**Journal of Deep Learning, Computer Vision and Digital Image Processing**
**E-ISSN: 2986-8939 P-ISSN: 2986-8920**
**Vol. 4, No. 1, March 2026**
**DOI. https://doi.org/10.61255/decoding.v4i1.849**

# Deep Learning for Precision Farming: MobileNetV2-Based Soybean Disease Recognition

**Muhammad Usman Javeed**✉
(Department of Computer Science, COMSATS University of Islamabad, Sahiwal, Pakistan)

**Muhammad Nauman**
(School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu, China)

**Shafqat Maria Aslam**
(School of Artificial Intelligence and Computer Science, Shaanxi Normal University, Xi'an, China)

**Yousuf Ramay**
(Department of Computer Science, Cholistan University of Veterinary and Animal Sciences, Bahwalpur, Pakistan)

**Muhammad Hamza**
(Department of Computing, University of Okara, Okara, Pakistan)

**Saifullah**
(Department of Zoology, Wildlife and Fisheries, University of Agriculture, Depalpur Campus, Okara, Pakistan)

OPEN ACCESS

**ARTICLE HISTORY**
Received: January 28, 2026
Revised: February 21, 2026
Accepted: March 18, 2026

**KEYWORDS**
Convolutional neural networks;
Image classification;
Lightweight deep learning models;
MobileNetV2;
Soybean disease detection;

## ABSTRACT

**Purpose** – Soybean is a globally significant legume, yet it remains underutilized in Pakistan, where low farmer adoption, limited institutional support, and diseases such as Charcoal Rot and Mosaic Virus widen the yield gap. Visual diagnosis is subjective and slow. This study develops a robust deep learning model for early, accurate classification of soybean leaf diseases.

**Methods** – The SDD-2025 dataset comprises 2,000 soybean leaf images from field research stations at the University of Okara and Agriculture University Faisalabad: Charcoal Rot (666), Mosaic Virus (666), and Multiple Pests (668). Images were resized to 256×256 and split 60:20:20 for training, validation, and testing. A pretrained MobileNetV2 network was trained with data augmentation, the Adam optimizer, a batch size of 16, and 6 epochs.

**Findings** – The model achieved 96.14% classification accuracy (3.86% error). Training and validation accuracy rose rapidly after the first epoch and stabilized by the fifth.

**Research implications** – The data cover three categories and limited environments and soybean varieties, so generalization to other regions and field conditions remains untested. The model has not been evaluated in field or on-device settings. Its lightweight architecture nonetheless suggests potential for low-cost, timely disease identification by farmers.

**Originality** – This study offers a regionally collected soybean disease dataset and a MobileNetV2 baseline for Pakistani conditions, where such resources are scarce. Future work should add healthy leaves, more disease classes, and varietal and environmental diversity, and test on-device performance.

**Correspondence Author**: ✉ usmanjavveed@gmail.com
**To cite this article** : Javeed, M. U., Nauman, M., Aslam, S. M., Ramay, Y., Hamza, M., & Saifullah. (2026). Deep learning for precision farming: MobileNetV2-based soybean disease recognition. *Journal of Deep Learning, Computer Vision and Digital Image Processing*, *4*(1), 103-116. Doi. https://doi.org/10.61255/decoding.v4i1.849

*This is an open access article under the CC BY-SA license*

# INTRODUCTION

## Background of the study:

Agriculture Department is the main supplier of food to a rapidly growing population. Agriculture diseases, nevertheless, endanger the continuity of this highly prized resource. Agricultural production losses and forest economy are immense. Farmers can recover some portions of $11 million losses caused by soya bean rust (a plant disease of soybeans) just by cleansing 20% of the infected crop [1]. Soybeans are pressed for oil and used-up for their protein as an Asian staple. Soya bean ingredients are increasing in importance as food and drink companies attempt to reduce their expenses while holding their protein supplies for human intake steady or improving them. Soybeans have been long held in esteem as a top-notch oil seed based on their great protein calm and other alimental values (abundant mineral salts and vitamins A, B, and D). So far, we have found that Pakistan Usage consumed 1.09 million tons of pure oil, 29% of which was self-produced domestically and 71% was imported. Every environment, at every stage of growth, and under every mechanical, physiological, and biological stress can retard the regular growth and expansion of oil seeds crops. A combination of several abiotic and biotic factors such as pollutants, pests, and diseases represent serious though variable threats to cropping depending upon region, season, and affected plant component. Due to its high protein complacent (40 percent), content for oil (20 percent) and meal industry based on the seed, soybean is conveniently used in human and animal nutriment and even in industrial applications. Soya bean oil has been pressed and tested and found to have no toxins. Unsaturated fatty acids that are rich in the oil are responsible for its lowering effect on cholesterol levels [2]. Regrettably, edible oil production in Pakistan is not enough to meet the demands. Palm oil is imported from Malaysia while soybean oil is imported from America and Argentina. Importation of edible oil is an expensive undertaking that is valued at 1 billion and is second most expensive import after fossil fuels. In Pakistan, effort is being made in cultivating non-conventional oil seed crops like soybeans and sunflowers in an attempt to meet the growing demands in the country that cannot be met by conventional oil-seed crops like rape seed and mustard ray. A shortage of edible oils has existed since the 1970s, despite the fact that per capita consumption and local edible oil production have both increased rapidly along with the global population.

Macrophomina phaseout Carbon rot in soybean is due to fungus Architectonic sabbatical and Sclerotic sabbatical [3], [4]. Plants with weaker vitality had their seeds infected with diseases like chromium, cryptosporidium, and metallic element. Soybean seeds were infected with macrophage phaseolina due to a precipitation deficit along with excessive heat during early soybean growth. [3] Therefore, there are countless methods for plant pathology detection, and early catching and diagnosis of works simplicity is imperative for the implementation of appropriate measures. Where there are no observable symptoms or symptoms are not manifested until too far gone to do any good, a detailed analysis, frequently involving powerful microscopes, must be conducted! In other instances, signs can only be manifested in invisible areas [5]. Symptoms due to diseases can be manifested on leaves, stems, fruits, seeds, etc. This study focuses on detection and soybean plant disease organizing via signs borne by diseases on leaves of plants. Diagnosis, or at least preliminary screening, is almost invariably done by a human person. [5] Conventional crop disease detection is done by skilled pathologists and producers, who look at leaves to spot symptoms. This procedure is manual work intensive, costly, and inconsistent due to the fact that human detection is relied upon. The value of this process may also be eroded due to environmental distractions as well as the result of optical illusions such as that due to parallax effect. Thus, instead of time-consuming and possibly harmful eye assessments of plant health, sensing remote sensing application have also been found to be a consistent method for separating ill from healthy leaves. Soybean diseases have been automated with digital image processing [8, 9], pattern recognition [6], and computer vision as per a series of researchers. Furthermore, a series of various varieties of crops and fruits are investigated with the latest techniques, such as grapes, pomegranates tomatoes, and maize. Convolutional neural networks have been applied recently in object identification and image categorization amongst others. Deep neural networks (DNNs) like the convolutional kind are increasingly being utilised to process images in ways that mimic the human visual system. There have been many proposed CNN designs for

application in object perception. The LeNet and Alex Net models are among those used as standards for diverse tasks [7].

Soybean (Glycine max) is leguminous growth in the worldwide, serving as a valuable beginning of protein and oil for human intake and animal feed. However, soya bean production faces significant challenges from various biotic and abiotic stressors, including fungal, bacterial, and viral diseases that affect crop yield and quality. Among these diseases, leaf diseases such as soybean rust, Septoria leaf spot, and bacterial blight are particularly detrimental, causing substantial economic losses to farmers and posing threats to global food security. Efficient approaches to disease management such as, crop rotation, chemical medication, and plant propagation for resistance. Mannual methods methods of disease identification rely on eye's r by agronomists and laboratory-based techniques such as microscopy and pathogen isolation, a process that can demand significant time, manual work and individual interpretation.

**Literature review:**

Current and emerging applications of CNNs and deep learning architecture in farming are discussed here. Image process and machine basic cognitive process approaches were employed to categories plant diseases before more in-depth research was conducted. The following actions are typical for such systems: For starters, a digital camera is used to take a picture. Prior to advanced analysis, an image typically undergoes several preprocessing operations, including enhancement, segmentation, filtering, and color space conversion. Once these steps are completed, the most relevant features are extracted and provided to a classifier for decision-making [8]. Thus, the effectiveness of the classification process is strongly influenced by the selected preprocessing and feature extraction strategies. Recent literature highlights that models trained on publicly available datasets are capable of achieving progressive performance. In comparison, Convolutional Neural Networks (CNNs) represent a deep supervised architecture that independent learns hierarchical features directly from raw data.

CNN has surpassed the progressive in nearly all important categorization challenges in recent years. Its architecture allows for simultaneous feature extraction and categorization [9]. technology such as CNN, AV, SP, etc. Many researchers have continuously constructed Convolutional Neural Networks (CNNs), which are a specialized form of neural networks that has found widespread use in pattern recognition tasks. CNNs were initially motivated by the work of Hubel and Wiesel in 1962 [10]. For CNNs to be shift-, scale-, and distortion-invariant, their architects incorporate the following three concepts: the local receptive field, the general weight, and the spatial or temporal displacement [11]. A wide variety of CNN models have been designed and applied in the domain of object recognition. There's the LeNet, the AlexNet, the GoogLeNet, and so on. LeNet was the first convolutional neural network (CNN) architecture designed by LeCun et al. [11] for recognizing handwritten digits. After fully integrating MLP, it features two convolutional layers and two additional sub-layers. Convolutional neural networks (CNN) have been proposed by certain researchers for use in leaf recognition and plant disease categorization. In order to recognize plants from pictures of their leaves, Atabay devised a convolutional neural network design. There are five distinct components to the proposed architecture.

After each pool layer, we use either the Regenerative stimulation functions such as ReLU (Rectified Linear Unit) and ELU (Exponential Linear Unit), combined with the Max-Pooling layer. In total, 32 plant species and 1,907 samples were analyzed in the Flavia [12] and Swedish leaf databases using the suggested approach. Images of individual leaves captured against a white background make up the database's contents. All input pictures are grayscale and 160 pixels square. Classification accuracy for the model was 97.24%, while accuracy for both data sets was 99.11%. Comparing the suggested architecture for leaf classification tasks based on CNNs, where cutting-edge and integrated methods are applied for feature construction and classification of leaves reveals promising results. In their investigation of complex systems, Angie K. Reyes et al. [13] employ deep learning techniques that eliminate the need for manually generated parts. The suggested architecture consists of 5 Conv layers and 2 FC ones.

The convolutional neural network (CNN) was taught using 1.8 million photos from the 2012 ILSVRC dataset, and it employed a finite element technique to transfer recognition learning from the broad domain to the narrow task of detecting vegetation. A dataset consists of photographs of plants or plant parts collected in the wild or under controlled conditions. A 0.486 accuracies were the mean they achieved. Sarada P. Mohanty et al. classified plant illnesses using preexisting CNN architectures as Alex Net and Google Net [14]. CNN training involved the use of damaged plant images in order to recognize 14 crop species and 26 distinct diseases and healthy plant leaves from a publicly available dataset of photographs collected under controlled conditions. (or absence). Overall, the model was accurate to the tune of 99.35%. However, this model's accuracy plummeted to 31.4% when evaluated on a batch of photographs captured in operating under environmental conditions different from the training dataset. General, the breadth of CNN demonstrates the possibility of plant disease classification. Uuz et al. proposed the use of a Deep Convolutional Neural Network (DCNN) to detect olive peacock spot and aculeus olearius diseases. In total, there are 3,400 samples of olive leaves, divided into three groups: healthy, olive peacock spots, and aculeus olearius illness. DCNN model is utilized as the basis for the VGG16 and VGG19 architectures. By applying the RMSProp optimizer in combination with Stochastic Gradient Descent (SGD), network performance is enhanced. Esgario et al. presented a CNN-based multi-objective network. Purpose: To quantify the biotic stress experienced by coffee leaves. Through the use of computational trials, a more precise and trustworthy detection system was created. With a multi-problem CNN network built on the ResNet50 framework, we were able to attain recorded 86.51% accuracy for severity assessment and 95.24% for biotic stress categorization. Thus, this approach is appropriate for detecting and measuring biotic stress in coffee farms. Escario et al. [15] suggested using a Convolutional Neural Network (CNN) that had been pre-trained to identify plant illnesses.

DenseNet201, ResNet101, Google Net, VGG19, Visual Pure Mathematics Group 16 (VGG16), and Alex Net are all examples of pre-trained CNN models. The authors chose four species that were affected by ten diseases: leaf spot, Yellow Vein Mosaic Virus, Citrus Canker, Orange Indian Ant, Brownish Spot, Cercospora Leaf Spot, Two Months, Epilachna Blight, Small Leaf Blight, and Tobacco Mosaic Virus (TMV). Prediction scores and picture classifications were continuously assessed for each disease category. A Bayesian deep learning approach was proposed by Hernández et al. [16] to identify plant diseases. In terms of classification performance, Bayesian reasoning outperforms state-of-the-art approaches.

The posterior density function is evaluated with respect to plant disease identification, and the prediction uncertainty is quantified. Chen et al. first proposed using leveraging DCNN transfer learning in order to identify plant leaf blight. For evaluation purposes, they choose Inception and VGGNet modules that had already been trained on ImageNet. Experiment results showed that 92.00% accuracy was achieved while classifying photos of rice plants. Extraction of high-dimensional features for categorization allows for improved experimental outcomes in terms of both the user's personal data and the public data set. This allowed for a disease categorization confirmation accuracy of 91.83 percent. A CNN-based approach was recommended by Lee et al. [17] to identify Ginkgo leaf diseases.

The models Inception V3 and VGGNet-16 are used together. Accuracy of 98.44% was attained in controlled laboratory settings, with the VGG model providing an additional 92.19 percent. But as processing time grows longer, data redundancy becomes more of a problem. An accuracy of 96.19 percent was achieved by the proposed categorization model.

**Table 1**. Summary of Related Work

| Remark No | Research Methodological analysis | Illness Work | Dataset Collect | Accuracy % |
|---|---|---|---|---|
| [62] | R-CNN | Virus disease Frogeye leaf spot Bacterial null | synthetic soybean leaf disease image dataset | 83.34 |
| [63] | CNN-SVM | Grade 0<br>Grade 1<br>Grade 2 | Hyper spectral images of a soybean leaf sample | 96.67 |

| Remark No | Research Methodological analysis | Illness Work | Dataset Collect | Accuracy % |
|---|---|---|---|---|
| [64] | HDL-AOA | Beanhalo blightPythium, Rhizoctonia rootrot, Anthracnose | SoybeanLarge DataSet | 98.23 |
| [65] | Alex Net and Google Net CNNs | bacterialblight, frogeyeleaf spot (FLS), brown spot, | Maharashtra, India | 98.75 |
| [66] | (CNN) model | soybean leaf diseases, soybean pod, soybean flower | ICAR (Indian Institute - Soybean Research) | 99.56 |
| [67] | PCR (QPCR) method | Brown spot Bacterial blight | USDA National Institute of Food and Agriculture | 98 |
| [68] | PCR | Fusarium root rot, Fusarium sporotrichioides, Amino acid max | root rot were collected from fields in Manitoba, Canada, in 2017 | |
| [69] | spectral analysis and linear discriminant analysis (LDA) | Phakopsora pachyrhizi (Asian soybean rust) | greenhouse State University | 93 |
| [70] | CNN model VGG19, Google LeNet, Dense121, Xception Net, LeNet, and ResNet50 | Anthracnose, Bacterial Blight, Soybean Mosaic Virus, Phytophora Rot, Rust, and Brown Spot | Image Database of Plant Disease Symptoms | 98.14 |
| [71] | CNN | Healthy Leaf Septorial leaf blight Frogeye leaf spot Downy Mildew | The dataset for training is downloaded fromPlantVillage (Hughes, Salathe, and others 2015) | 99.32 |

With the advancement of computational systems, more computer vision approaches have been applied to plant disease classification. Agricultural prediction and categorization problems [18] are progressively being tackle with the use of artificial powerfulness and data mining method. Image process know-how were employed by multiple groups of person to recognize species and diagnose diseases by recover the Region of Interest. (ROI). Decision Tree (DT), k-Nearest Neighbor, Support Vector Machine (SVM), utilized, and their performances were plumbed using tenfold cross-validation [19]. The recent emergence of deep learning within the realm of machine learning (ML) has yielded groundbreaking advancements in a variety of scientific disciplines, including computer vision, drug design, and bioinformatics. Without the need for manually constructed features, raw data. The past few years have witnessed rapid growth in deep learning-related methods in the field of computer vision [20].

**Gap analysis:**

However, despite the potential of deep learning in plant disease diagnosis, there are challenges to overcome, including the availability of annotated datasets, model generalization across different environments and disease severities, and interpretability of model predictions.

**Rationale of the study:**

In deep learning techniques for automated disease diagnosis in agriculture. These methods offer the potential to revolutionize disease detection by analyzing large-scale image datasets and extracting discriminative features from plant images with unprecedented accuracy and efficiency. Deep learning, in particular, has achieved significant happening in a range of computer imagery applications such as, including object recognition, segmentation, and classification for automated

plant disease diagnosis. Tackling these issues calls for collaborative efforts between different fields of plant pathologists, computer scientists, and agronomists to build a slid and reliable deep learning-based classification systems for soybean leaf diseases. By leveraging the potential of technology, it is possible to improve disease management practices, ensure effective resource distribution and ultimately enhance soybean productivity and food security globally.

**Purpose or Hypotheses of the study:**

The objective of this thesis is to use a deep learning-based classification system for identifying soybean leaf diseases from digital images. By leveraging deep learning architectures and techniques, the proposed system aims to accurately classify soybean leaf images into different disease categories, enabling rapid and reliable diagnosis in agricultural settings. Through empirical evaluation and validation using real-world datasets, this research seeks to demonstrate the effectiveness and practical applicability of deep learning in soybean disease management, ultimately contributing to the sustainability and resilience of soybean production systems.

In this inquiry I will give “Method for Soybean simplicity detection from Pakistani.We will develop the Soybean dataset with the help of Agriculture university, Faisalabad sub-campus okara’s pathologist.

We will develop an efficient deep learning approach for soybean diseases detection.

We will have performed the existing methods in term of diagnose and cross validation.

We will develop a system which work fast with maximum accuracy.

The organization of this material is as follows: Section 2 explores prior research, Section 3 outlines the experimental framework and methodological approach, Section 4 presents the findings along with analysis, and Section 5 delivers concluding insights and suggests opportunities for future work, followed by the reference list.

## METHOD

**Research Design:**

Artificial Intelligence is rooted in self-learning algorithms. As information is more forthcoming with regard to artifacts, the algorithms continuously evolve [12]. This method for addressing the dilemma continues to improve and adapt. Having been inspired by the human brain, this is a framework that can have self-learning algorithms used with it. Trying to imitative the way that human nerve cell communicates, Artificial Neural Networks (ANNs) are sets of inter-connected nodes. Data is stored, fed via positive or negative weights (in effect, both depending on interaction), and then output in this neural network. By their multi-layered construction and ability to perceive patterns, ANNs have a great deal of promise. The definition "deep learning" is taken here to identify the process that occurs in ANNs [21] is used to depict the type of learning that these networks are capable of.

In order to categories coal root, numerous pests, and mosaic virus attacks, this research introduces a deep transfer learning method. The pre-processing and multiplication method is used to address class inequality in the primary database and to generate a more representative sample. Second-level classification of malignant melanoma from benign skin lesions involves automatic feature extraction followed by the implementation of the pre-trained MobileNetV2 model. The process flow for the suggested solution is depicted Fig. 1.

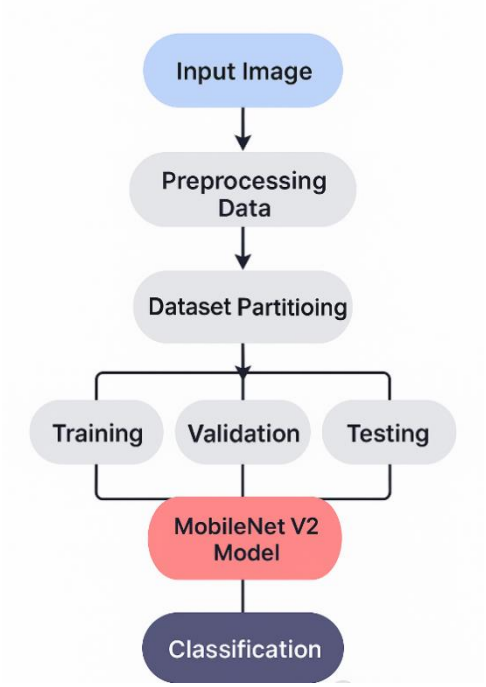



**Figure 1**. Proposed method flowchart

**Population and the methods of sampling Instrumentation:**

Deep learning models used to perform dataset operations on which tuned for the purpose of accuracy. Soybean lead disease data set was generated by researchers from Okara University and the University of Agriculture in Faisalabad, Okara. The following is a across-the-board breakdown of the database:

| No | Disease | Size |
|---|---|---|
| 1 | Charcoal Root | 666 |
| 2 | Mosaic Virus | 666 |
| 3 | Binary Pestis | 668 |

**Table 2**. Summary of SDD-2025 dataset

Several organizations contributed data from various soybean' fields of different diseases to the SDD-2025. The dataset includes 2000 disease images. We divide all these datasets into three classes: charcoal root, Mosaic virus, and multiple pests. The images of the benign class were chosen haphazardly from the entire set of images (see Figure 2 for some samples and Table 1 for information)

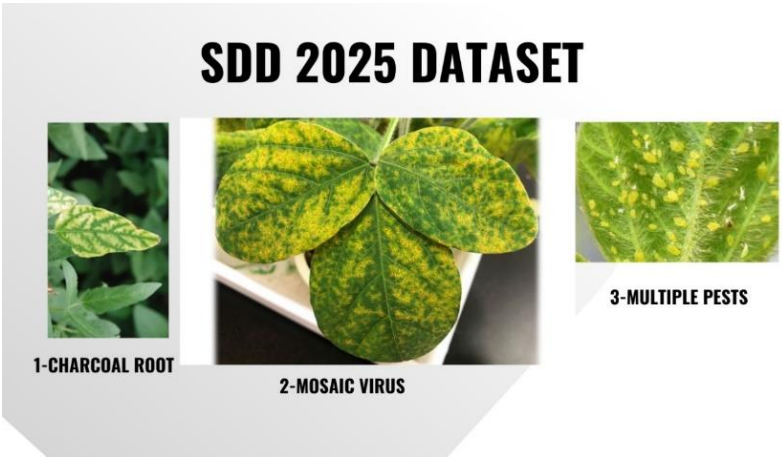



**Figure 2**. All diseases shown

**Instrument:**

A thorough analysis of the MobileNetV2 structure [22] is under way as part of the ongoing research project to tackle the problem of soybean disease classification. The choice of employing a MobileNetV2 model is stuck between a range of different considerations. The basic design of MobileNetV2 is based on that of the previous model, designated as MobileNetV1. Dissociable convolution and linear whirl that are deeply dissociable, as well as inverse residuals, will be discussed in the following section since they are two of the most profound concepts that define the MobileNetV2 architecture.

Other very efficient models that employ depth-wise partitioning are Xception and Shuffle Net, introduced. MobileNetV2 [23] also employs the same depth-wise segmentation employed in MobileNetV1. Deep-resolvable convolution can be exchanged with a normal convolution in two steps. First, we apply a separable convolution individually on each of the feature maps in map-based convolution. The resulting feature map is then acted upon and represented by a second map. Pointwise convolution with a 1 x 1 kernel is computed in parallel across all maps here. Images are acted upon in a standard convolution along their height, width, and channel axes simultaneously, as illustrated in Figure 3. Instead, a separable convolution's initial process is a diligent examination of the height and breadth parameters. Channel dimensions are under its control during the second process, something that is attributed to the factorization aspect of the conventional process.

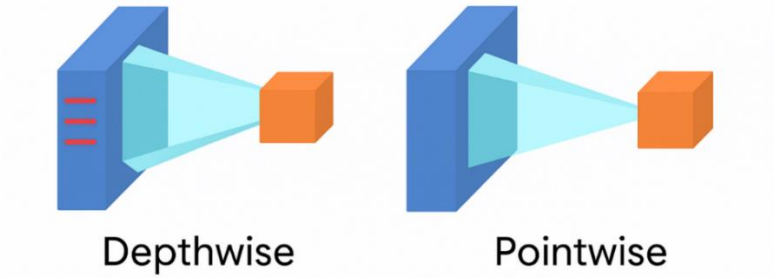



**Figure 3**. Traditional convolution and depth-wise separable convolution

The building pieces that make up the ResNet network are described and contrasted with inverted residues, using three convolution operators and impedance and residual coupling, both blocks achieve similar results. To convert information from the input domain to the intermediate representation and from the intermediate representation to the output domain. In order to filter the interval representation, as depicted in Figure 4, we apply a 3 x 3 filter [24] [22].

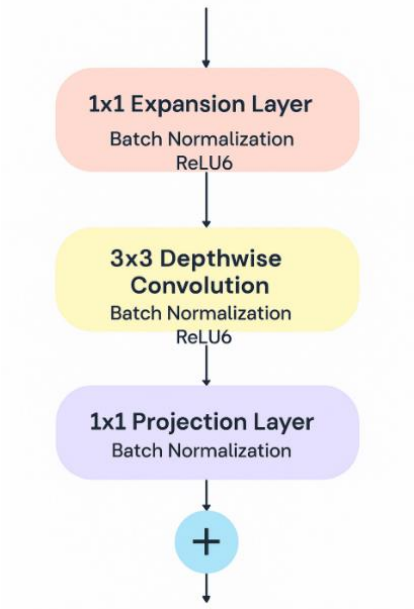



**Figure 4**. Bottleneck residual block

Intra-block convolution output fewer feature maps than the first and final matter block solutions, while reverse convolution's initial and final mixture yield the fewest feature maps. When comparing MobileNetV2 to ResNet, the residual property between the training and test feature maps (channels) is lower. Layer sizes can vary from modest to enormous when many units are joined in a design. Using MobileNetV2's residuary block settings helps cut down the network's memory footprint.

For each edifice block, MobileNet V2 features an extra depth-based and point-to-point conversion layer. Data with numerous channels can be changed into sensors with a small number of maps using the projection layer, a new improver to MobileNetV2 that wasn't present in V1. (channels). The system's block that stores the results of each individual block. After pass through a 1 x 1 diffusive convolution layer, the number of feature mappings (channels) will rise proportionately to the spreading factor used in the ensuant depth-based convolution. The second major addition to the MobileNetV2 architecture is the concept of residual connectivity. A system's gradient flow is aided by the constitution of residual connections. In MobileNetV2, we use ReLU6 as the activation mathematical function for every tier of the architecture to ensure that our models are consistently robust. The projection layer's output, on the other hand, does not contain the activation function. MobileNetV2's full structure consists of 17 residual block blocks, a 1 x 1 sequential convolutional layer, a global average pool layer, and a classification layer. In Figure 5, we can see the finalized MobileNetV2. The pointwise convolution is a 1x1 convolution, which means that it has a kernel size of 1x1.

In this example, the input mental image x is first passed through two convolutional layers. The convolutional layers extract features from the image and reduce its size by half. The features extracted by the convolutional layers are then passed through a fully connected layer with 128 neurons. The fully connected layer combines the features extracted by the convolutional layers and uses them to classify the image into one of the 10 known classes. The class with the highest probability is the class that the image is classified as. Table 3 shows the model and parameters that gave the best results with an accuracy of 96.14 percent.

**Table 3**. Parameters used

| Parameters Used | Values of Parameters |
|---|---|
| Architecture use Mobile Net V2 | Architecture use Mobile Net V2 |
| Types of Education Transfer from scratch | Types of Education Transfer from scratch |
| Layers All Train | Layers All Train |
| Adam Learning Algorithm | Adam Learning Algorithm |
| Default Alpha Rate is Learning rate | Default Alpha Rate is Learning rate |
| ReLu & Sigmoid is Activation Function | ReLu & Sigmoid is Activation Function |
| binary-cross-entropy is Loss Function | binary-cross-entropy is Loss Function |
| Batch Size 16 | Batch Size 16 |
| Epochs 6 | Epochs 6 |

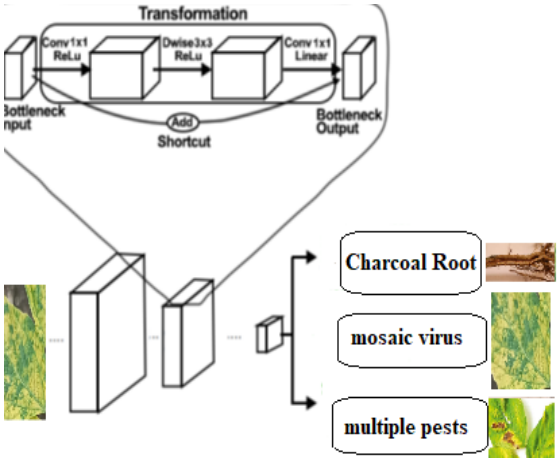



**Figure 5**. Classifier based on mobilenetv2

**Procedures and if relevant, the time frame:**

To obtain high consistence in categorization results and built features, process is applied to all SDD-2025 input images. The CNN approach requires a large amount of iterative training; For this purpose, a large-scale image data set is needed to avoid the peril of overfitting.

All images in the first SDD database are easy in a resolve of 6000 x 4000. The database size is exchanged to 256*256. This will insignificantly reduce the show of the model and speed up process.

**Analysis plan:**

The MobileNetV2 model is trained using the dataset supply in Section 3.1.1. The SDD-2025 dataset employs 1200, 400, and 400 images for training, testing, and validation, respectively. The MobileNetV2 model is trained on the training set and its performance evaluated using test and validation data. Therefore, we split the dataset into three sets: 70% for preparation, 15% for testing, and 15% for proof.

**Scope and/or limitations of the methodology you used:**

The training data set used for a model is very small, which can cause overfitting. A way to mitigate this effect is through the use of a smaller but more efficient network, like MobileNetV2. MobileNetV2 is a design that optimizes the cost associated with errors while maximizing the rate of execution along with memory usage. The low memory consumption is also another attractive feature, aside from the fast execution rate, which increases the feasibility of parameterization and testing.

## RESULTS AND DISCUSSION

**Results:**

An experiment was run with the parameters provided in Table 4 to assess the efficiency of the newly proposed mobilenetv2 architecture: An Adam optimizer with a binary-cross-entropy loss function over the course of 6 epochs, 16 batches, and the standard alpha values. The results of this new approach indicated a 96.19 percent success rate in detecting foliar diseases in soybeans. In terms of soybean leaf blight, the transfer training model achieved a recall of 94% and an accuracy of 97%. The F1 ratings for the three disorders ranged between 94% and 96%. In Table 4, we can see that the gap between our test set's accuracy and the accuracy seen in the leaderboard tests for SDD 2022 is not large enough to warrant abandoning the suggested strategy. Due to the lack of readily available ground truth for SDD-2025, the test is embedded in The leaderboard is 2000 photos.

Area under the receiver operating characteristic curve (AUC), a measure developed by the Kaggle organizer. Figure 6 depicts the accuracy and loss at each time point during training and validation. It demonstrates that there is a rapid improvement in training and validation accuracy after the first session, and that this improvement is maintained until the fifth period. However, both training and validation losses quickly decline after the first period, and then level off after four. By applying the

data augmentation strategy to the training set, the suggested method improves classification scores on the SDD-2025 database.

**Table 4**. Classification results and accuracy on sdd-2025 database

| Presentation Measure # | Charcoal-grey Rot % | Mosaic Virus % | Multiple Pests % | Average Accuracy % | Leader Accuracy % |
|---|---|---|---|---|---|
| Accuracy % | 96.14% | 96.14% | 96.14% | 96% | 96.14% |
| Recall | 84% | 99% | 99% | - | |
| F1-Score | 91% | 91% | 99% | - | |
| Precision | 99% | 83% | 99% | - | |

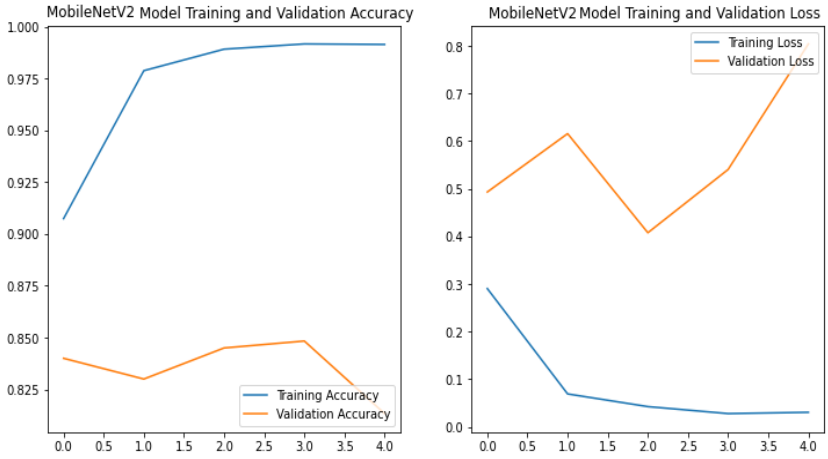



**Figure 6**. (a) Accuracy graph where (b) loss value graph of proposed model

Accuracy, preciseness, the ROK curve, and model correctness can be computed with the confusion matrix, a almighty ML tool. The disorder matrix was used to graphically assess the precision of the classification. This indicates that the categorization is more accurate in the case of MobileNetV2 than for the classes that are dark-colored, while the incorrectly predicted samples are the ones that are shown with light-colored representations. True predictions are signified along the diagonal in the confusion matrix, while false predictions are indicated along the other diagonal.

Figure 7 shows that when the feature based method is applied to the SDD-2025 dataset, the recommended MobileNetV2 subject performs better. The overall consistency of the suggested MobileNetV2 system is 96.14%, and the defect is 3.86% also the Iterative improvement of the several model prototype

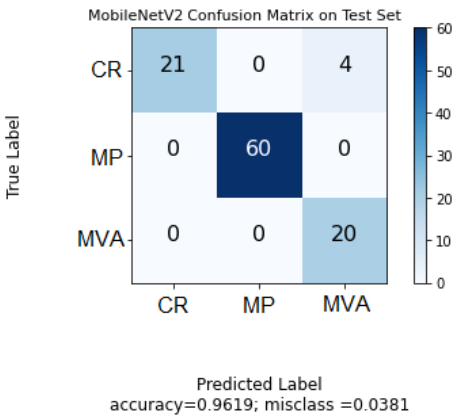



**Figure 7**. Confusion matrix of MobileNetV2

The MobileNetV2 model has a higher area nether the curve (about 96.14%) and attempted to promote success in the validation and test classes. Figure 8 shows the ROK slope as a measuring of the efficaciousness of these up-to-date know-how, with black stand for the ROK curve and red the being of outlier articulation.

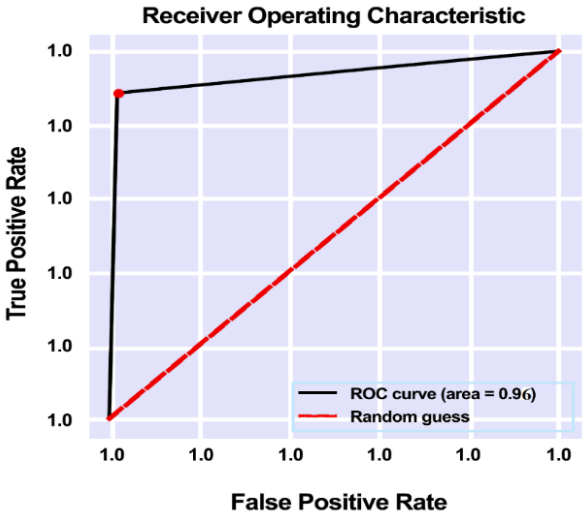



**Figure 8**. Proposed model ROC Curve on SDD-2025

When the feature activity strategy was applied to the training phase, measurement methods such as preciseness, F1-measure, recall, clarity, and ROK curve show that the proposed method conducted highly well in the SDD-2025 database.

**Discussion:**

Utilizing an proposed method, we evaluate our model against the others. There are just fewest studies that have been able to detect soyabean . That's why we may use Google to compare these studies to the few other articles that have been published on the topic. As a result, we check our predictions against these of the study. Table 7 displays the results of a survey given in proposing an MobileNetv2 model for the detection of soyabean disease. Compared to the analysis provided in , which only reached 96% accuracies, the proposed model significantly outperformed expectations, achieving an impressive 96% accuracy. To our knowledge, no previous research has examined any of these factors, either singly or in combination, that reached a higher level of accuracy than our results. Thus, we evaluate our model against the literature. The MobileNetV2 model in table 7 of can accurately classify a wide range of features with a 96% success rate. The findings demonstrated that the suggested MobileNetV2 model outperformed the referenced study [25]. That made it impossible to evaluate the results of this study against those of others.

## CONCLUSION

This study proposed a deep learning model based on the mobilenetv2 model for sensing leaf illnesses in soybean. The dataset was created by photographing soybean effectively detects more than one disease. Compared to conventional foliar disease detection technologies, this one help farmers save money by eliminating future outbreaks of infectious illnesses and increasing soybean yields. Moreover, the suggested model is totally automated, meaning that neither human nor machine is involved in the training process. Furthermore, the proposed model can be utilized for real-time foliar disease diagnosis in plants. The suggested deep learning model, however, shows great promise as a model for early diagnosis of leaf diseases in plants in real time. However, substantial improvement is needed due to the lack of data on the effects of the environment on various plants. When exposed to the same pathogen, various plant species develop unique lesions. Different regions' and varieties' worth of soybean plants were used to collect the data sets. Environmental shifts, variations in soybean varieties, and outbreaks of disease in various soybean varieties are all variables that will need to be handled in the future if we are to eradicate their negative impacts.

## ACKNOWLEDGMENT

The authors declare that this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors thank the researchers of the University of Okara and the University of Agriculture Faisalabad for their support in the collection of the SDD-2025 dataset.

## AUTHOR CONTRIBUTION STATEMENT

MUJ: Conceptualization, Methodology, Software, Formal Analysis, Writing - Original Draft, Project Administration. MN: Methodology, Software, Validation, Visualization, Writing - Review & Editing. SMA: Investigation, Formal Analysis, Data Curation, Writing - Original Draft. WYR: Supervision, Validation, Writing - Review & Editing. MH: Software, Investigation, Visualization, Writing - Review & Editing. SA: Resources, Investigation, Data Curation, Writing - Review & Editing.

## AI DISCLOSURE STATEMENT

The authors used ChatGPT and QuillBot during the preparation of this work for drafting, paraphrasing, and language editing of the manuscript. After using the tools/services, the authors thoroughly reviewed and edited the content as needed and take full responsibility for the content of the publication.

## REFERENCES

[1] R. J. Hay, M. Augustin, C. E. M. Griffiths, and W. Sterry, "The global challenge for skin health," Br. J. Dermatol., vol. 172, no. 6, pp. 1469–1472, Jun. 2015, doi: 10.1111/bjd.13854.

[2] C. Melenotte, M. Drancourt, J. P. Gorvel, J. L. Mège, and D. Raoult, "Post-bacterial infection chronic fatigue syndrome is not a latent infection," Médecine et Maladies Infectieuses, vol. 49, no. 2, pp. 140–149, Mar. 2019, doi: 10.1016/j.medmal.2018.11.006.

[3] J. Rashid, I. Khan, G. Ali, S. H. Almotiri, M. A. AlGhamdi, and K. Masood, "Multi-level deep learning model for potato leaf disease recognition," Electronics, vol. 10, no. 17, p. 2064, Aug. 2021, doi: 10.3390/electronics10172064.

[4] M. Sohail, G. Ali, J. Rashid, I. Ahmad, S. H. Almotiri, M. A. AlGhamdi, A. A. Nagra, and K. Masood, "Racial identity-aware facial expression recognition using deep convolutional neural networks," Appl. Sci., vol. 12, no. 1, p. 88, Dec. 2021, doi: 10.3390/app12010088.

[5] B. G. Blackburn, S. Hosain, T. Greenhow, G. M. Green, C. Lomen-Hoerth, M. GoEVen, and P. S. Mead, "Serious bacterial infections acquired during treatment of patients given a diagnosis of chronic Lyme disease—United States," MMWR Morb. Mortal. Wkly. Rep., vol. 66, no. 23, pp. 607–609, Jun. 2017, doi: 10.15585/mmwr.mm6623a3.

[6] J. Kluger, "Nearly 15% of people worldwide have had Lyme disease, study says," Time, Jun. 14, 2022.

[7] A. Kirubarajan, A. Taher, S. Khan, and S. Masood, "Artificial intelligence in emergency medicine: A scoping review," JACEP Open, vol. 1, no. 6, pp. 1691–1702, Dec. 2020, doi: 10.1002/emp2.12277.

[8] H.-A. Joung, Z. S. Ballard, J. Wu, D. K. Tseng, H. Teshome, L. Zhang, E. J. Horn, P. M. ArnaboEVi, R. J. Dattwyler, O. B. Garner, "Point-of-care serodiagnostic test for early-stage Lyme disease using a multiplexed paper-based immunoassay and machine learning," ACS Nano, vol. 14, no. 1, pp. 229–240, Jan. 2020, doi: 10.1021/acsnano.9b08151.

[9] D. Koundal, A. A. AlZubi, S. Tiwari, K. Walia, J. M.Alanazi, F. I. AlZobi, and R. Verma, "An efficient stacked deep transfer learning model for automated diagnosis of Lyme disease," J. Healthc. Eng., vol. 2022, p. 2933015, Feb. 2022, doi: 10.1155/2022/2933015.

[10] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," J. Physiol., vol. 160, no. 1, pp. 106–154, Jan. 1962.

[11] A. A. R. H. E. N. Hassan Ibrahim, Moustafa El-Taieb, "Dermoscopy versus skin biopsy in diagnosis of suspicious skin lesions," Al-Azhar Assiut Med. J., vol. 15, no. 4, pp. 203–209, 2017.

[12] S. Pandiaraj, T. Jose, S. Velliangiri et al., "Investigation of smart methodologies in Lyme disease detection," in Proc. 2021 Int. Conf. Comput. Commun. Inform. (ICCCI), Jan. 2021, pp. 1–5, doi: 10.1109/ICCCI50826.2021.9402359.

[13] J. Vendrow, J. Haddock, D. Needell, and L. Johnson, "Feature selection from Lyme disease patient survey using machine learning," Algorithms, vol. 13, no. 12, p. 334, Dec. 2020, doi: 10.3390/a13120334.

[14] S. I. Hossain, J. d. G. de Herve, D. Abrial, R. Emillion, I. Lebertb, Y. Frendo, D. Martineau, O. Lesens, and E. M. Nguifo, "Expert opinion elicitation for assisting deep learning based Lyme disease classifier with patient data," arXiv preprint arXiv:2208.14384, Aug. 2022.

[15] D. Chumachenko, P. Piletskiy, M. Sukhorukova, and T. Chumachenko, "Predictive model of COVID-19 epidemic process using machine learning approach," CEUR Workshop Proc., vol. 12, no. 9, 2021.

[16] S. Akbarian, M. P. NeEVer, C. B. Russell, T. Cawston, L. Moreno, S. N. Patel, V. G. Allen, and E. Dolatabadi, "A computer vision approach to identifying ticks related to Lyme disease," IEEE Access, vol. 10, pp. 8831–8841, 2022, doi: 10.1109/ACCESS.2021.3138883.

[17] D. Jacob Jerrish, O. Nankar, S. Gite, S. Patil, and K. Kotecha, "Deep learning approaches for Lyme disease detection: leveraging progressive resizing and self-supervised learning models," 2022.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[19] Q. Rao and J. Frtunikj, "Deep learning for self-driving cars: Chances and challenges," in Proc. 1st Int. Workshop Softw. Eng. for AI in Autonomous Syst. (SEfAIAS), 2018, pp. 35–38.

[20] H. Sun and R. Grishman, "Employing lexicalized dependency paths for active learning of relation extraction," in Proc. 34th AAAI Conf. Artif. Intell., vol. 34, no. 03, pp. 1415–1423, Apr. 2020.

[21] F. A. Radtke, N. Ramadoss, A. Garro, J. E. Bennett, M. N. Levas, W. H. Robinson, P. A. Nigrovic, and L. E. Nigrovic, "Serologic response to borrelia antigens varies with clinical phenotype in children and young adults with Lyme disease," Pediatr. Infect. Dis. J., vol. 59, no. 11, pp. e01–e34.

[22] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," IEEE Trans. Med. Imaging, vol. 36, no. 4, pp. 994–1004, Apr. 2017, doi: 10.1109/TMI.2016.2642839.

[23] E. Zhang, "Lyme disease rashes," 2021.

[24] H. Yang and P. A. Bath, "The use of data mining methods for the prediction of dementia: evidence from the English longitudinal study of aging," IEEE J. Biomed. Health Inform., vol. 24, no. 2, pp. 345–353, Feb. 2020.

[25] B. Levine, "Why is it so hard to find a rheumatologist?"