An Efficient Convolution Neural Network-based Novel Framework for Potato Leaf Diseases Classification and Identification

Md Yasir Arafat, *Member, IAENG*, Sawrnali Khatun, Monishanker Halder, Mohammad Nowsin Amin Sheikh, Md. Nasim Adnan and Sk. Shalauddin Kabir

Abstract—**Agriculture is one of the indispensable fields for the survival of mankind. Potatoes also play an important role in agriculture. Several leaf diseases, including early and late blight, have a significant impact on the quality and quantity of potatoes. The manual interpretation of these leaf diseases is time-consuming and inconvenient. Fortunately, potato plants use the appearance of the leaves to detect diseases. The early detection of infections significantly improves productivity. The early recognition of these diseases from leaf images employs various image processing and deep learning techniques, significantly reducing production losses. The Convolutional Neural Network (CNN) is the most popular deep learning method that extensively recognizes leaf diseases from images due to its incredible and marvelous performance. We use several pre-trained deep learning models, including VGG16, MobileNetV2, ResNet50, InceptionV3, Xception, and a proposed novel CNN-based deep learning model on the plant village dataset to classify and identify potato leaf early and late blight diseases. We apply the transfer learning technique to the pre-trained models and employ data augmentation for the proposed model. Compared to these pre-trained models, the proposed novel model offers the lowest loss and highest accuracy for potato leaf disease detection using fewer parameters and layers. It effectively addresses the overfitting and underfitting problems that occur in pre-trained models. It also achieves the best performance with a test accuracy of 99.67% compared to these pre-trained models used in the diagnosis of potato leaf early and late blight diseases.**

Index Terms—**Deep learning, Potato leaf disease detection, classification, pre-trained models, transfer learning, CNN.**

I. Introduction

OUR survival is dependent upon agriculture. Our country boasts a diverse range of agricultural food varieties. Diseases destroy numerous foods and crops. The potato (Solanum Tuberosum) serves as a fundamental dietary staple in our nation. Potatoes play a significant part in advancing our economy and meeting our food requirements. Diverse kinds of disease substantially diminish potato yield. The primary pathogens affecting potato leaves are early blight (Alternaria Solani) and late blight (Phytophthora Infestans). These diseases impact potato plants, resulting in significant economic losses in potato agriculture. To reduce potato product loss, it is essential to identify the disorders that necessitate appropriate treatment. Without knowledge of the specific illnesses affecting potato plants, we will be unable to implement appropriate measures to mitigate production losses. We employ diverse machine learning and deep learning methodologies to identify, locate, and categorize these disorders. CNN is a prevalent technique for the identification and classification of various disease categories. Any machine learning system, such as support vector machine (SVM), random forest (RF), decision tree (DT), logistic regression (LR), or artificial neural network (ANN), can classify potato leaf diseases. However, those machine learning systems require a different procedure for feature extraction. All machine learning methods are required to extract the appropriate features for categorization. Feeding the classifier unprocessed images will reduce its ability to identify the photographs, resulting in diminished accuracy. CNN efficiently addresses the issue by directly extracting features from the images; hence, it ensures optimal accuracy. In CNN, a distinct feature extraction technique is unnecessary. It possesses a superior accuracy rate compared to a conventional classifier and effectively addresses all of these challenges. We utilized CNN to identify the disease on potato leaves due to the exceptional efficacy of the detection method. We conducted a performance comparison between various pre-trained models and a custom model to identify a specific disease within a designated dataset. We utilized five pre-trained models to evaluate performance.

Large datasets like ImageNet initially train pre-trained models, which frequently show excellent performance. However, when applied to new data, they do not consistently deliver high performance. In contrast, using a custom-designed model allows us to address this issue effectively. This approach enables us to fine-tune the model structure by adjusting parameters and layers to suit the specific dataset. As a result, we can pinpoint an ideal model for our particular dataset, which leads to enhanced performance for the targeted class. Our experimentation with five pre-trained models alongside our unique proposed CNN model has confirmed this crucial insight, which is a highlight of our study. We also employed data augmentation to significantly enhance our model's performance, leading to a notable reduction in model loss. In this context, the custom-designed model exhibits superior performance on our specific dataset,

Manuscript received March 2, 2024; revised April 17, 2025.

Md. Yasir Arafat is an Assistant Professor in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (Tel.:+8801743754783; e-mail:arafatcse@just.edu.bd)

Sawrnali Khatun is a Postgraduate Research Scholar in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (e-mail: m180104.cse@student.just.edu.bd)

Monishanker Halder is an Assistant Professor in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (e-mail:m.halder@just.edu.bd)

Mohammad Nowsin Amin Sheikh is an Assistant Professor in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (e-mail:n.amin@just.edu.bd)

Md. Nasim Adnan is an Assistant Professor in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (e-mail:nasim.adnan@just.edu.bd)

Sk. Shalauddin Kabir is a Lecturer in the Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh (e-mail:sks.kabir@just.edu.bd)

achieving the highest accuracy and the lowest loss when compared to the pre-trained model. Pre-trained models do not consistently perform well on specific datasets. Previous studies have not provided a clear explanation for these issues. Therefore, previous studies did not present a comparison between the performance of custom models and different pre-trained models. Previous studies did not illustrate CNN's working process using real-time potato images. We address this gap by separately demonstrating the performance using custom and pre-trained models, providing a comprehensive and well-explained analysis. We also clearly demonstrated and explained CNN's working process using real-time potato image pixels. By utilizing a custom-designed model, we can determine an ideal model for our dataset, which is not achievable with pre-trained models. We have provided a clear explanation of this concept. Our assessment of dataset and model performance involved experimenting with different layers and dropout rates. Our goal is to mitigate the issues of underfitting and overfitting while increasing test accuracy and reducing test loss for potato leaf disease detection. The pre-trained models we employed on our specific potato leaf disease dataset exhibit these two specific issues in detecting various potato leaf diseases. ResNet50 experiences under-fitting, whereas VGG16, MobileNetV2, Inception, and Xception struggle with over-fitting. To address both over-fitting and under-fitting, we have proposed a novel CNN model with fewer parameters and layers compared to these pre-trained models. While increasing the number of hidden layers to address underfitting is relatively straightforward, managing overfitting is more challenging. To combat the overfitting problem, we have employed data augmentation techniques and incorporated dropout layers into our CNN model. Our proposed CNN model has outperformed all the pre-trained models, achieving the highest test accuracy of 99.33% and the lowest test loss of 1.43%.

This research has the following main objectives:

- Develop a deep learning model using CNN for potato leaf disease detection.
- Utilize transfer learning techniques with various pre-trained models.
- Detect and classify different diseases of the potato leaf, such as healthy, early blight, and late blight.
- Enhance the robustness of the model by applying normalization and data augmentation techniques to the dataset.
- Attempt to minimize loss during the testing phase to accurately detect potato leaf disease.
- Mitigate overfitting by using data augmentation strategies and incorporating dropout layers.
- Compare the classification accuracy between different pre-trained models and the proposed CNN model.
- Identify and visualize potato leaf diseases using the proposed CNN model, demonstrating it as the best-performing model.
- Develop a mobile application that integrates the proposed CNN model for the real-time detection and diagnosis of potato leaf disease.

II. Literature Review

Kumar Sanjeev et al. [1] used an ANN classifier for the early prediction of potato diseases based on leaf images. The FFNN model, with an accuracy of 96.5%, serves this purpose. Dr. Tejashree T. Moharekar et al. [2] used a CNN model for potato leaf disease detection and achieved an accuracy of 94.6%. Rabbi Mahuma et al. [3] used the pre-trained DenseNet for potato leaf disease detection with an accuracy of 97.2%. Mosleh Hmoud Al-Adhaileh et al. [4] used a CNN architecture for detecting potato late blight disease, and the model accuracy is 99%. Tahira Nazir et al. [5] used a deep learning method to classify potato leaf disease and achieved 98.12% accuracy. Deep Kothari et al. [6] made a CNN model to find potato leaf disease and tested it against VGG, ResNet, and GoogleNet on the same dataset. The CNN model was 97% accurate for the first 40 CNN epochs. Sindhuja Bangari et al. [7] used a CNN model for the detection of potato leaf diseases with an accuracy of 99.07%. For leaf disease identification, Dr. N.ANANTHI et al. [8] used image preprocessing and image enhancement (CLAHE, Gaussian blur). CNN performed the classification, and the accuracy was 98.54%. A. Singh and H. Kaur [9] used SVM for potato leaf disease detection classification, and the accuracy is 95.99%. N. Tilahun and B. Gizachew ([10]) detected two types of potato diseases, early blight and late blight, using the pre-trained CNN models MobileNet and EfficientNet. With EfficientNet, a higher accuracy of 98% is attained than with MobileNet. Using the transfer learning method, Birhanu Gardie et al. [11] identify potato disease from leaf images. The comparison of different model accuracies is shown in this work using the same dataset, where InceptionV3 acquired the best accuracy at 98.7%. Md. A. Iqbal and K. H. Talukder [12] used image processing and machine learning methods to detect potato leaf diseases, where RF obtained a higher accuracy of 97%. Abdul Jalil Rozaqi and Andi Sunyoto ([13]) use a customized CNN model with 4 convolution layers and 4 MaxPooling layers for potato to recognize the guava leaf disease automatically at 97% for training data and 92% for validation data using 20 batches at 10 epochs. Md. Asif et al. [14] built a customized CNN model named the sequential model and data augmentation technique for potato leaf disease detection. The customized model achieved the best result with an accuracy of 97% compared to pre-trained models. R. A. Sholihati et al. [15] use the deep learning CNN models VGG16 and VGG19 to classify potato leaf disease, and the average accuracy is 91%. Chaojun Hou et al. [16] used the machine learning method for this process, and the SVM achieved the best result with the highest accuracy of 97.4%. Al-Amin et al. [17] used a deep convolution neural network technique to predict potato disease from leaves with an accuracy of 98.33%. Junzhe Feng et al. [18] used ShuffleNetV2 to detect potato late blight disease. Feilong Kang et al. [19] used the machine learning technique to identify potato blight diseases from leaf images. Hritwik Ghosh et al. [20] used convolutional neural networks to recognize and predict potato leaf disease. P. Enkvetchakul and O. Surinta [21] used two deep CNN models, MobileNetV2 and NasNetMobile, with the data augmentation technique. Krishnan et al. [22] detected potato blight disease from leaf images using machine learning and compared the performance of three classifiers: SVM, RF, and ANN. Al-Akkam and Altaei [23] used deep learning techniques to detect potato leaf diseases. To detect potato diseases, Islam et al. [24] used segmentation and multiclass support vector machines. Liu et al. [25] used a deep learning CNN model named AlexNet for the detection of apple leaf disease with an overall accuracy of 97.62%. Arora et al. [26] used a machine learning method named deep forest for the identification of disease from maize leaves and compared it with other methods. The highest accuracy is 96.25% achieved from the deep forest. Zhang et al. [27] used improved deep CNN models, e.g., GoogLeNet and Cifar10 models, to identify maize leaf disease. The obtained accuracy for GoogLeNet is 98.9% and for Cifar10, it is 98.8%. Howlader et al. [28] proposed an eleven-layer deep CNN model to automatically recognize the guava leaf disease, and the achieved average accuracy is 98.74%. Rangarajan Aravind and Raja [29] used transfer learning and data augmentation methods for automated crop disease classification in agriculture. Hassan et al. [30] used the CNN model with transfer learning techniques to identify many types of plant diseases from leaf images. Shrivastava and Pradhan [31] used the transfer learning method to identify diseases in rice plants. Andrew et al. [32] used a model based on deep learning to identify crop leaf disease. Srinivasu et al. [33] identified the skin disease using MobileNetV2. Abuhayi and Mossa [34] classified the coffee leaf disease using CNN. Vasavi et al. [35] detected the crop leaf disease using machine learning and deep learning. Yadav & Jadhav [36] used deep CNN to create medical images for disease diagnosis. Sun et al. [37] used CNN to recognize tea leaf diseases. Kundu et al. [38] used image-processing techniques to identify plant diseases based on leaf appearance. Shama et al. [39] used deep learning techniques to detect plant leaf diseases. Md. Nabobi Hasan et al. [40] used image-processing techniques to detect plant diseases from leaf images. Mustafa Abed et al. [41] used machine learning techniques to predict pan evaporation. Kevin Aliffanova Ardisa et al. [42] used the CNN method to detect vegetables and produced an accuracy of 95.78%. Shumpei Takezaki and Kazuya Kishida [43] used the data augmentation technique to detect abnormal heart sounds. Using hybrid data mining techniques, Vemuri Bharath Kumar et al. [44] predicted and categorized diabetes.

III. Materials and Methods

To reduce time complexity, we developed a deep learning model using CNN for potato leaf disease detection with fewer parameters and layers. This section describes the different layers and activation functions of the CNN. We also utilized pre-trained models such as VGG16, MobileNetV2, ResNet50, InceptionV3, and Xception to classify potato leaf diseases through a transfer learning process.

A. Process of Potato Leaf Disease Classification

The classification of potato leaf diseases involves the identification and categorization of diverse diseases impacting potato plants based on visual symptoms observed on the leaves. We employ several strategies, such as picture augmentation, to increase the dataset's diversity and enhance the model's performance. It encompasses multiple essential approaches, including image acquisition, preprocessing, augmentation, feature extraction, and classification. Applications for feature extraction and classification frequently employ CNN. Figure 1 illustrates the comprehensive stages involved in the classification of potato leaf disease.

TABLE I
NUMBER OF LEAF SAMPLES IN THE TRAINING, VALIDATION, AND TESTING SETS

Label	Category	Number	Training Sample	Validation Sample	Test Sample
1	Healthy	500	300	100	100
2	Early Blight	500	300	100	100
3	Late Blight	500	300	100	100
Total		1500	900	300	300

1) Image Acquisition: The dataset is essential for identifying potato leaf diseases. Initially, an appropriate image dataset must be collected. We employed a curated dataset from Kaggle, accessible at https://www.kaggle.com/datasets/muhammadardiputra/potato-leaf-disease-dataset, for our implementation. Kaggle offers a diverse array of online image repositories. It serves as a prominent repository for several categories of image collections. We possess three categories of potato leaf datasets: those exhibiting early blight, those exhibiting late blight, and those displaying healthy leaves. Both steep and lowland areas harbor Alternaria solani, the principal agent of early blight. The unique angular, oval form of the brown-to-black necrotic lesions is characterized by concentric rings. Numerous dots coalesce and subsequently disperse on the leaves. Phytophthora infestans is the causative agent of late blight in potatoes, impacting tubers, leaves, and stems. As the illness advances, the leaf spots enlarge, proliferate, and ultimately transition from purple-brown to entirely black. Below the leaf surface, a white growth becomes apparent. We randomly selected the example image from our potato leaf disease dataset utilizing Python programming. This is presented in Fig. 3.

We partitioned the dataset into a training set, a validation set, and a test set. The dataset comprises 1500 images: 900 designated for training, 300 for validation, and 300 for testing. All datasets about potato leaf disease include dimensions of 256x256 pixels. Maintaining uniform dataset sizes throughout training is crucial for illness classification with CNN. Inconsistent visual dimensions can cause issues during CNN training. In a CNN, it is essential that each input image is of uniform dimensions to facilitate efficient processing by the network. The datasets are presented in Table I.

2) Image Pre-processing: Prior to the feature extraction of an image, we implement several image preprocessing techniques to enhance performance, including image resizing, filtering, noise reduction, color transformation, data augmentation, normalization, and image segmentation. Photographs of plant leaves are generally characterized by noise after capture. These distorted visuals are exceedingly difficult to identify. Consequently, we must eliminate the noise from the first acquired noisy image collection. It also offers elevated training accuracy. Afterward, we need to resize the image to match its original dimensions. If the image dimensions do not match those specified by the model's inherent code using a framework such as TensorFlow, then an error is likely to occur. Normalization scales pixel values to a uniform range, often between 0 and 1, facilitating algorithmic learning from the data. It can accelerate model inference while diminishing the necessity for model training.

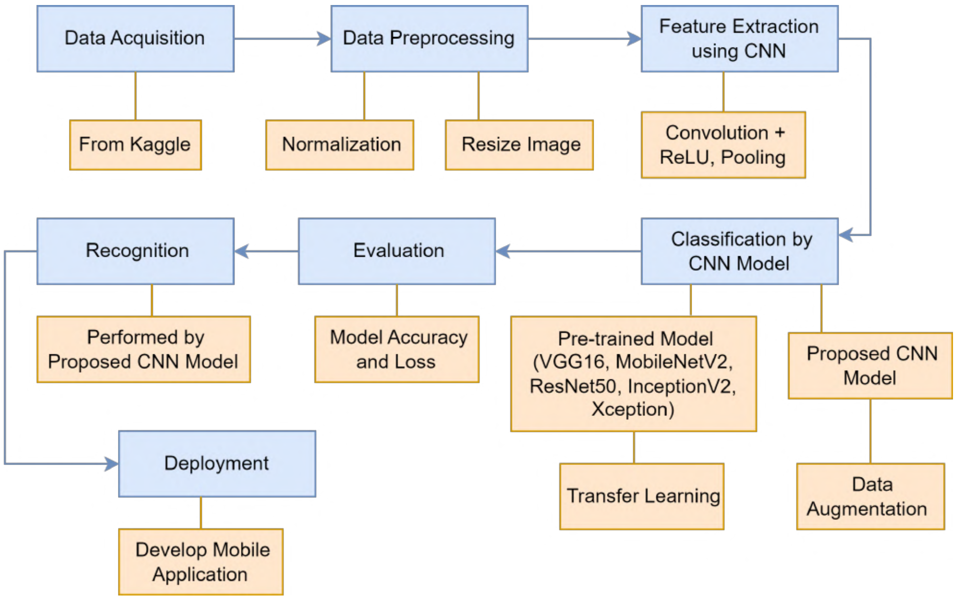

Fig. 1. Flow diagram of all processes for potato leaf disease detection

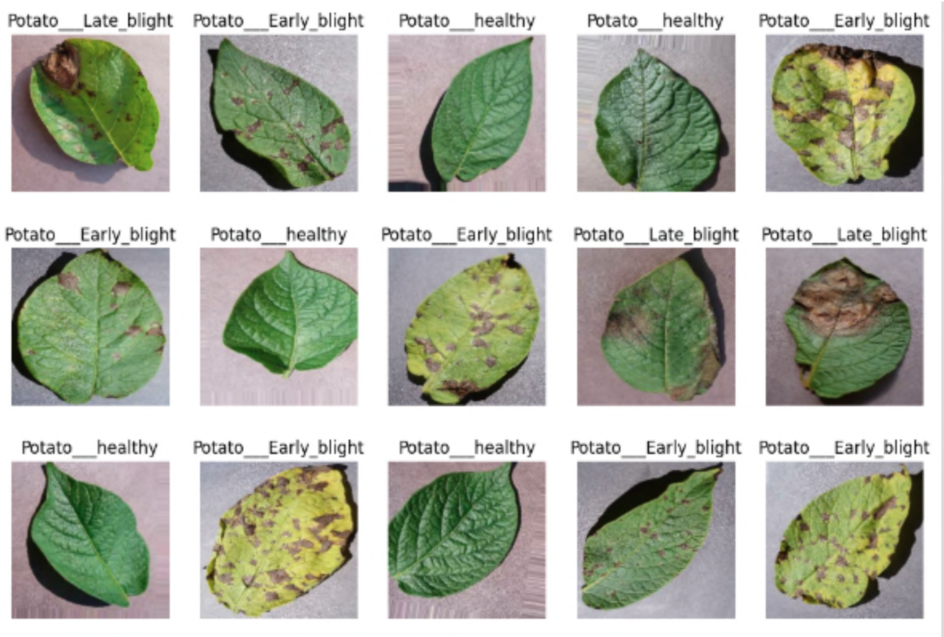

Fig. 2. Potato leaf image samples, e.g., Healthy, Early Blight, Late Blight

3) Image Augmentation: Augmentation involves enhancing the dataset through various approaches. Various approaches, including rotation, flipping, shifting, random brightness adjustment, and zooming, augment the number of images or data in disease categorization. Nonetheless, a notable distinction exists between image augmentation and image preprocessing. Both the training and test sets employ picture preprocessing techniques, while the training data exclusively uses image augmentation techniques. It is impractical to comprehensively represent an image that encompasses every conceivable real-world occurrence for a model. By augmenting the images, we may increase the sample size of our training data and incorporate novel scenarios that may be challenging to identify in reality. The model can acquire knowledge from a broader spectrum of events by augmenting the training data to generalize across various circumstances. Image augmentation is a critical strategy for mitigating the overfitting issue in deep learning. We utilize it to mitigate overfitting concerns and improve classification precision. Multiple data augmentation examples are illustrated in Fig. 4.

We specifically employ this augmentation strategy for training data, not for test data, as the model benefits from an increased volume of training data. We can employ a data augmentation strategy to increase the quantity of the dataset, thereby improving the model's accuracy. If we possess a limited dataset, we can utilize this data. We utilize this information to augment the quantity of photographs. The increased dataset is illustrated in Table II.

4) Feature Extraction: The feature extraction method is essential for recognizing patterns in images that assist in illness identification. This procedure integrates convolutional

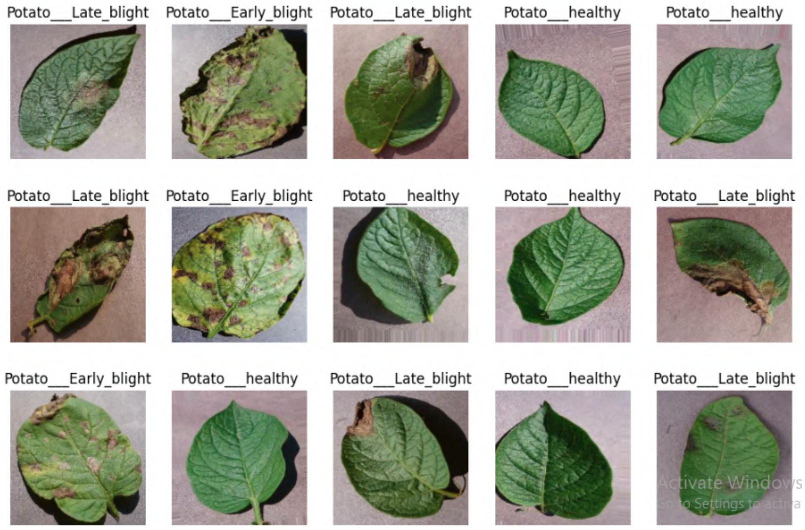

Fig. 3. Potato leaf image samples, e.g. Healthy, Early Blight, Late Blight.

Fig. 4. Potato leaf image samples after augmentation, e.g., Healthy, Early Blight, and Late Blight

TABLE II
NUMBER OF LEAF SAMPLES IN TRAINING, VALIDATION, AND TESTING SETS AFTER AUGMENTATION

Label	Category	Number	Training Sample	Validation Sample	Test Sample
1	Healthy	900	700	100	100
2	Early Blight	1200	1000	100	100
3	Late Blight	‘1200	1000	100	100
Total		3300	2700	300	300

and pooling layers to extract significant features, subsequently advancing to fully connected layers and softmax classification layers for decision-making. The softmax classifier generates predictions from the input data by diminishing dimensionality and removing superfluous information. This reduction does not compromise any crucial or pertinent image features, facilitating more efficient processing without diminishing accuracy. Object recognition and classification utilize the retrieved features as the basis for a feature vector. Convolutional neural networks (CNNs) are distinguished by their ability to directly extract features from source images, in contrast to other neural networks such as Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF), which depend on previously extracted features for classification and do not engage in direct feature extraction. A CNN retrieves information from the input image for classification, which is particularly efficient for image-related tasks. The use of CNNs accelerates learning and enhances the machine learning process by optimizing the feature extraction and classification pipeline. This is emphasized in the initial segment of the CNN operational procedure, as depicted in Fig. 5.

5) Classification: The process of image categorization entails assigning an image to a specific category. Our study categorizes potato leaf diseases into three separate picture classes. To do this, we use a convolutional neural network (CNN), which is recognized as the most efficient method of diagnosing potato leaf diseases, because of its robust feature extraction abilities. The standard classification procedure occurs in the second phase of the CNN workflow, as seen in Fig. 5.

6) Evaluation and Recognition: In machine learning and deep learning, we employed many performance indicators to evaluate the efficacy of potato leaf disease classification tasks or any classification challenges. We employ a confusion matrix to accomplish this. It is a recognized standard for assessing accuracy or error metrics in classification tasks. We compute the performance metrics of accuracy, precision, and recall utilizing a confusion matrix. The confusion matrix presents a summary of actual vs. expected values, including true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) in a matrix format. Evaluating a model's performance is essential as it indicates the model's accuracy and error rate. Various assessment criteria, such as accuracy, precision, and recall, are available for assessing a model's performance. We assess our model utilizing accuracy metrics (test accuracy and test loss). Ultimately, our suggested CNN model is capable of identifying potato leaf disease.

B. CNN

Convolutional neural networks consist of three primary layers: the convolutional layer, the pooling layer, and the fully connected layer (FC) or dense layer. In addition to these layers, the stride, padding, and activation functions are significant concepts in convolutional neural networks.

1) Convolutional Layer: Convolution is the procedure forr integrating a kernel or filter with an image,which involveses summation of the products of related pixels in the kernel and the image. The genuine potato leaf dataset comprises an array of pixels, facilitating a clearer comprehension of the convolution process in the Python code. The pixels of the array are depicted in Fig. 6.

The kernel traverses the entire image. This convolution approach, with a filter, has been employed to extract characteristics from the original photos. Convolution extracts image characteristics from a compact square of input data while preserving the spatial relationships among pixels. Examine a 5-by-5 image as an example of how convolution interprets each image as a matrix of pixel values. A standard neural network links each input neuron to the subsequent hidden layer. The convolution process is illustrated in Fig. 7.

2) Pooling Layer: The pooling layer aims to decrease the dimensionality of the feature map, minimize spatial size, reduce the number of parameters, and mitigate overfitting. This is encompassed by the hidden layer of CNN. CNN uses many pooling techniques, including max, average, and sum pooling. We used the MaxPooling technique. The MaxPooling procedure is illustrated in Fig. 8.

3) Fully Connected Layer: The fully connected layers, or dense layers, constitute the last levels of the network. They extract data from the convolution layer post-convolution to provide output utilizing classifiers such as Softmax, Sigmoid, etc. The final convolutional layer, called the pooling layer, transmits the output to the fully connected layer after it is flattened. The phrase "fully connected" denotes a link between each neuron in the lower layer and every neuron in the higher layer. This procedure is illustrated in Fig. 9.

4) Stride and Padding: The stride refers to the distance that the filter matrix moves across the input matrix. The filters advance one pixel per movement with a stride of 1, and they advance two pixels per movement with a stride of 2. A longer stride yields reduced feature maps. Padding is used to preserve the dimensions of the input and output images. CNN incorporates padding into an image to improve analytical precision. Zero padding involves augmenting our input photos with layers of zeros.

5) Activation Function: Deep learning employs many activation functions for classification, including Sigmoid, TANH, ReLU, and Softmax. We used the ReLU and softmax activation functions.

Sigmoid is used for binary categorization. To distinguish between two distinct categories, we employ the sigmoid function. It employs a probabilistic methodology to make decisions with values ranging from 0 to 1. It works exclusively for positive numbers. It is denoted as in (1).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

TANH serves as an additional activation function. In contrast to sigmoid, TANH has a range of -1 to +1. Using the TANH function, we may address negative values. The TANH activation function possesses superiority over the sigmoid activation function. It functions for both positive and negative values. It is denoted as in (2).

$$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{2}$$

We denote ReLU as a Rectified Linear Unit. We employ an additional process known as ReLU after each convolution operation. We add ReLU after each convolution operation. ReLU utilizes a non-linear mechanism. It substitutes each pixel in the feature map with a zero to eradicate any negative pixel values. It is a linear function that outputs the input directly when it is positive but outputs zero when it is negative. The concealed layer employs it to enhance computational efficiency and mitigate the vanishing gradient issue. It is denoted as in (3).

$$f(x) = \max(0, x) \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } x \geq 0. \end{cases} \tag{3}$$

The dense layer uses Softmax to identify multi-class image groupings. The dense layer must employ the softmax activation function to detect several classes. It computes the likelihood of belonging to each category. We shall determine the actual class according to the highest probability. In a dataset comprising three image classes—healthy leaf, early blight, and late blight—the softmax function yields class probabilities: healthy leaf: 1.5%, early blight: 96%, and late blight: 2.5%, indicating a prediction of early blight's presence. Ultimately, Softmax identifies the image class with the greatest likelihood. In the last layer of our model, we employed Softmax as a classifier for the identification of potato leaf diseases. In a classification problem, the potential classes or categories are represented by y_i, whereas the input features required for predictions are marked by x, as illustrated in (4).

$$\text{Predicted class} = \arg\ max_i P(y_i \mid x) \tag{4}$$

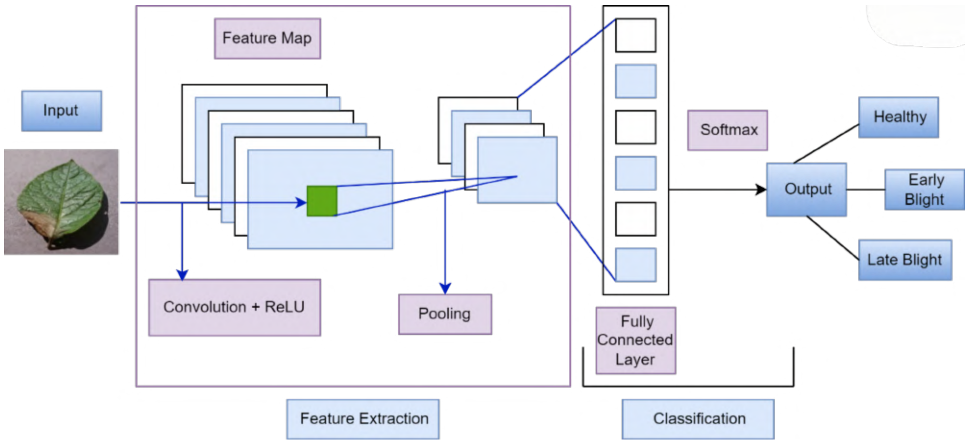

Fig. 5. CNN working process

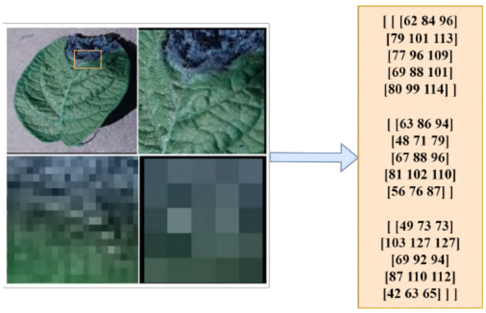

Fig. 6. Array of pixels from potato leaf samples

6) Dropout Layer: We employ dropout layers to enhance test accuracy and mitigate overfitting. This progress is promising and demonstrates the model's efficacy in both the training and testing phases. Dropout, a training methodology, randomly disregards specific neurons. The approach randomly deactivates these neurons. This indicates that the neuron momentarily nullifies its influence on activating subsequent neurons during the forward pass and refrains from conveying any weight modifications to it during the backward trip. A dropout can be applied after both pooling layers, such as MaxPooling2D, and convolutional layers, such as Conv2D. Our model utilized a dropout rate of 0.5 in the fully connected layers.

C. Transfer learning

Effective and precise training of a neural network often necessitates a substantial dataset. Nevertheless, a more extensive dataset may not always be available, in which scenario transfer learning proves to be highly beneficial and pragmatic for enhancing accuracy. It is a method that employs a validated training model. It entails utilizing an established paradigm to address a novel issue. Deep learning and machine learning are presently prevalent owing to their capacity to train deep neural networks with minimal data while attaining superior accuracy. This approach proves to be highly effective for photo classification when it utilizes a small dataset with an appropriate model. We utilize the proficiency of a pre-trained model on an extensive dataset. Moreover, employing a pre-trained model from an extensive dataset diminishes training duration, lessens data prerequisites, and typically improves neural network efficacy, even when data is scarce. We employed transfer learning for the categorization of potato diseases using pre-trained models.

D. Pre-trained Network Models

We used pre-trained network models, such as VGG16, ResNet50, MobileNetV2, InceptionV3, and Xception. All pre-trained models use transfer learning to detect potato leaf disease. To fine-tune the pre-trained models, we removed their final classification layer(s) and added our classification layer(s) with the appropriate number of output units.

1) VGG16 : Only the 16 layers of VGG16 possess weights, distinguishing them from other algorithms that predominantly depend on hyperparameters. VGG16 comprises a configuration of 13 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. The architecture has 13 convolutional layers and three fully linked layers, totaling 16 layers with adjustable parameters. Subsequently, a softmax layer is implemented, followed by two fully linked layers, each including 4096 nodes. We utilized pre-trained network models such as VGG16, ResNet50, MobileNetV2, InceptionV3, and Xception.

2) MobileNetV2: MobileNetV2 is a convolutional neural network with 53 layers. It is a more compact and lightweight network model. A crucial component of MobileNetV2 is the depthwise separable convolutions found inside the inverted residual blocks. MobileNetV2 does not have the conventional, completely linked layers at the network's end that some other neural network architectures do. Instead, it uses a final, fully linked softmax layer and global average pooling (GAP) for classification.

3) ResNet50: ResNet50 is a variant of the ResNet architecture with 48 convolutional layers, 1 max-pooling layer, and 1 average pooling layer. It organizes its 50 layers into 5 blocks, each including a series of residual blocks. Each block integrates a convolutional block with an identity block. Despite being far deeper than VGG, the model's actual weights are fewer than those of the VGG family due to the utilization of global average pooling instead of the fully connected layer.

4) InceptionV3: InceptionV3 comprises 48 layers. These layers comprise convolutional layers, max-pooling layers, fully connected layers, and auxiliary classifiers. The architecture is defined by its heavy use of inception modules,

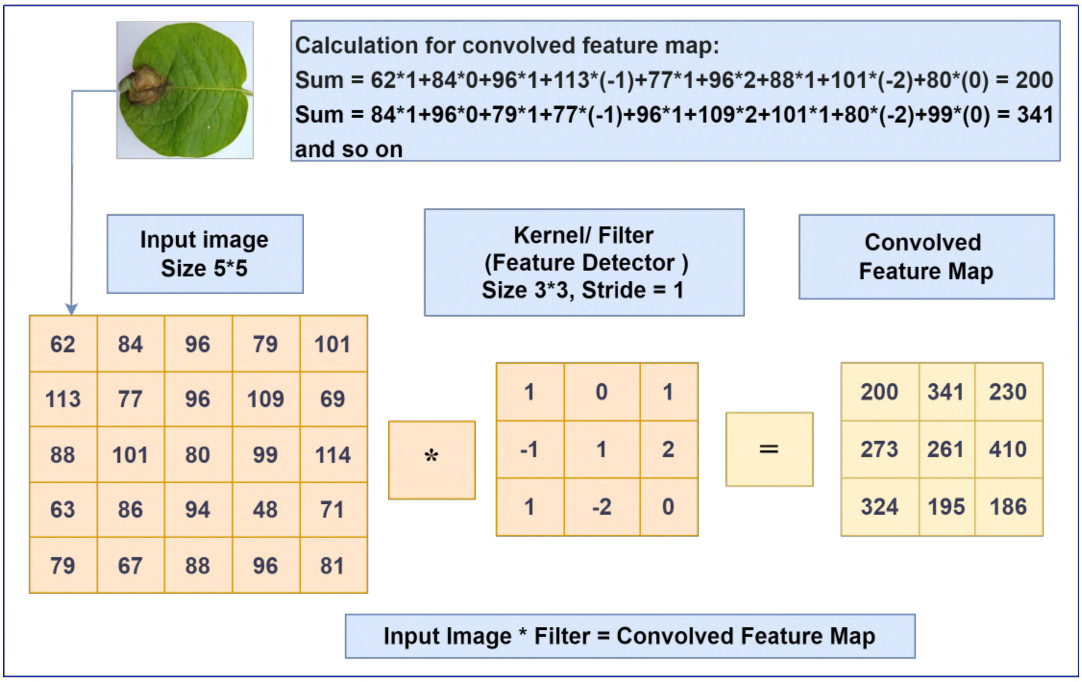

Fig. 7. Convolution Process

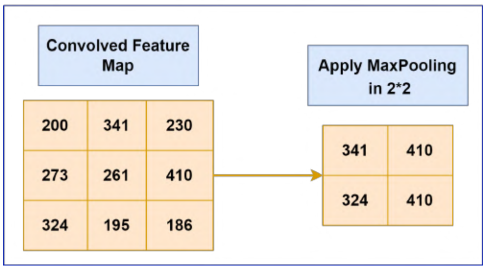

Fig. 8. Applying MaxPooling 2*2 in the Polling Process

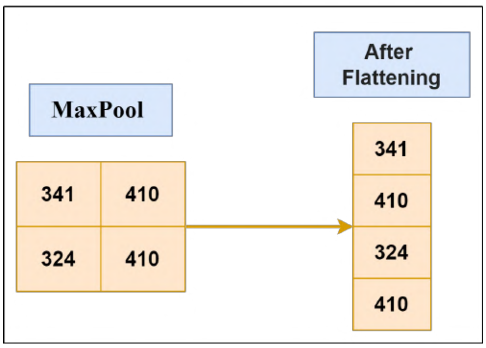

Fig. 9. Flattening Process in Fully Connected Layer

which are made up of many parallel convolutional filters with different kernel sizes.

5) Xception: Xception is a convolutional neural network with 71 layers. Excluding the first and last modules, the other 14 modules, each consisting of 36 convolutional layers, include linear residual connections encircling them. It employs depth-wise separable convolutions instead of traditional pooling layers, such as max-pooling or average-pooling. In the traditional context, it lacks any fully connected layers. It employs a softmax layer and global average pooling for classification instead of a fully linked layer.

E. Proposed CNN Model

We proposed a deep learning model employing convolutional neural networks (CNN). The proposed CNN model has three dense layers and seven convolutional layers. We utilized a Softmax classifier for the categorization of leaf diseases in the final layer. We utilized the ReLU nonlinear activation function in every convolutional layer. This reduced the likelihood of the vanishing gradient problem and converted negative numbers to zeros. We employ a pooling layer after each convolution to reduce computational cost and spatial dimensions. We mitigate the overfitting issue, enhance the efficacy of the activation function, and improve convergence through image downsampling via max pooling. The output layer, also known as the fully connected or dense layer, uses a softmax classifier with three outputs to identify the class of a potato leaf image. The comprehensive architecture of our suggested CNN model is illustrated in Fig. 10 and the experimental setup[45].

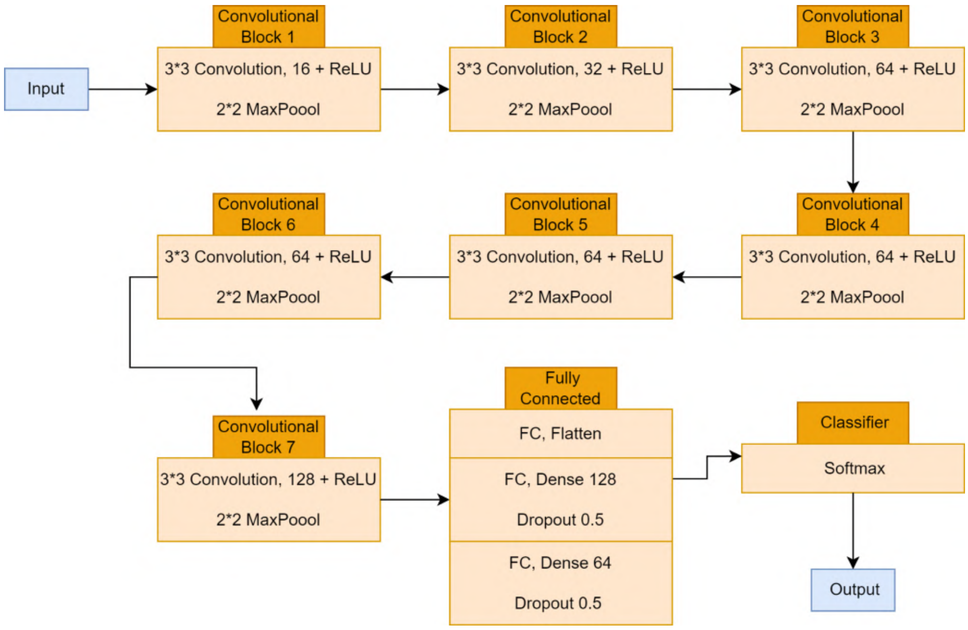

Fig. 10. The Proposed CNN Model Architecture and the experiments [45]

1) Input layer

- Our model's input image size is 256×256.
- Convolution and ReLU operations using CNN directly extract features from this input image.

2) Convolutional and pooling layers

- We employ three sets of convolutional layers with ascending filter sizes (16, 32, 64, and 128) and utilize ReLU activation functions to capture hierarchical features.
- To reduce spatial dimensions, we add MaxPooling layers after each convolutional layer.

3) Flattening layer

- The flattened layer transforms the output derived from the convolutional layers into a one-dimensional vector.

4) Fully connected layers

- We add two fully connected layers with 128 and 64 neurons each and use ReLU activation functions to make feature extraction more in-depth.
- Dropout layers with a rate of 0.5 are inserted after each fully connected layer to mitigate the over-fitting problem.

5) Output layer

- Three neurons make up the concluding layer, which corresponds to the three distinct disease classes: potato healthy, early blight, and late blight.
- We use the softmax activation function to classify three types of potato diseases.

6) Model compilation

- To compile the model, we employ the Adam optimizer in conjunction with categorical cross-entropy loss, a suitable choice for tasks involving multi-class classification.
- Throughout training, we use the accuracy metric to monitor the model's performance.

All of the model architectures are summarized in Table III.

We employ categorical cross-entropy as the loss function and Adam as the optimizer for classification in all models. We employ a batch size of 32 for categorization in all instances. We trained all models over 80 epochs. The hyperparameters utilized for the categorization of potato leaf disease across all models are presented in Table IV.

IV. Experimental Results and Discussion

This study utilizes many pre-trained models alongside a suggested deep-learning CNN model for the detection and classification of potato leaf diseases. We classified potato leaves into three categories: healthy, early blight, and late blight. Given that this is a multi-class classification assignment for potato leaves, we utilized a softmax classifier in the dense layer. All models applied the softmax classifier, maintained a constant batch size of 32, and utilized categorical cross-entropy as the loss function. We trained all models for 80 epochs using Google Colab to identify potato leaf diseases.

We utilized five pre-trained models (VGG16, ResNet50, MobileNetV2, InceptionV3, and Xception) for the identification of potato leaf diseases using transfer learning. Although these models are typically resilient, they may not represent the most suitable option for every dataset. Pre-trained models, originally trained on vast datasets, may exhibit poor generalization to smaller, domain-specific datasets. Consequently, we abstained from employing data augmentation strategies for these five pre-trained models. Notwithstanding their elevated training accuracy, these models failed to adapt effectively to our potato leaf disease dataset. Both the training and validation accuracy demonstrated deficiencies, and their performance regarding validation and test accuracy was disappointing. Evaluating these models with novel data revealed elevated loss performance, indicating potential issues with overfitting or underfitting. These models significantly elevated the test loss for potato leaf disease detection.

To mitigate these issues, we implemented a bespoke CNN model featuring optimized layers and a constrained

TABLE III
SUMMARY OF THE MODEL USED IN POTATO LEAF DISEASE DETECTION

Model	Parameter	Layer	Input Size	Filter Size	Polling	Activation Function	Number of Convolution Layer	Number of MaxPool Layer	Number of Fully Connected Layer
VGG16	138.4M	16	224*224	3*3	2*2 MaxPool	ReLU	13	5	3
MobileNetV2	3.5M	53	224*224	3*3, 1*1	GAP	ReLU	43	-	-
ResNet50	25.6M	50	224*224	3*3	3*3 MaxPool	ReLU	49	5	1
InceptionV3	23.62 M	48	299*299	3*3, 5*5, 1*1	3*3 MaxPool	ReLU	47	4	2
Xception	22.85M	71	299*299	3*3	GAP	ReLU	36	-	-
Proposed CNN Model	233K	10	256*256	3*3	2*2 MaxPool	ReLU	7	7	3

TABLE IV
THE HYPERPARAMETER USED TO DETECT POTATO LEAF DISEASES

Batch Size	32
Epochs	80
Loss Function	Categorical Cross Entropy
Optimizer for Model Training	Adam
Dropout	0.5
Classifier in Output Layer	Softmax

parameter count. Furthermore, to mitigate overfitting, we enhanced the potato leaf disease dataset with techniques such as shifting, rotating, cropping, zooming, and shearing. This enhancement expanded our initial collection from 1500 to 3300 photos. We additionally used a dropout rate of 0.5 in the fully linked layer to help alleviate the overfitting issue. Our model's performance was superior to that of the pre-trained models. It attained a remarkable test accuracy of 99.33% and a test loss as minimal as 1.43%. We substantially enhanced the training and validation accuracy, thereby resolving the challenges encountered by the pre-trained models. The quality and quantity of the dataset, along with the GPU's processing capacity for model classification, determine a model's performance.

Our study presents a graph of training and validation accuracy, as well as the loss, for all models. Fig. 11 shows the accuracy and loss plot for VGG16. Fig. 12 shows the accuracy and loss plot for ResNet50. Fig. 13 shows the loss plot for MobileNetV2. Fig. 14 shows the loss plot for InceptionV3. Fig. 15 shows the loss plot for Xception. The loss plot for the proposed CNN model is shown in Fig. 16.

In Fig. 11, VGG16 attains 70% training accuracy and 79% validation accuracy during the initial epoch. Thereafter, the training accuracy consistently increases to 100% by the fifth epoch, whereas the validation accuracy attains 96%. From the fifth epoch to the final epoch, the training accuracy constantly maintains a level of 100%; however, the validation accuracy exhibits fluctuations. This significant discrepancy between training and validation accuracy is also apparent in the loss graph. Notwithstanding the elevated training accuracy, the model demonstrates diminished validation accuracy, indicating an overfitting issue.

In Fig. 12, ResNet50 initiates with a training accuracy of 42% and a validation accuracy of 43% during the initial epoch. Nonetheless, as training advances, both accuracy metrics demonstrate enhancement. By the concluding epoch, the training accuracy has ascended to an amazing 97%, while the validation accuracy remains at 84%. Both the mean training accuracy and validation accuracy are below average. Unfortunately, the loss remains consistently high for both training and validation across all epochs, making it unsuitable for our potato leaf dataset. The significant gap between the training and validation measures is evident. This model demonstrates inadequate performance in both training and validation accuracy, signifying an underfitting issue.

In Fig. 13, MobileNetV2 initiates with an 85% training accuracy and a 96% validation accuracy during the initial epoch. Nonetheless, as training advances, the model's efficacy is enhanced. After the fourth epoch, the training accuracy attains 100%, while the validation accuracy stands at 97%. Regrettably, from the fourth epoch onward, the training accuracy maintains at 100%, whereas the validation accuracy displays variability. The significant gap between training and validation accuracy is apparent. Furthermore, the loss graph displays a similar issue, which is marked by consistently high loss values in both the training and validation datasets, making it unsuitable for our potato leaf dataset. Although the model attains elevated training accuracy, its diminished validation accuracy indicates a potential overfitting issue.

InceptionV3 initiates with a training accuracy of 62% and a validation accuracy of 72% in the initial epoch, as illustrated in Fig. 14. Analogous to MobileNetV2, the training accuracy progressively ascends and attains 100% by the seventh epoch, while the validation accuracy stands at 93%. From the seventh epoch onward, the training accuracy continuously remains at 100%; however, the validation accuracy exhibits fluctuations and fails to attain 100%. The disparity between training and validation accuracy is readily

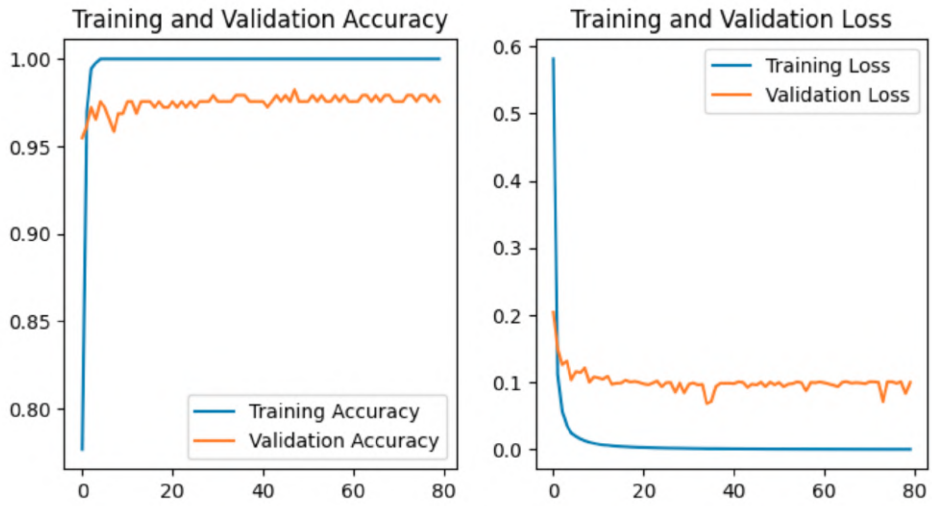

Fig. 11. The Accuracy and Loss Plots Using VGG16

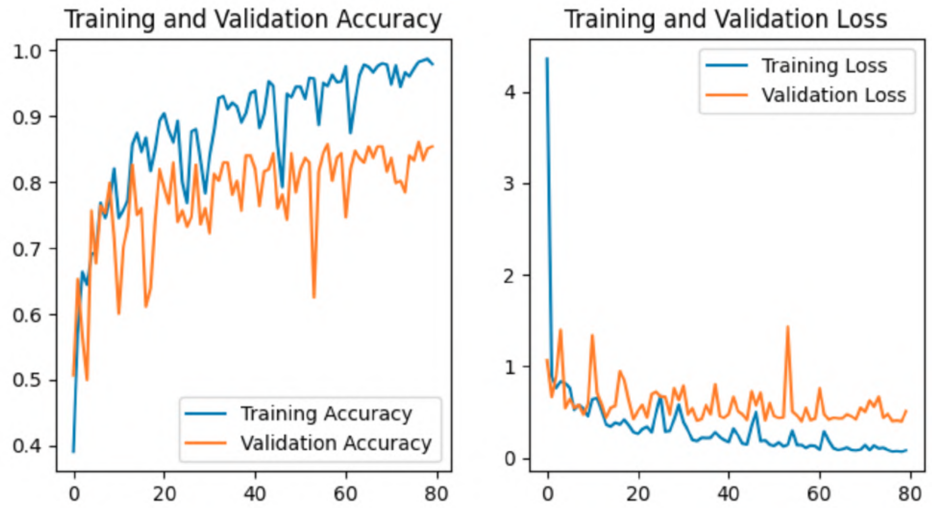

Fig. 12. The Accuracy and Loss Plots Using ResNet50

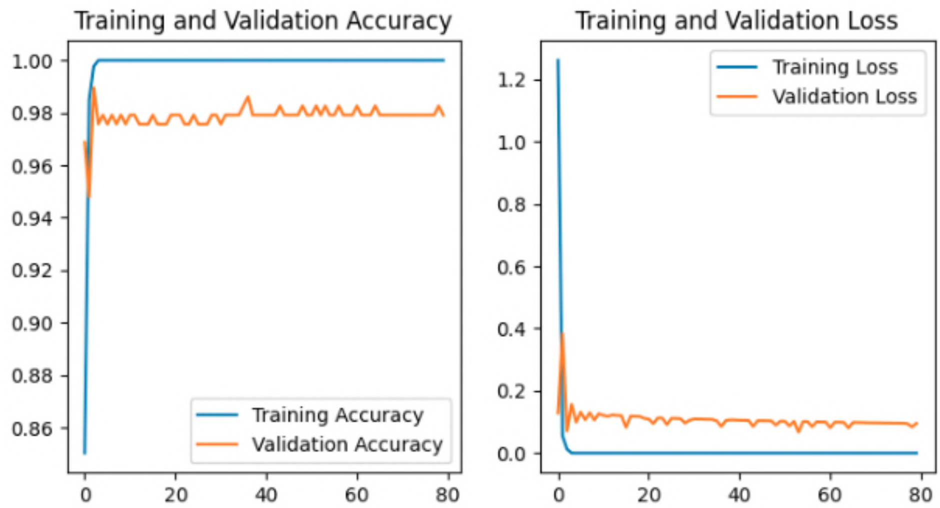

Fig. 13. The Accuracy and Loss Plots Using MobileNetV2

apparent. The loss graph consistently displays elevated loss values across all epochs for both training and validation. Despite attaining the best training accuracy, the model's poor validation accuracy indicates an overfitting issue, rendering it unsuitable for our dataset.

In Fig. 15, Xception exhibits an initial training accuracy of

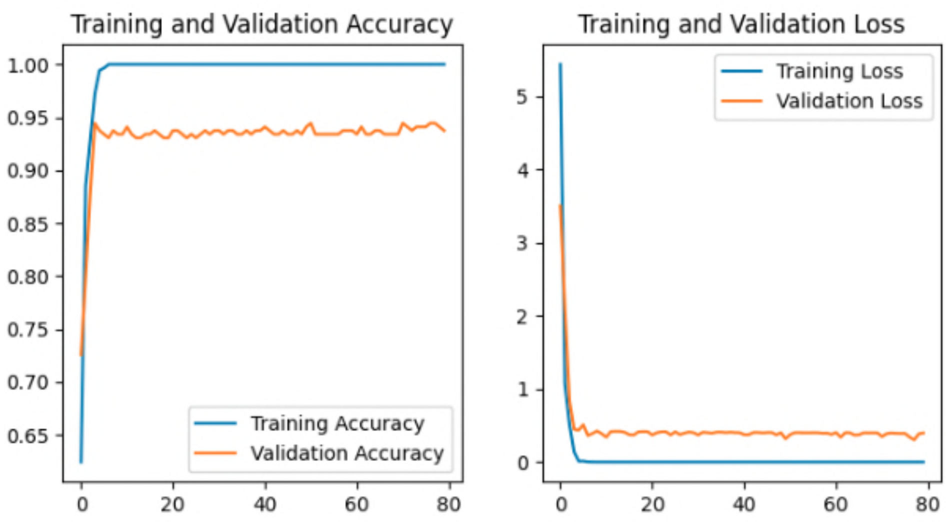

Fig. 14. The Accuracy and Loss Plots Using InceptionV3

83% and a validation accuracy of 95% during the first period. As training advances, the model's efficacy is enhanced. After the sixth epoch, the training accuracy attains a flawless 100%, but the validation accuracy is recorded at an admirable 98%. From the sixth epoch onward, the training accuracy constantly maintains 100%; however, the validation accuracy displays fluctuations. The significant disparity between training and validation accuracy is clearly evident. The loss graph also indicates the issue, with persistently elevated loss values for both training and validation datasets. Although the model attains elevated training accuracy, its diminished validation accuracy and substantial loss during both training and validation indicate an overfitting issue, rendering it less appropriate for our dataset.

Despite the impressive accuracy of all pre-trained models, issues of overfitting and underfitting render them unsuitable for our potato leaf disease dataset. To address this challenge, we developed a specialized deep learning model incorporating several modifications, such as layer tweaks, dropout layers, and data augmentation techniques. The proposed CNN model exhibits a training accuracy of 40% and a validation accuracy of 44% during the initial epoch, as illustrated in Fig. 16. Subsequently, the accuracy of both training and validation progressively improves concurrently while the loss in both training and validation diminishes. The training accuracy attains an impressive 99% by the last epoch, whereas the validation accuracy stands at 98%. The model consistently demonstrates suitable training and validation performance over all epochs, indicating its compatibility with our dataset. This model provides an appropriately tailored solution for our dataset by effectively addressing overfitting and underfitting issues.

The training and validation accuracy graphs indicate that VGG16, MobileNetV2, InceptionV3, and Xception attained 100% training accuracy; nevertheless, their validation accuracy is comparatively suboptimal. This signifies a case of overfitting when the models excel on the training data yet falter with the validation data. Conversely, ResNet50 has inadequate performance in both training and validation accuracy, indicating an underfitting issue. It does not align effectively with our particular potato leaf disease dataset, despite demonstrating outstanding performance on the ImageNet dataset. This illustrates that a model may perform exceptionally well on one dataset while underperforming on another.

We created a customized deep-learning CNN model to address these difficulties, optimizing it through the adjustment of various parameters. Our proposed CNN model achieved excellent performance for our specific potato leaf disease dataset. The training and validation accuracy is significant, indicating a robust model fit.

The pre-trained models employ transfer learning for classification. These pre-trained models, originally trained with a comprehensive array of parameters on the ImageNet dataset, do not undergo retraining of all parameters. It is impractical to train such a vast number of parameters within a limited timeframe. Retraining all parameters in our dataset would lead to increased time complexity. Transfer learning thus appears as the optimal method for image classification, as it obviates the necessity of retraining all parameters.

The paper proposed a deep-learning CNN model with fewer parameters for classifying potato leaf diseases, utilizing data augmentation techniques. We assessed the efficacy of various models in classifying potato leaf disease. MobileNetV2 achieved an accuracy of 98.67% with fewer parameters than the others. Xception attained 96% accuracy; VGG16 reached 97% accuracy; ResNet50 exhibited 80% accuracy, indicating potential underfitting issues; InceptionV3 secured 94.67% accuracy; and the proposed CNN Model achieved the highest accuracy at 99.67%.

The pre-trained models incorporate both trainable and non-trainable parameters as a result of transfer learning implementation. Transfer learning does not optimize all parameters. A custom model trains all parameters, unlike transfer learning. During model training, Python scripts can retrieve these parameters from the model summary. The accuracy in question refers to the test accuracy derived from test data after model training.

Table V presents a summary of the test accuracy, layer, and parameter of all models used in potato leaf disease identification.

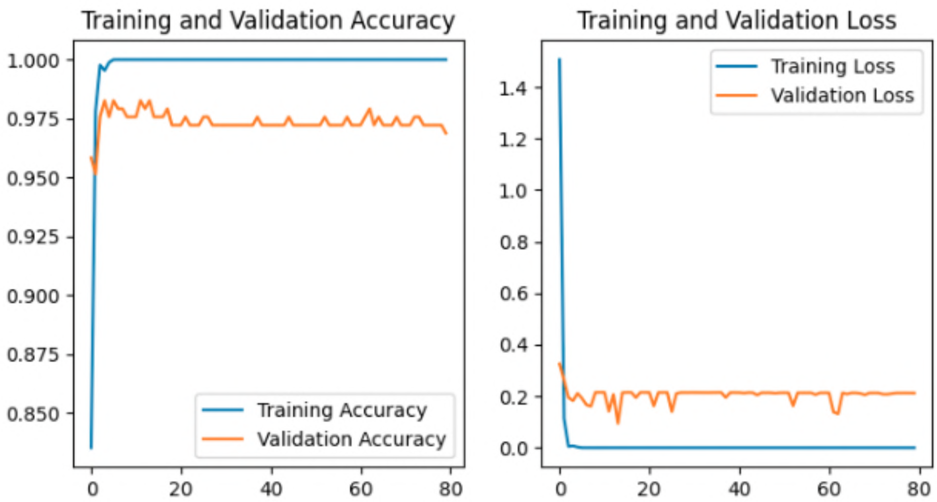

Fig. 15. The Accuracy and Loss Plots Using Xception

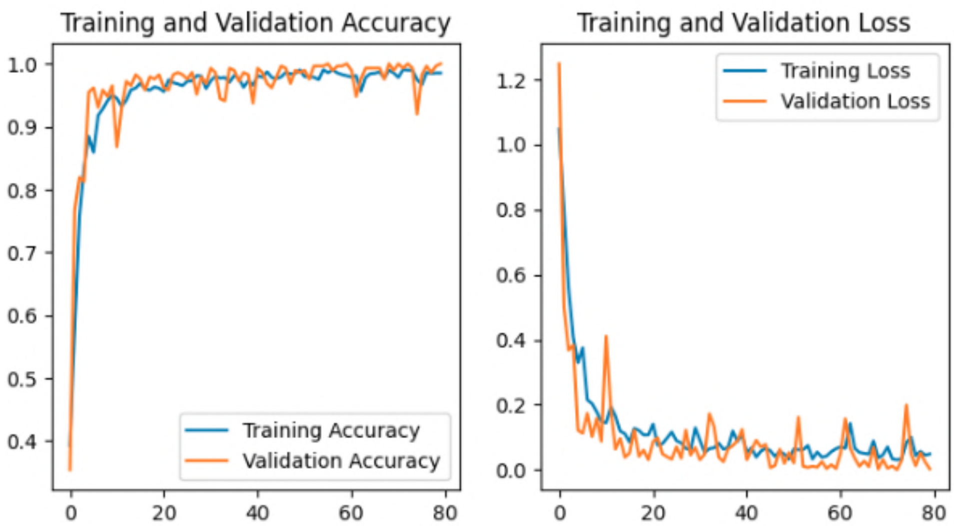

Fig. 16. The Accuracy and Loss Plots Using Proposed CNN Model

TABLE V
COMPARISON AMONG DIFFERENT MODEL RESPECT OF ACCURACY AND PARAMETER

Model	Test Accuracy	Layer	Parameter		
			Total Parameter	Trainable Parameter	N on-trainable Parameter
VGG16	97.00%	16	14,789,955	75,267	14,714,688
ResNet50	85.00%	50	23,888,771	301,059	23,587,712
MobileNetV2	98.67%	53	2,503,747	245,763	2,257,984
InceptionV3	94.67%,	48	21,956,387	153,603	21,802,784
Xception	98.00%	71	21,475,883	614,403	20,861,480
Proposed CNN Model	99.33%	10	233,187	233,187	0

The VGG16 model achieved a test accuracy of 97% on the potato leaf disease dataset, utilizing 75K trainable parameters across 16 layers. ResNet50, on the same dataset, attained 85% test accuracy with 301K trainable parameters and 50 layers. MobileNetV2 reached a test accuracy of 98.67%, employing 245K trainable parameters and 53 layers. InceptionV3 achieved a test accuracy of 94.67% using 153K trainable parameters and 48 layers, while Xception attained 98% test accuracy with 614K trainable parameters and 71 layers. In contrast, our proposed CNN model outperformed all pre-trained models, achieving an impressive test accuracy of 99.67% with only 233K trainable parameters and 10 layers. While VGG16 has fewer trainable parameters than our model, its test accuracy and test loss are inferior. Similarly, MobileNetV2, despite having fewer parameters, falls short in terms of training accuracy and test loss. Additionally, both VGG16 and MobileNetV2 exhibit higher layer counts and greater model complexity compared to our CNN model. No-

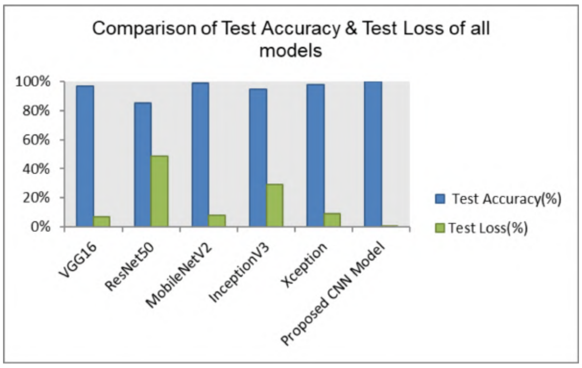

Fig. 17. Comparison of Test Accuracy and Test Loss of the Model

tably, our proposed CNN model combines low computational cost, minimal layer count, and reduced complexity while maintaining an optimal number of trainable parameters and delivering the best accuracy among all the models tested.

Table VI provides an overview of the training accuracy and loss, as well as the test accuracy and loss for each model.

The pre-trained CNN models shown in Table VI are used on the same dataset to find diseases in potato leaves using transfer learning. VGG16, ResNet50, MobileNetV2, InceptionV3, and Xception are pre-trained models that employ transfer learning. Additionally, we employ various augmentation techniques to create a tailored CNN model for the classification of potato leaf diseases on the identical dataset. Except for ResNet50, all other pre-trained models exhibit small training losses and achieve 100% training accuracy. However, their performance varies concerning the test dataset. MobileNetV2 attains 98.67% test accuracy with a test loss of 7.44%, InceptionV3 records 94.67% test accuracy with a test loss of 29.26%, Xception accomplishes 98% test accuracy with a test loss of 9.12%, and VGG16 reaches 97% test accuracy with a test loss of 6.85%. In our unobserved potato leaf disease dataset, these pre-trained models exhibit inferior performance due to their significantly higher test loss. Their reduced test accuracy and increased test loss indicate overfitting issues despite their exceptional training accuracy.

However, ResNet50's operation differs significantly. It achieves a training accuracy of 97% and a training loss of 9.39%. The test accuracy of ResNet50 is 85%, accompanied by a high test loss of 48.82%, demonstrating very poor performance on our dataset. The combination of low training and test accuracy, along with high test loss, suggests an underfitting problem. None of the pre-trained models seem to be well-suited for our potato leaf disease dataset. Our dataset on potato leaf disease seems incompatible with any of the pre-trained models.

Our proposed CNN framework, however, significantly mitigates these issues. It attains an exceptionally low test loss of 0.8%, resulting in a notable test accuracy of 99.67% and a training accuracy of 98.71%. Our suggested approach excels at diagnosing potato leaf disease, especially when confronted with unfamiliar data. The proposed CNN model effectively addresses the issue of identifying potato leaf disease. Fig. 17 presents a comparison of test accuracy and loss for all models.

In the majority of instances, the proposed CNN model outperforms the pre-trained CNN models regarding accuracy. Moreover, in the pre-trained models, the test loss increases; however, in the proposed CNN model, it diminishes by 0.8%, thereby enhancing our model's performance. The training accuracy surpasses all pre-trained models, excluding ResNet50, although the testing accuracy is inferior to that of our suggested CNN model. Our proposed CNN model attains the best test accuracy and the lowest test loss.

The confusion matrix analyzed the performance of the potato leaf disease classification model. It displayed the performance of our proposed CNN model in identifying three varieties of potato leaves. In this instance, 0 denotes potato early blight, 1 signifies potato late blight, and 2 represents a healthy potato leaf. It matched the actual target values with the predicted ones for each class of the potato leaf dataset. It helps us figure out how well our CNN model can tell the difference between early blight, late blight, and healthy leaves in our potato leaf disease recognition application. This matrix displays the performance of each class on the potato leaf dataset. It delivers outstanding outcomes for the potato early blight and potato healthy categories in our test dataset. It also demonstrates excellent performance in the potato late blight category within the test image dataset. We can comprehensively evaluate the exceptional performance of our proposed CNN model in identifying potato leaf diseases through the confusion matrix. Figure 18 illustrates the confusion matrix of the proposed CNN model.

The classification report is a table summarizing multiple evaluation metrics. We use it to evaluate the performance of our model. It shows a detailed summary of the performance of our classification model. It typically shows metrics like precision, recall, F1-score, and support for each class. We present the classification report for our proposed CNN model, which evaluates three classes of potato leaves: potato early blight, late blight, and healthy. Our CNN model demonstrates excellent performance across all classes, achieving a remarkable test accuracy of 99.67%. We achieved this performance by using data augmentation techniques, which required fewer parameters and layers than other pre-trained models. The precision, recall, F1-score, and support for our proposed CNN model on the potato leaf disease dataset are presented in Table VII. With precision ratings of 100%, 100%, and 98% for the early blight, late blight, and healthy categories, respectively, the testing findings show how successful the suggested method is. Likewise, the early blight, late blight, and healthy categories had 100%, 99%, and 100% recall scores, respectively. It also scored 100%, 98%, and 99% F1-scores, respectively, for the early blight, late blight, and healthy categories. Additionally, the method demonstrated its resilience and dependability by achieving a weighted-average accuracy of 100% and a macro-average accuracy of 99% on the test dataset. The graph of the precision, recall, and F1-score of our model on the potato dataset is represented in Fig. 19.

Our proposed CNN model outperformed all other pre-trained models, with the greatest test accuracy of 99.33% and the lowest test loss. It successfully mitigated the overfitting problem by integrating a dropout layer of 0.5 into the fully linked layers and other data augmentation methods. MobileNetV2 produced the optimal outcome, attaining an accuracy of 98.67% while utilizing fewer parameters than

TABLE VI
TRAINING ACCURACY AND LOSS, TEST ACCURACY AND LOSS COMPARISON

Model	Technique		Training Accuracy(%)	Training Loss(%)	Test Accuracy(%)	Test Loss(%)	Model Fit
	Transfer Learning	Data Augmentation					
VGG16	Yes	No	100.00	0.03087	97.00	6.85	Overfit
ResNet50	Yes	No	97.00	9.39516	85.00	48.82	Under fit
MobileNetV2	Yes	No	100.00	0.00006	98.67	7.44	Overfit
InceptionV3	Yes	No	100.00	0.00155	94.67	29.26	Overfit
Xception	Yes	No	100.00	0.00018	98.00	9.12	Overfit
Proposed CNN Model	No	Yes	98.71	3.35295	99.67	0.8	Good fit

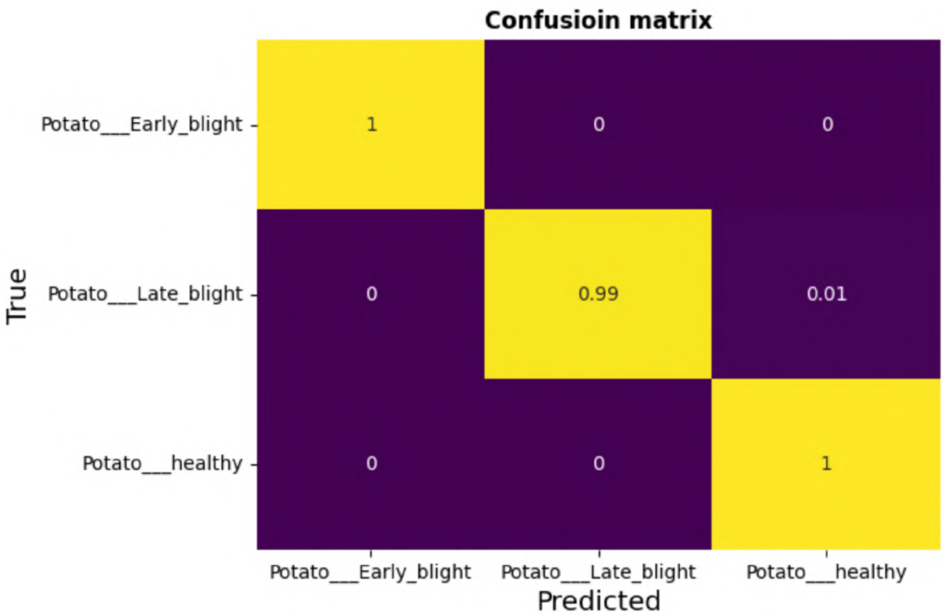

Fig. 18. The confusion matrix of our proposed CNN model

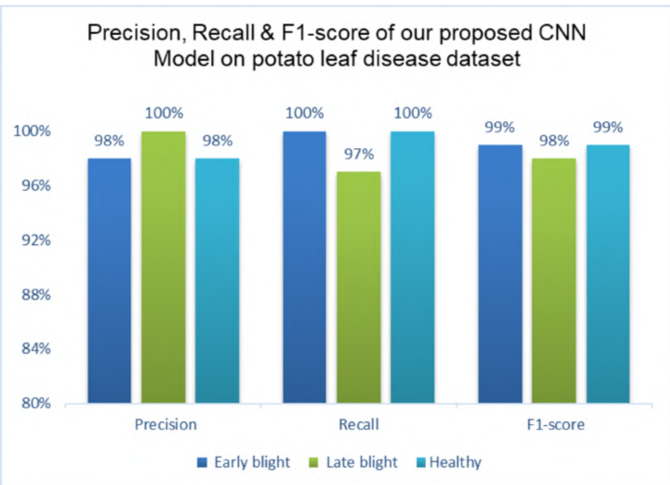

Fig. 19. Comparison chart of the precision, recall and F1-score for three potato leaf classes using our proposed CNN model

TABLE VII
CLASSIFICATION REPORT OF OUR PROPOSED CNN MODEL

	Precision	Recall	F1-score	Support
Early blight	1.00	1.00	1.00	205
Late blight	1.00	0.99	0.98	199
Healthy	0.98	1.00	0.99	126
Accuracy			1.00	530
Macro avg	0.99	0.99	1.00	530
Weighted avg	1.00	1.00	1.00	530

other pre-trained CNN models. Conversely, ResNet50 had the least effective performance among the models in our dataset. The pre-trained models faced difficulties associated with overfitting and underfitting. Overfitting occurs when the training accuracy increases, but the test accuracy significantly decreases. Underfitting occurs when both training and test accuracy are inadequate. Except for ResNet50, all pre-trained models attained elevated training accuracy; nevertheless, their test accuracy was comparatively poor, accompanied by considerable test loss. Nonetheless, our suggested CNN model successfully addressed these challenges, attaining the highest test accuracy and the lowest test loss. It exhibited exceptional performance with an accuracy of 99.67% on the test dataset, utilizing data augmentation approaches while employing fewer parameters and layers than other pre-trained models. Our proposed CNN model attains the highest test accuracy in the identification of potato leaf diseases. It is illustrated in Fig. 20.

Fig. 20. Recognition of potato diseases from the leaf using Proposed CNN Model

Fig. 21. Sample of potato leaf dataset collected from regional fields

To help untrained farmers accurately identify potato illnesses and implement appropriate treatment measures, we created a mobile application. Our application has improved potato yields by facilitating accurate disease classification. To verify its reliability, we evaluated the application using actual photos obtained from nearby potato farms. Several images of potato leaves are collected from various potato fields to test our mobile application. Sample images from the real-time dataset are shown in Fig. 21.

It presents the confidence level for each category after classifying potato leaf images. It precisely recognizes previously unobserved potato leaf images, indicating that our suggested CNN model successfully reduces overfitting. The subsequent images depict the interface of the Android application that identifies three categories of potato leaf conditions: early blight, late blight, and healthy. The application consistently differentiates between sick and healthy leaves. We evaluated the Android application with both web-derived photos and real-time images captured in the field. Clean, closely taken, and noise-free photos significantly increase classification accuracy and boost confidence levels. However, distorted or ambiguous images significantly diminish confidence levels.

Figure 22 demonstrates the potato early blight recognition using the mobile application. In this case, all three images are correctly classified as early blight, but with varying confidence levels. The first image is very clear, resulting in 100% confidence. However, the other two images are less clear, leading to lower confidence scores compared to the first one.

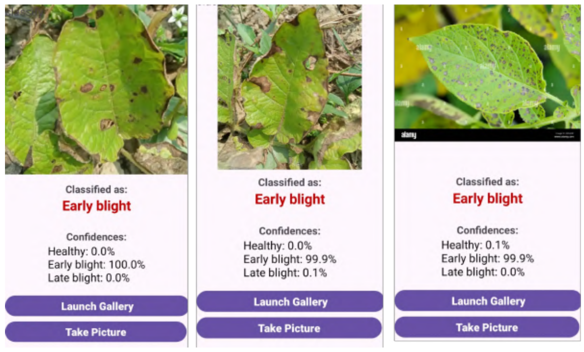

Fig. 22. Potato early blight recognition using mobile application

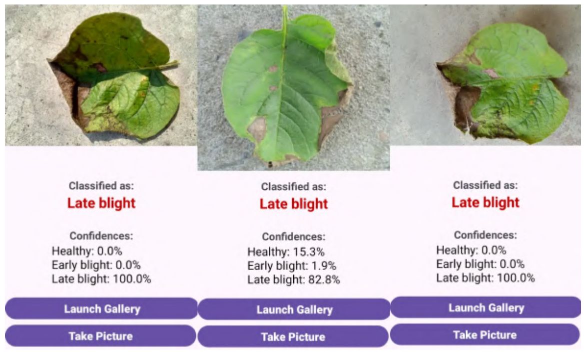

Fig. 23. Potato late blight recognition using mobile application

Figure 23 demonstrates the recognition of potato late blight using the mobile application. Similar to early blight, late blight is correctly identified across all images. However, the confidence levels differ due to variations in image quality. The first and third images are clear, achieving 100% confidence, while the remaining image is noisy and reversed, resulting in reduced confidence levels during identification.

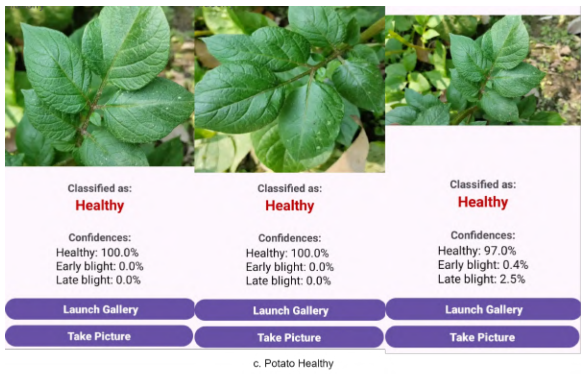

Fig. 24. Potato healthy leaf recognition using mobile application

Fig. 24 demonstrates the potato leaf recognition using the mobile application. The first and second images are sharp and clear, resulting in 100% confidence. In contrast, the third image was taken from a distance and contains additional objects in the frame, making it less clear and causing a lower confidence score compared to the first two images.

We tested the app using unseen images with class labels from external websites to identify any potential errors. For new and unseen data, the model's confidence may decrease, while for training data, the confidence remains consistently high. An ideal model should accurately predict class labels for unseen data while maintaining high confidence. In our case, the model successfully predicts the correct class labels for most of the unseen data. However, in some cases, although the confidence is lower, the predictions remain accurate.

The main contribution is focused on our study.

- We introduce a convolutional neural network (CNN) model that uses fewer parameters and methods to classify potato leaf diseases.
- Normalizing pixel values to a range of 0 to 1 and enriching the data with techniques like random flipping, shifting, zooming, brightness enhancement, rotation, and shearing improve the model's robustness.
- We address the issues of underfitting and overfitting to increase the efficacy of our proposed CNN model.
- We employ batch normalization, data augmentation, and dropout layers to enhance convergence and mitigate overfitting in our proposed CNN model.
- The pre-trained models utilize transfer learning to detect potato leaf disease.
- The fundamental criteria for evaluating all models are tested loss and accuracy.
- We compare the accuracy of pre-trained models with the suggested CNN model.
- The proposed CNN model surpasses previous pre-trained models by mitigating overfitting and underfitting, achieving the maximum test accuracy of 99.67% and the lowest test loss of 0.8%.
- Our proposed CNN model significantly reduces test loss in comparison to pre-trained models used in potato leaf disease recognition.
- The proposed CNN model accurately identified potato leaf disease and demonstrated strong performance on our test dataset.
- We create a mobile application utilizing our proposed CNN model and evaluate it on a potato leaf data set sourced from a local potato field.

V. Conclusion and Future work

This study employs leaf samples to distinguish between healthy and infected potato leaves, focusing on early and late blight. We attain this distinction by using a deep learning methodology, utilizing five transfer-learning models in conjunction with a bespoke CNN model augmented by data enhancement approaches.

However, adjustments to several parameters and layers enabled our suggested CNN model to perform the task more efficiently. This change markedly diminished the overfitting problem that had consistently plagued the pre-trained models.

We suggested a CNN model that excels in detecting potato leaf diseases, attaining maximum accuracy with low loss. Among the pre-trained models, VGG16 and MobileNetV2 yielded robust results while necessitating fewer parameters than their counterparts. Xception and InceptionV3 exhibited notable performance; nonetheless, ResNet50 was the least effective on our dataset. Nonetheless, all the pre-trained models demonstrated a certain level of overfitting. In contrast, our CNN model effectively addressed this issue, surpassing the pre-trained models in both test accuracy and test loss by employing data augmentation approaches.

For this study, we employed a publicly accessible dataset. We intend to create a custom dataset by gathering photos of healthy and damaged potato leaves from multiple potato fields. We intend to enhance accuracy and minimize loss with this augmented and customized dataset. Furthermore, we anticipate the integration of software and hardware solutions for enhanced precision and real-time outcomes. We intend to work with additional pre-trained models to create a more robust one for improved accuracy.

Acknowledgment

This work was supported by the Computer Science and Engineering Discipline, Jashore University of Science & Technology.

References

[1] K. Sanjeev, N. K. Gupta, W. Jeberson, and S. Paswan, "Early prediction of potato leaf diseases using ann classifier," *Oriental Journal of Computer Science and Technology*, vol. 13, no. 2, 3, pp. 129–134, 2021.

[2] T. T. Moharekar, U. R. Pol, R. Ombase, T. J. Moharekar *et al.*, "Detection and classification of plant leaf diseases using convolutional neural networks and streamlit," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 7, pp. 4305–4309, 2022.

[3] R. Mahum, H. Munir, Z.-U.-N. Mughal, M. Awais, F. Sher Khan, M. Saqlain, S. Mahamad, and I. Tlili, "A novel framework for potato leaf disease detection using an efficient deep learning model," *Human and Ecological Risk Assessment: An International Journal*, vol. 29, no. 2, pp. 303–326, 2023.

[4] M. H. Al-Adhaileh, A. Verma, T. H. Aldhyani, and D. Koundal, "Potato blight detection using fine-tuned cnn architecture," *Mathematics*, vol. 11, no. 6, p. 1516, 2023.

[5] T. Nazir, M. M. Iqbal, S. Jabbar, A. Hussain, and M. Albathan, "Efficientpnet—an optimized and efficient deep learning approach for classifying disease of potato plant leaves," *Agriculture*, vol. 13, no. 4, p. 841, 2023.

[6] D. Kothari, H. Mishra, M. Gharat, V. Pandey, M. Gharat, and R. Thakur, "Potato leaf disease detection using deep learning," *Int. J. Eng. Res. Technol*, vol. 11, no. 11, 2022.

[7] S. Bangari, P. Rachana, N. Gupta, P. S. Sudi, and K. K. Baniya, "A survey on disease detection of a potato leaf using cnn," in *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. IEEE, 2022, pp. 144–149.

[8] N. Ananthi, K. Kumaran, V. Madhushalini, S. G. Moorthi, and P. Harish, "Detection and identification of potato plant leaf diseases using convolution neural networks," *Eur J Mol Clin Med*, vol. 7, no. 4, 2020.

[9] A. Singh and H. Kaur, "Potato plant leaves disease detection and classification using machine learning methodologies," in *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1. IOP Publishing, 2021, p. 012121.

[10] N. Tilahun and B. Gizachew, "Artificial intelligence assisted early blight and late blight potato disease detection using convolutional neural networks," *Ethiopian Journal of Crop Science*, vol. 8, no. 2, 2020.

[11] B. Gardie, S. Asemie, K. Azezew, and Z. Solomon, "Potato plant leaf diseases identification using transfer learning," *Indian Journal of Science and Technology*, vol. 15, no. 4, pp. 158–165, 2022.

[12] M. A. Iqbal and K. H. Talukder, "Detection of potato disease using image segmentation and machine learning," in *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*. IEEE, 2020, pp. 43–47.

[13] A. J. Rozaqi and A. Sunyoto, "Identification of disease in potato leaves using convolutional neural network (cnn) algorithm," in *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*. IEEE, 2020, pp. 72–76.

[14] M. K. R. Asif, M. A. Rahman, and M. H. Hena, "Cnn based disease detection approach on potato leaves," in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, 2020, pp. 428–432.

[15] R. A. Sholihati, I. A. Sulistijono, A. Risnumawan, and E. Kusumawati, “Potato leaf disease classification using deep learning approach,” in *2020 International Electronics Symposium (IES)*. IEEE, 2020, pp. 392–397.

[16] C. Hou, J. Zhuang, Y. Tang, Y. He, A. Miao, H. Huang, and S. Luo, “Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation,” *Journal of Agriculture and Food Research*, vol. 5, p. 100154, 2021.

[17] M. Al-Amin, D. Z. Karim, and T. A. Bushra, “Prediction of rice disease from leaves using deep convolution neural network towards a digital agricultural system,” in *2019 22nd International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2019, pp. 1–5.

[18] J. Feng, B. Hou, C. Yu, H. Yang, C. Wang, X. Shi, and Y. Hu, “Research and validation of potato late blight detection method based on deep learning,” *Agronomy*, vol. 13, no. 6, p. 1659, 2023.

[19] F. Kang, J. Li, C. Wang, and F. Wang, “A lightweight neural network-based method for identifying early-blight and late-blight leaves of potato,” *Applied Sciences*, vol. 13, no. 3, p. 1487, 2023.

[20] H. Ghosh, I. S. Rahat, K. Shaik, S. Khasim, and M. Yesubabu, “Potato leaf disease recognition and prediction using convolutional neural networks,” *EAI Endorsed Transactions on Scalable Information Systems*, vol. 10, no. 6, 2023.

[21] P. Enkvetchakul and O. Surinta, “Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition,” *Applied Science and Engineering Progress*, vol. 15, no. 3, pp. 3810–3810, 2022.

[22] P. Patil, N. Yaligar, and S. Meena, “Comparision of performance of classifiers-svm, rf, and ann in potato blight disease detection using leaf images,” in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 2017, pp. 1–5.

[23] D. Tiwari, M. Ashish, N. Gangwar, A. Sharma, S. Patel, and S. Bhardwaj, “Potato leaf diseases detection using deep learning,” in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 2020, pp. 461–466.

[24] M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, “Detection of potato diseases using image segmentation and multiclass support vector machine,” in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2017, pp. 1–4.

[25] B. Liu, Y. Zhang, D. He, and Y. Li, “Identification of apple leaf diseases based on deep convolutional neural networks,” *Symmetry*, vol. 10, no. 1, p. 11, 2017.

[26] J. Arora, U. Agrawal *et al.*, “Classification of maize leaf diseases from healthy leaves using deep forest,” *Journal of Artificial Intelligence and Systems*, vol. 2, no. 1, pp. 14–26, 2020.

[27] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, “Identification of maize leaf diseases using improved deep convolutional neural networks,” *Ieee Access*, vol. 6, pp. 30 370–30 377, 2018.

[28] M. R. Howlader, U. Habiba, R. H. Faisal, and M. M. Rahman, “Automatic recognition of guava leaf diseases using deep convolution neural network,” in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2019, pp. 1–5.

[29] K. Rangarajan Aravind and P. Raja, “Automated disease classification in (selected) agricultural crops using transfer learning,” *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, vol. 61, no. 2, pp. 260–272, 2020.

[30] M. N. Hasan, M. Mustavi, M. A. Jubaer, M. T. Shahriar, and T. Ahmed, “Plant leaf disease detection using image processing: A comprehensive review,” *Malaysian Journal of Science and Advanced Technology*, pp. 174–182, 2022.

[31] V. K. Shrivastava, M. K. Pradhan, S. Minz, and M. P. Thakur, “Rice plant disease classification using transfer learning of deep convolution neural network,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 631–635, 2019.

[32] J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, “Deep learning-based leaf disease detection in crops using images for agricultural applications,” *Agronomy*, vol. 12, no. 10, p. 2395, 2022.

[33] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, “Classification of skin disease using deep learning neural networks with mobilenet v2 and lstm,” *Sensors*, vol. 21, no. 8, p. 2852, 2021.

[34] B. M. Abuhayi and A. A. Mossa, “Coffee disease classification using convolutional neural network based on feature concatenation,” *Informatics in Medicine Unlocked*, vol. 39, p. 101245, 2023.

[35] P. Vasavi, A. Punitha, and T. V. N. Rao, “Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: A review,” *International Journal of Electrical and Computer Engineering*, vol. 12, no. 2, p. 2079, 2022.

[36] S. S. Yadav and S. M. Jadhav, “Deep convolutional neural network based medical image classification for disease diagnosis,” *Journal of Big data*, vol. 6, no. 1, pp. 1–18, 2019.

[37] S. Xiaoxiao, M. Shaomin, X. Yongyu, C. Zhihao, and S. Tingting, “Image recognition of tea leaf diseases based on convolutional neural network,” in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*. IEEE, 2018, pp. 304–309.

[38] R. Kundu, U. Chauhan, and S. Chauhan, “Plant leaf disease detection using image processing,” in *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)*, vol. 2. IEEE, 2022, pp. 393–396.

[39] R. Sharma, A. Singh, N. Jhanjhi, M. Masud, E. S. Jaha, S. Verma *et al.*, “Plant disease diagnosis and image classification using deep learning.” *Computers, Materials & Continua*, vol. 71, no. 2, 2022.

[40] M. N. Hasan, M. Mustavi, M. A. Jubaer, M. T. Shahriar, and T. Ahmed, “Plant leaf disease detection using image processing: A comprehensive review,” *Malaysian Journal of Science and Advanced Technology*, pp. 174–182, 2022.

[41] M. Abed, M. A. Imteaz, and A. N. Ahmed, “Utilising machine learning for pan evaporation prediction - a case study in western australia,” in *Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2023*, 5 - 7 July, 2023, Hong Kong, pp. 14–18.

[42] K. A. Ardisa, W. A. E. Prabowo, and S. Rustad, “Implementation of convolutional neural network method for detecting vegetables as recommendation for vegetarian food recipes,” in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2022*, 6 - 8 July, 2022, London, U.K., pp. 38–41.

[43] S. Takezaki and K. Kishida, “Construction of cnns for abnormal heart sound detection using data augmentation,” in *Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021*, 20-22 October, 2021, Hong Kong, pp. 18–23.

[44] V. B. Kumar, K. Vijayalakshmi, and M. Padmavathamma, “A hybrid data mining approach for diabetes prediction and classification,” in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2019*, 22-24 October, 2019, San Francisco, USA, pp. 298–303.

[45] “Supplementary info (si) for the experimental scripts:,” https://github.com/YasirArafat91/Deep-learning/blob/8293316b10794c072b98e31d24890c0e6d7856cf/PLD_CNN.ipynb, accessed: 02 April 2025.