al., 2021; Chakravarthy & Raman, 2020). The ImageNet collection, which consists of more than one million photos, was used to pretreat Xception, a convolutional neural network with 71 layers. Animals, computer parts, and commonplace items are just a few examples of the thousands of possible classifications that this pre-trained model can achieve. The requirement of 299 pixels on each side of the pictures increases the ability of the network to learn complex feature representations in different image formats (Mohameth et al., 2020; Karthik et al., 2020). The distinctive architectural structure of an inception network is determined by the existence of repeated inception modules. For the ImageNet data set, the image detection model Inception v3 had a precision of more than 78.1%. With the help of several novel concepts formulated by different researchers over the years, this model has been improved to achieve improved picture recognition tasks. To facilitate even comparison between all of the baseline models and the proposed model, they were all trained under the same set of experimental conditions, with the same data splits. The baselines of transfer-learning were initialized with pretrained ImageNet weights on the soybean dataset. This regular training protocol has been used to maintain that performance variation is based on behavior of architecture and not variation in optimization settings.

3.3 Evaluation Parameters

The time parameter limits the maximum time a job can use the processor to a more precise degree, which gives more precise control over the use of the CPU. Efficient usage of computing resources is guaranteed at the step and job level by adjusting this value based on the needs of the task (Wspanialy & Moussa, 2020).

- The training accuracy of a model is a measure of how good the model is in learning the training data and how good the model is in applying to an original data set. The quality of the model in a validation data set is revealed by checking the quality of the model, which is also known as validation accuracy. In order to understand whether the model was under- or over-fitting, the training and validation accuracies were compared (Liu et al., 2023; Cheng et al., 2022).
- The training loss is a metric which is used to measure the performance of the algorithm in fitting the training data. This reveals how much the model learns from the data, and is hence an important measure of how effective the model is during its training. The generalization capability of the model to generalize new data is measured by the validation error which assesses the performance of the model on new unseen validation data (Liu et al., 2023).
- The accuracy of the model to identify the patterns and connections in the test data in relation to the training data was measured using testing accuracy. It is an important measure for the assessment of the generalizability and overall effectiveness of the model (Chen et al., 2021; Gonzalez-Huitron et al., 2021)

Quantifying the inaccuracy of the model using the test dataset is called the testing loss. This indicates the level of agreement between the actual values at some location versus what the model predicts. When the testing loss is low, the model is correct and when it is high, prediction errors are high (Haque et al., 2022; Abbas et al., 2021). To be able to assess classification performance more comprehensively, the proposed model was assessed based on accuracy, precision, recall, and F1-score alongside the values of loss. Whereas accuracy is a summary of absolute correctness, precision and recall are class-sensitive measures of false positive and false negative behavior respectively and F1-score summarizes the balance between these two. These measures are especially significant in the context of plant diseases detection, where false identification of similar in appearance diseases can be a factor in the choice of treatment.

3.4 Implementation Details and Reproducibility

All experiments were implemented in TensorFlow using Python 3.7. To facilitate reproducibility, a fixed random seed (42) was used for all libraries involved in data shuffling and weight initialization, and identical train–validation–test splits were maintained across all compared models. Unless otherwise stated, the primary experiments used a 70/20/10 split for training, validation, and testing, respectively, while additional experiments with 65/25/10 and 60/30/10 splits were performed to evaluate robustness to changes in data partitioning in Table 3. During training, standard on-the-fly data augmentation was applied to improve generalization under field conditions. In real-world application, the images of soybean leaves can be varied in orientation, scale, framing and lighting based on the angle of the capture and the environmental factors around the point of capture. Thus, augmentation through random rotation, horizontal flipping, width and height shifting, and moderate zoom were applied to create realistic variations that might happen during the process of taking pictures with your handheld and to decrease the possibility of overfitting. All models were trained using the Adam optimizer with a learning rate of 0.0001, batch size of 32, and up to 100 epochs, as summarized in Table 2.

3.5 Proposed System Training Workflow

The following Algorithm outlines the training process for the proposed Separable-CNN with GAP architecture.

```
Algorithm : Training Pipeline for Separable-CNN with GAP
Input: Dataset D, Epochs E = 100, Batch Size = 32,
        Learning Rate lr = 0.0001
Output: Trained Model M

1:  Set random seed = 42
2:  Load dataset D and preprocess images (resize 224×224×3)
3:  Apply augmentation: rotation, flip, width and height shift,
                              zoom
4:  Split D into Train(70%), Validation(20%), Test(10%)
5:  Initialize model M using Separable Convolution blocks +
    GAP layer
6:  Initialize optimizer = Adam(lr),
    loss = categorical cross-entropy
7:  For epoch = 1 to E do
8:      Train M on Train set and compute training accuracy/loss
9:      Validate M on Validation set and compute validation
        accuracy/loss
10: End For
11: Evaluate M on Test set and record metrics
12: Save model weights and training log
13: Return trained model M
```

3.6 Proposed System

Examining all of the huge convolutional ResNets and VGGs, there is a question of how to reduce the number of parameters in these networks without losing accuracy or even increasing the generalizability of the model, as shown in Fig. 2 of the proposed model architecture, which breaks down the networks into layers. The solution is a separable CNN layer and GAP, as follows:

3.6.1 Separable Convolution

Standard convolution simultaneously learns spatial and cross-channel relationships but this results in a high number of parameters and computational operations. For an input feature map of size $H \times W \times C_{in}$ and C_{out} filters of size $k \times k$, a conventional convolutional layer requires $k^2 C_{in} C_{out}$ parameters. Depthwise separable convolution factorises this operation into two steps:

(i) a **depthwise convolution**, where a separate $k \times k$ filter is applied to each input channel independently, and (ii) a **pointwise convolution**, where 1×1 filters linearly combine the resulting channels (Howard et al., 2019;

Zhou et al., 2024). The total parameter count becomes $k^2 C_{in} + C_{in} C_{out}$, which is much smaller than that of a standard convolution when $k > 1$. This factorization has significant counts of fewer parameters and fewer multiply-accumulate operations relative to traditional convolution and has the ability to learn discriminative visual patterns. All the convolutional blocks in the proposed architecture are performed with depthwise separable convolutions to efficiently extract disease related texture information. The design is also suitable to the classification of soybean leaf disease since it minimizes the complexity of the models, but still has the ability to reproduce visual symptoms like rust pustules, mildew structures, mosaic patterns, and lesion boundaries. This design is based on the principle of the Mobile friendly CNN namely the MobileNet and EfficientNet, but it is intended to be used for the specific purpose of a soybean leaf disease classification in the Indian field condition (Howard et al., 2019; Parez et al., 2023).

The approach has been successfully used for various computer vision applications, such as hyperspectral image classification (Gao et al., 2023; Ghaderizadeh et al., 2021; Wei et al., 2023), semantic segmentation (Dai et al., 2023; Wang et al., 2021), and sound source localization (Krause et al., 2021). In summary, depthwise separable convolutions offer a powerful alternative to standard convolutions in deep learning models, providing a balance between computational efficiency and feature extraction capabilities. Fig. 3 illustrates the concept using a 3×3 kernel that is decomposed into two 1D convolutions. This decomposition reduces computation while preserving the representational capacity needed to capture disease-specific spatial structures.

The results are shown in Fig. 3. Nine parameters with a 3 × 3 kernel were would be 9 parameters. In a spatially separable convolution, the kernel is split into two smaller kernels: one with a 3 × 1 form and the other with a 1 × 3 shape. Three plus three equaled six parameters. This implies that matrix multiplication is not necessary. The computational cost is reduced by breaking down the 2D convolution into two 1D convolutions, and the overall computational cost is reduced. This design is much lower than common convolution in terms of trainable parameter and computation cost, and thus it is very appropriate in low-weight image-classification models.

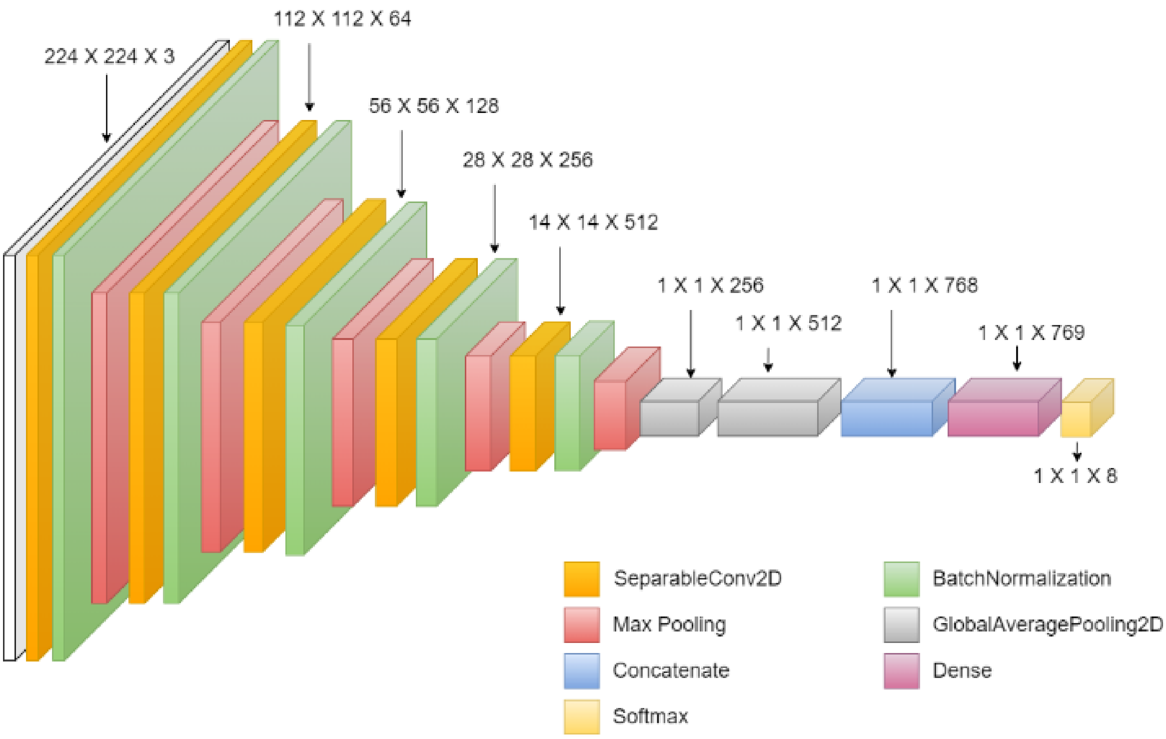


Fig. 2. Proposed Model Architecture

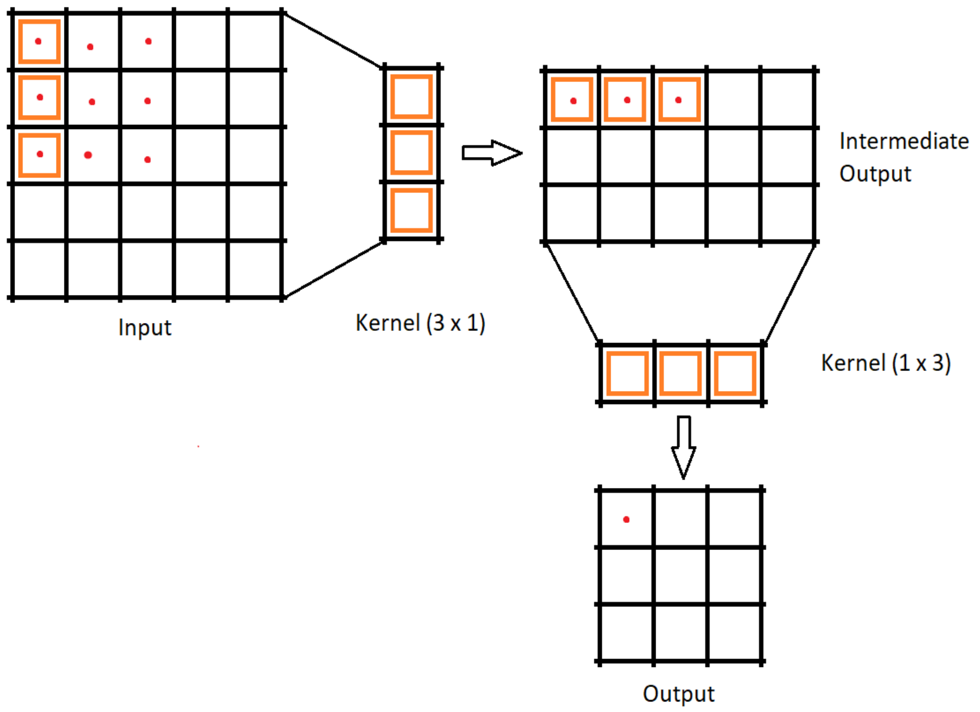


Fig. 3. Depthwise Separable Convolutions

- GAP Layer: This works with global average pooling (GAP) instead of the flattened layer in the proposed model to improve the performance of the model by increasing the accuracy and significantly reducing the trainable parameters.
- Flatten Layer: Used to reshape the tensor of the input data into a one-dimensional array. This results in a larger output size because it combines all elements into a single dimension. This can result in several parameters that can cause overfitting.
- Global Average Pooling (GAP): GAP performs an average pooling operation, reducing the spatial dimensions. Global Average Pooling of feature maps obtained on two features networks depths and then features are fused. The design allows the model to maintain both complementary information of the intermediate representation and deeper representation. Their concatenation gives a compact yet informative way to represent them to classify them, whilst still not increasing the large number of parameters normally by flattening-based fully connected layers. This helps prevent overfitting and reduces computational cost.

Table 2. Model Hyper Parameters

Category	Hyperparameter	Setting / Value
Dataset	Crops	Soybean
Layers	Separable Convolutional	5
Layers	Normalization	5
Layers	Max Pooling	5
Layers	GAP (Global Avg. Pooling)	2
Training	Activation	Softmax, ReLU
Training	Epochs	100
Training	Optimizer	Adam
Training	Learning Rate	0.0001
Input	Image Resolution	224 × 224 × 3
Training	Batch Size	32

4. Results and Discussion

Here, we discuss the comparison of baseline models and old-fashioned machine-learning methods. The proposed Separable-CNN with GAP was compared to a number of robust CNN baselines (Inception-ResNetV2, Xception, ResNetV2, VGG- 16, VGG- 19, and Inception-V3) using various train-validation-test splits (70/20/10, 65/25/10, 60/30/10) as summarised in Table 3. In all splits, we got the best validation and test accuracy with our model as well as the lowest loss values. Indicatively, the proposed model achieved 96.58 percent validation accuracy and a loss of 0.10096 and 95.605 percent test accuracy and a loss of 0.15041 in the 65/25/10 split, which is the best result compared to all the transfer-learning baselines. Such stability in various data partitions indicates that Separable CNN + GAP architecture generalises well on unseen soybean leaf images. Here, we present the results of several models and a comparative study that takes into account different factors.

Table 3. Complexity comparison of the proposed model and popular deep architectures

Model	Parameters (M)	Model size (MB)	Depth/Layers
VGG16	138.4	528	16
VGG19	143.7	549	19
ResNet50	25.6	98	107
ResNet152V2	60.4	232	307
InceptionV3	23.9	92	189
InceptionResNetV2	55.9	215	449
Xception	22.9	88	81
EfficientNet-Lite0	≈4.7	≈18	–
MobileNetV3-Small	≈2.5	≈10	≈14
ShuffleNetV2 1.0×	≈2.3	≈9	–
Proposed Model	0.78	9.04	19

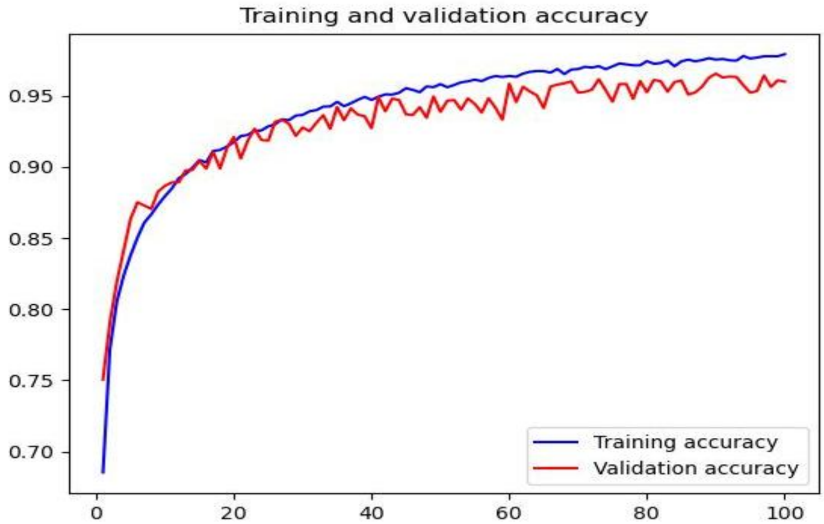
To gather the data, the workstation used a GPU Nvidia Quadro P2000 - 5GB, 32 GB of memory, and an Intel(R) Xeon(R) CPU E5-1650 @ 3.60 GHz. By using the suggested Separable-CNN with Global Average Pooling model, which has been trained on massive datasets and incorporates the learned weights and architecture, we can now include learning in our issue statement.

As shown in Table 3, the suggested Separable-CNN with two GAP heads is very sparse (0.78 M parameters, 9.04 MB), which is approximately 60 to180 times thinner than heavy transfer-learning backbones (VGG16/VGG19 and ResNet152V2) and yet other Separable-CNNs can be more accurate in soybean leaf disease classification. Our architecture requires approximately 3 to 6 times fewer parameters than new lightweight architectures, even when compared to recent models like EfficientNet-Lite, MobileNetV3-Small and ShuffleNetV2, and provides comparable performance. The ratio between accuracy and complexity is favourable and therefore the proposed model is especially appealing to memory and power constrained edge deployments in real time agricultural applications.

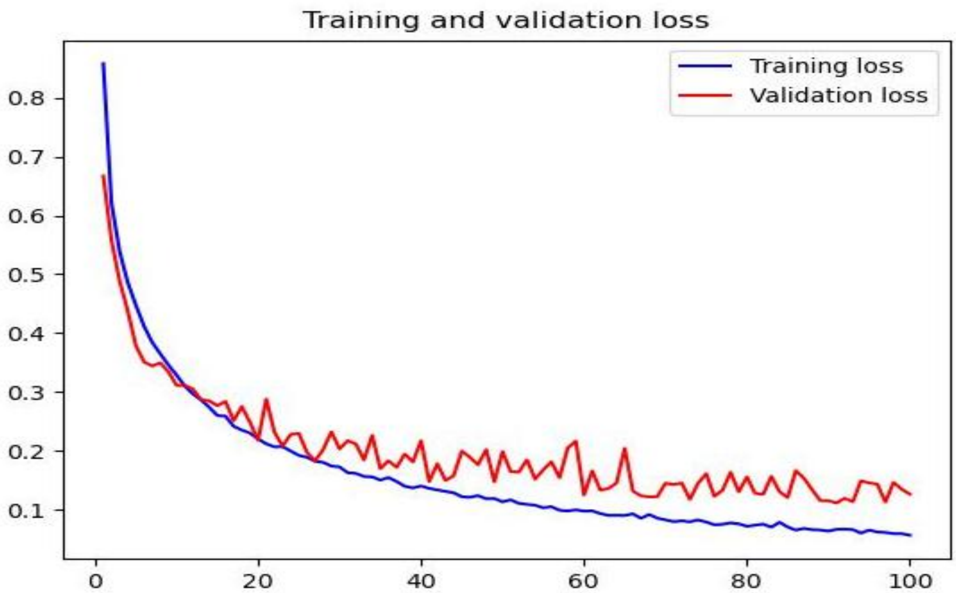
As shown in Fig. 4(a), the accuracy increased as the epoch increased; however, at a certain limit, it became linear. Therefore, we stopped the epochs. As shown in Fig.4(b), the loss decreases as the epoch increases; however, at a certain limit, it becomes linear. Therefore, we stopped the epochs.

As shown in Fig. 5(a), the confusion matrix indicates the testing of the model using test samples. Fig. 5(b) of the classification report shows that the proposed model was trained using a split-image dataset, with 70% of the time spent on training, 20% for validation, and 10% for testing.

As shown in Table 4, transfer learning is related to this. Table 5 lists the complexity of time. The accuracy and loss metrics for the test and validation data indicated that the Proposed Model was the best. Regardless of the data split ratio, the best results were obtained in terms of the accuracy and loss. For validation performance, the proposed model achieved accuracies of 95.61%, 96.58%, and 96.39% for the 70/20/10, 65/25/10, and 60/30/10 splits, respectively, with correspondingly low loss values of 0.1176, 0.1009, and 0.1219. On the test sets, the model achieved accuracies of 95.70%, 95.61%, and 95.31%, respectively, again with the lowest losses among the compared models. There is a wide range of CNN models and data-split ratios in terms of accuracy and loss values. As an example, ResNet152's performance is decent on the 70/20/10 split, but it suffers on the 65/25/10 and 60/30/10 splits owing to increased loss and decreased accuracy. Although VGG-16 and VGG-19 show some variation in performance across data split ratios, they are generally not as accurate as the other models. The success of the Proposed Model in employing pre-trained models for a given job is demonstrated by the fact that it obtains the maximum accuracy across all data split ratios according to the transfer learning findings. The Proposed Model stands out as the best performer in this analysis. Its low loss and excellent accuracy make it seem like it a good fit for the current job.

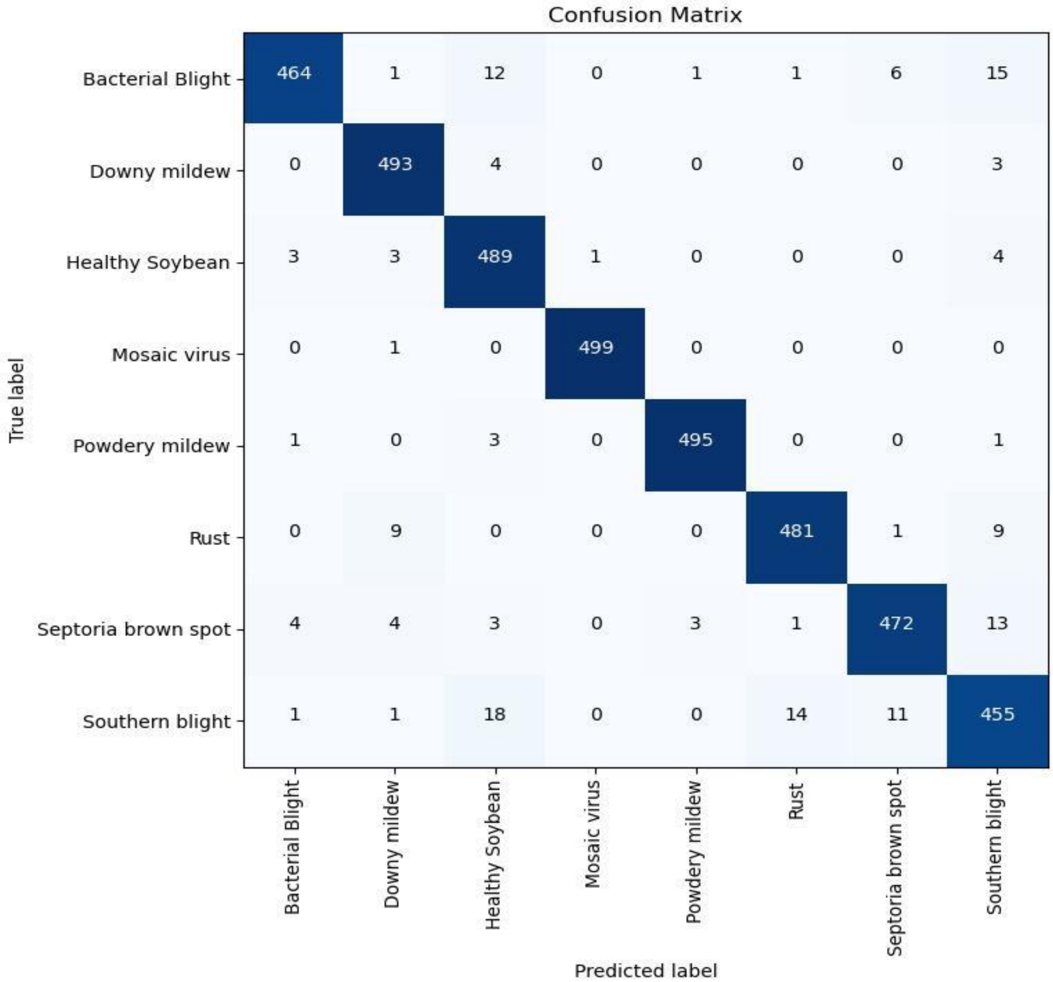


(a)



(b)

Fig. 4. Training dynamics of the proposed model (a) Training and validation accuracy (b) Training and validation loss



(a)

Classification Report

	precision	recall	f1-score	support
Bacterial_Blight	0.98	0.93	0.95	500
Downy_mildew	0.96	0.99	0.97	500
Healthy_Soybean	0.92	0.98	0.95	500
Mosaic_virus	1.00	1.00	1.00	500
Powdery_mildew	0.99	0.99	0.99	500
Rust	0.97	0.96	0.96	500
Septoria_brown_spot	0.96	0.94	0.95	500
Southern_blight	0.91	0.91	0.91	500
accuracy			0.96	4000
macro avg	0.96	0.96	0.96	4000
weighted avg	0.96	0.96	0.96	4000

(b)

Fig. 5. Evaluation on the test set (a) Normalized confusion matrix (b) Classification report

Table 4. Models Comparative Analysis

Data to Validate						
	TR/VA/TE 70/20/10		**TR/VA/TE 65/25/10**		**TR/VA/TE 60/30/10**	
Model	**_loss_**	**ACC**	**_loss_**	**ACC**	**_loss_**	**ACC**
Inception-ResNetV2	0.7192	0.7748	0.569	0.8089	0.6559	0.7845
XceptionNet	0.8738	0.7123	0.7958	0.7357	1.0826	0.6712
ResNetV2	0.9271	0.808	1.3653	0.7279	1.4283	0.7347
VGG-16	0.7797	0.7201	0.687	0.7484	0.7777	0.7201
VGG_19	0.7643	0.7025	0.884	0.6888	0.666	0.7445
Inception_V3	0.7204	0.7543	0.8286	0.7318	1.0675	0.6986
Proposed Model	0.1176	0.9560	0.1009	0.9658	0.1219	0.9638
Data to Test						
	TR/VA/TE 70/20/10		**TR/VA/TE 65/25/10**		**TR/VA/TE 60/30/10**	
Model	**loss**	**accuracy**	**loss**	**accuracy**	**loss**	**accuracy**
Inception-ResNetV2	0.5823	0.856	0.4129	0.8843	0.5398	0.8511
XceptionNet	0.5644	0.8472	0.4986	0.8687	0.7294	0.8082
ResNetV2	0.9394	0.8203	1.3776	0.7402	1.4406	0.7470
VGG-16	0.5132	0.8404	0.4323	0.8697	0.4877	0.8472
VGG-19	0.5234	0.8365	0.5691	0.8111	0.5021	0.8335
Inception-V3	0.5519	0.8501	0.6662	0.8257	0.6193	0.8169
Proposed Model	0.1472	0.9570	0.1504	0.9560	0.1707	0.9531

Table 5. Time Complexity Analysis

Parameters	**70/20/10**	**65/25/10**	**60/30/10**
Models	**(Hours)**	**(Hours)**	**(Hours)**
Proposed CNN	09:12:56	09:01:42	08:12:56

The results are consistent with and supplementary to new lightweight CNN and transformer-based models of plant disease detection. MobileNetV3 and EfficientNet powered networks like E-GreenNet and Dise-Efficient have proven to be accurate and fast inference on generic multi-crop datasets with depthwise separable convolutions and well-scaled architectures to run on edges (Parez et al., 2023; Gole et al., 2023). The variants of ShuffleNetV2 and similar designs also have demonstrated strong performance in the presence of complex field backgrounds, with small-sized models (Zhou et al., 2024). In more recent times lightweight Vision Transformer hybrids like TrIncNet and PMvt have demonstrated that transformer-style global context can additionally be used to enhance performance, but at the expense of more complex architectures and increased compute demands (Gole et al., 2023; Li et al., 2023). In contrast, we use convolutional operations and GAP as the sole architectural elements of our model, which is architecturally low but achieves competitive accuracy in a field-collected soybean dataset. Its simplicity is appealing to be used on mid-range CPUs, low-cost GPUs, or embedded systems where both memory and power constraints are limited.

Even though the dataset is gathered in the real fields, backgrounds are blacked out manually, which removes clutter and makes it easier to train models, but potentially partially removes exposure to the more complicated backgrounds that occur in uncontrolled environments. Furthermore, whereas conventional augmentation causes differences in perspective and light, the current experiments do not explicitly measure robustness to high shadows, partial occlusions (e.g. overlapping leaves), or even early/late disease stages. It will be significant in the future to test the model with both raw and unsegmented field images on raw datasets and on multi-year and multi-location datasets to establish that the model can generalise additional farming conditions. To ensure that the network targets biologically significant lesions when making predictions, which is essential to the practical use of the network in advisory services, the incorporation of explainability tools, including Grad-CAM-based explainability and perturbation-based saliency maps, will also be important.

5. Conclusion and Future Work

Several important insights were obtained by comparing the different models using different data splits for the validation and test datasets. This study introduced a light Separable-CNN with Global Average Pooling to detect soybean leaf disease. The architecture proposed aimed at reducing the computational complexity and also maintaining the discriminative power with field-relevant classification of soybean diseases. Experimental analysis on a dataset of 40,000 images in eight classes demonstrated that the model has achieved high performance levels in comparison to a number of established transfer-learning baselines on various train-validation-test splits. The suggested model obtained high test performance such as 95.70% accuracy on the 70/20/10 split, although the model size is 9.04 MB with only around 0.78 million parameters in the model. Its features particularly render it very appealing to deployment-oriented agricultural applications where memory, storage, and power can be limited. The findings thus indicate that a well-crafted task-specific light-weight model can provide a more desirable approach to complexity as compared to performance relative to heavier general-purpose CNN backbones on soybean disease detection. The proposed model is the most suitable in the diagnosis of soybean leaf disease because it has a tolerable training time, a low loss, and a high accuracy in various data splits. These properties position it as a potential candidate to be part of useful disease observation tools to farmers.

There are different directions that future work will be based on. To test the model on other public benchmark datasets (e.g., PlantVillage) and across-location field datasets we will first and foremost test cross-domain generalisation. Second, we will conduct systematic edge-device assessments on platforms like Raspberry Pi 4 and low-power AI accelerators, quantifying inference latency, memory usage, and energy consumption under realistic field conditions (Gonzalez-Huitron et al., 2021; Demilie, 2024). and, in parallel, we plan to integrate explainable AI techniques (e.g., Grad-CAM based visualisations) to provide transparent, farmer-friendly explanations of model predictions. Third, systematic ablation studies and statistical validation should be conducted to examine the contribution of architectural components such as depthwise separable convolutions, GAP layer and classifier design. Lastly, extension of the work to multi-task, e.g., simultaneously estimating disease severity or multi-spectral/IoT sensor data, could further make it useful in precision agriculture.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript

Competing Interests

Authors have declared that no competing interests exist.

References

Elmore, M. G., et al. (2022). Detection and discovery of plant viruses in soybean by metagenomic sequencing. *Virol J, 19*(1). https://doi.org/10.1186/s12985-022-01872-5

Wang, B., Zhang, C., Li, Y., Cao, C., Huang, D., & Gong, Y. (2023, April). An ultra-lightweight efficient network for image-based plant disease and pest infection detection. *Precis Agric, 24*(5), 1836–1861. https://doi.org/10.1007/s11119-023-10020-0

Parez, S., Lee, J. W., Dilshad, N., & Alanazi, T. M. (2023, January). Towards sustainable agricultural systems: A lightweight deep learning model for plant disease detection. *Computer Systems Science and Engineering, 47*(1), 515–536.

Xu, X., Zhang, Y., Yang, D., Yu, H., Zhou, L., & Cao, H. (2023, May). Recognition of edible fungi fruit body diseases based on improved ShuffleNetV2. *Agronomy, 13*(6), 1530. https://doi.org/10.3390/agronomy13061530

Vadivel, T., & Suguna, R. (2022). RETRACTED ARTICLE: Automatic recognition of tomato leaf disease using fast enhanced learning with image processing. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, *72*(1), 312-324. https://www.tandfonline.com/doi/abs/10.1080/09064710.2021.1976266

Chen, J., Wang, W., Zhang, D., Zeb, A., & Nanehkaran, Y. A. (2021, April). Attention embedded lightweight network for maize disease recognition. *Plant Pathol, 70*(3), 630–642. https://bsppjournals.onlinelibrary.wiley.com/doi/10.1111/ppa.13322

Hassan, S. M., Maji, A. K., Jasiński, M., Leonowicz, Z., & Jasińska, E. (2021, June). Identification of plant-leaf diseases using CNN and transfer-learning approach. *Electronics (Switzerland), 10*(12). https://doi.org/10.3390/electronics10121388

Hassan, S. M., Jasinski, M., Leonowicz, Z., Jasinska, E., & Maji, A. K. (2021, December). Plant disease identification using shallow convolutional neural network. *Agronomy, 11*(12). https://doi.org/10.3390/agronomy11122388

Mohameth, F., Bingcai, C., & Sada, K. A. (2020). Plant disease detection with deep learning and feature extraction using Plant Village. *Journal of Computer and Communications, 8*(6), 10–22. https://doi.org/10.4236/jcc.2020.86002

Agarwal, M., Singh, A., Arjaria, S., Sinha, A., & Gupta, S. (2020). ToLeD: Tomato leaf disease detection using convolution neural network. In *Procedia Computer Science* (pp. 293–301). Elsevier B.V. https://doi.org/10.1016/j.procs.2020.03.225

Sardogan, M., Tuncer, A., & Ozen, Y. (2018). Plant leaf disease detection and classification based on CNN with LVQ algorithm. https://doi.org/10.1109/UBMK.2018.8520559

Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., & Menaka, R. (2020, January). Attention embedded residual CNN for disease detection in tomato leaves. *Applied Soft Computing Journal, 86*. https://doi.org/10.1016/j.asoc.2019.105933

Duong, L. T., Nguyen, P. T., Di Sipio, C., & Di Ruscio, D. (2020, April). Automated fruit recognition using EfficientNet and MixNet. *Comput Electron Agric, 171*. https://www.sciencedirect.com/science/article/pii/S0168169919319787

Gonzalez-Huitron, V., León-Borges, J. A., Rodriguez-Mata, A. E., Amabilis-Sosa, L. E., Ramírez-Pereda, B., & Rodriguez, H. (2021, February). Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4. *Comput Electron Agric, 181*. https://www.sciencedirect.com/science/article/pii/S0168169920331562

Ghaderizadeh, S., Sharifi, A., Zhao, N., Abbasi-Moghadam, D., & Tariq, A. (2021). Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks. *IEEE J Sel Top Appl Earth Obs Remote Sens, 14*, 7570–7588. https://doi.org/10.1109/jstars.2021.3099118

Wei, S., Li, M., Guan, L., Liu, Y., Zheng, X., & Huang, H. (2023, June). DCCaps-UNet: A U-shaped hyperspectral semantic segmentation model based on the depthwise separable and conditional convolution capsule network. *Remote Sens (Basel), 15*(12), 3177. https://doi.org/10.3390/rs15123177

Wang, Y., Wei, Q., Xu, L., Chen, Y., Li, K., & Wang, F. (2021). A depthwise separable fully convolutional ResNet with ConvCRF for semisupervised hyperspectral image classification. *IEEE J Sel Top Appl Earth Obs Remote Sens, 14*, 4621–4632. https://doi.org/10.1109/jstars.2021.3073661

Krause, D., Kowalczyk, K., & Politis, A. (2021). Comparison of convolution types in CNN-based feature extraction for sound source localization. *Institute of Electrical Electronics Engineers*, 820–824.

Parez, S., Lee, J. W., Dilshad, N., & Alanazi, T. M. (2023, January). Towards sustainable agricultural systems: A lightweight deep learning model for plant disease detection. *Computer Systems Science and Engineering, 47*(1), 515–536.

Gole, P., Marwaha, S., Haque, M. A., Bedi, P., & Deb, C. K. (2023, July). TrIncNet: A lightweight vision transformer network for identification of plant diseases. *Front Plant Sci, 14*. https://doi.org/10.3389/fpls.2023.1221557

Li, G., Zhao, Q., Wang, Y., Chang, B., & Yuan, P. (2023, September). PMVT: A lightweight vision transformer for plant disease identification on mobile devices. *Front Plant Sci, 14*. https://doi.org/10.3389/fpls.2023.1256773

Gonzalez-Huitron, V., León-Borges, J. A., Rodriguez-Mata, A. E., Amabilis-Sosa, L. E., Ramírez-Pereda, B., & Rodriguez, H. (2021, January). Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4. *Comput Electron Agric, 181*, 105951.

Zhao, R., Niu, X., & Luk, W. (2018, February). Automatic optimising CNN with depthwise separable convolution on FPGA. *Association for Computing Machinery*, 285. https://doi.org/10.1145/3174243.3174959

Barbedo, J. G. A. (2019, April). Plant disease identification from individual lesions and spots using deep learning. *Biosyst Eng, 180*, 96–107. https://doi.org/10.1016/j.biosystemseng.2019.02.002

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019, June). A comparative study of fine-tuning deep learning models for plant disease identification. *Comput Electron Agric, 161*, 272–279. https://doi.org/10.1016/j.compag.2018.03.032

Howard, A., et al. (2019, October). Searching for MobileNetV3. *Institute of Electrical Electronics Engineers*, 1314–1324. https://doi.org/10.1109/iccv.2019.00140

KC, K., Yin, Z., Wu, M., & Wu, Z. (2019, October). Depthwise separable convolution architectures for plant disease classification. *Comput Electron Agric, 165*. https://doi.org/10.1016/j.compag.2019.104948

Nazki, H., Yoon, S., Fuentes, A., & Park, D. S. (2020, January). Unsupervised image translation using adversarial networks for improved plant disease recognition. *Comput Electron Agric*. https://dl.acm.org/doi/10.1016/j.compag.2019.105117

Chakravarthy, A. S., & Raman, S. (2020, February). Early blight identification in tomato leaves using deep learning. In *2020 International Conference on Contemporary Computing and Applications (IC3A 2020)* (pp. 154–158). Institute of Electrical and Electronics Engineers Inc. https://ieeexplore.ieee.org/abstract/document/9077010/

Darwish, A., Ezzat, D., & Hassanien, A. E. (2020, February). An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm Evol Comput*.

Chao, X., Sun, G., Zhao, H., Li, M., & He, D. (2020, July). Identification of apple tree leaf diseases based on deep learning models. *Symmetry (Basel), 12*(7). https://doi.org/10.3390/sym12071065

Argüeso, D., et al. (2020, August). Few-shot learning approach for plant disease classification using images taken in the field. *Comput Electron Agric*.

Wspanialy, P., & Moussa, M. (2020, November). A detection and severity estimation system for generic diseases of tomato greenhouse plants. *Comput Electron Agric*. https://www.sciencedirect.com/science/article/pii/S0168169920302052

Chen, J., Zhang, D., & Nanehkaran, Y. A. (2020, November). Identifying plant diseases using deep transfer learning and enhanced lightweight network. *Multimed Tools Appl, 79*(41–42), 31497–31515. https://doi.org/10.1007/s11042-020-09669-w

Abbas, A., Jain, S., Gour, M., & Vankudothu, S. (2021, August). Tomato plant disease detection using transfer learning with C-GAN synthetic images. *Comput Electron Agric*.

Wu, Y., & Xu, L. (2021, October). Image generation of tomato leaf disease identification based on adversarial-VAE. *Agriculture (Switzerland), 11*(10). https://doi.org/10.3390/agriculture11100981

Chen, Y., Zhao, Q., & Lu, L. (2021, December). Combining the outputs of various k-nearest neighbor anomaly detectors to form a robust ensemble model for high-dimensional geochemical anomaly detection. *J Geochem Explor, 231*. https://doi.org/10.1016/j.gexplo.2021.106875

Gong, L., Yu, C., Lin, K., & Liu, C. (2022, January). A lightweight powdery mildew disease evaluation model for its in-field detection with portable instrumentation. *Agronomy, 12*(1). https://doi.org/10.3390/agronomy12010097

Thakur, P. S., Khanna, P., Sheorey, T., & Ojha, A. (2022, July). Explainable vision transformer enabled convolutional neural network for plant disease identification: PlantXViT. https://doi.org/10.48550/arXiv.2207.07919

Cheng, Z., Liu, M., Qian, R., Huang, R., & Dong, W. (2022, August). Development of a lightweight crop disease image identification model based on attentional feature fusion. *Sensors, 22*(15). https://doi.org/10.3390/s22155550

Haque, M. A., et al. (2022, December). Deep learning-based approach for identification of diseases of maize crop. *Sci Rep, 12*(1). https://doi.org/10.1038/s41598-022-10140-z

Dai, Y., Li, C., Su, X., Liu, H., & Li, J. (2023, May). Multi-scale depthwise separable convolution for semantic segmentation in street–road scenes. *Remote Sens (Basel), 15*(10), 2649. https://doi.org/10.3390/rs15102649

Liu, Y., et al. (2023, June). A high-precision plant disease detection method based on a dynamic pruning gate friendly to low-computing platforms. *Plants, 12*(11). https://doi.org/10.3390/plants12112073

Gao, H., Zhu, M., Wang, X., Li, C., & Xu, S. (2023, June). Lightweight spatial-spectral network based on 3D-2D multi-group feature extraction module for hyperspectral image classification. *Int J Remote Sens, 44*(11), 3607–3634. https://www.tandfonline.com/doi/abs/10.1080/01431161.2023.2224099

Hu, Y., Tian, S., & Ge, J. (2023, October). Hybrid convolutional network combining multiscale 3D depthwise separable convolution and CBAM residual dilated convolution for hyperspectral image classification. *Remote Sens (Basel), 15*(19), 4796. https://doi.org/10.3390/rs15194796

Shaheed, K., et al. (2023, November). EfficientRMT-Net—An efficient ResNet-50 and vision transformers approach for classifying potato plant leaf diseases. *Sensors, 23*(23), 9516. https://doi.org/10.3390/s23239516

Demilie, W. B. (2024, January). Plant disease detection and classification techniques: A comparative study of the performances. *J Big Data, 11*(1). https://doi.org/10.1186/s40537-023-00863-9

Zhou, H., et al. (2024, March). Identification of leaf diseases in field crops based on improved ShuffleNetV2. *Front Plant Sci, 15*. https://doi.org/10.3389/fpls.2024.1342123

Zhou, H., et al. (2024, June). Maize leaf disease recognition based on improved convolutional neural network ShuffleNetV2. *Plants (Basel), 13*(12), 1621. https://doi.org/10.3390/plants13121621

Yuan, Y., Sun, J., & Zhang, Q. (2024, November). An enhanced deep learning model for effective crop pest and disease detection. *J Imaging, 10*(11), 279. https://doi.org/10.3390/jimaging10110279

Peer-review history:
The peer review history for this paper can be accessed here:
https://pr.sdiarticle5.com/review-history/155117