I. J. Engineering and Manufacturing, **2023, 6, 9-22**
Published Online on December 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijem.2023.06.02

Deep Learning-Based Potato Leaf Disease Detection Using CNN in the Agricultural System

Abdullah Walid
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh
Email: abdullahwalid.rcs.rcc@gmail.com

Md. Mehedi Hasan
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh
Email: mhsojib2014@gmail.com

Tonmoy Roy
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh
Email: tonmoyroydev@gmail.com

Md. Selim Hossain*
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh
Email: selim.ece@tch.hstu.ac.bd
*Corresponding Author

Nasrin Sultana
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh
Email: nasrinsultana@hstu.ac.bd

Received: 25 July, 2023; Revised: 20 August, 2023; Accepted: 05 September, 2023; Published: 08 December, 2023

Abstract: Potatoes play a vital role as a staple crop worldwide, making a significant contribution to global food security. However, the susceptibility of potato plants to various leaf diseases poses a threat to crop yield and quality. Detecting these diseases accurately and at an early stage is crucial for the effective management and protection of crops. Recent advancements in Convolutional Neural Networks (CNNs) have demonstrated potential in image categorization applications. Therefore, the goal of this work is to investigate the potential of CNNs in detecting potato leaf diseases. As neural networks have become part of agriculture, numerous researchers have worked on improving the early detection of potato blight using different machine and deep learning methods. However, there are persistent problems related to accuracy and the time it takes for these methods to work. In response to these challenges, we tailored a convolutional neural network (CNN) to enhance accuracy while reducing the trainable parameters, computational time and information loss. To conduct this research, we compiled a diverse dataset consisting of images of potato leaves. The dataset encompassed both healthy leaves and leaves infected with common diseases such as late blight and early blight. We took great care in curating and preprocessing the dataset to ensure its quality and consistency. Our focus was to develop a specialized CNN architecture tailored specifically for disease detection. To improve the performance of the network, we employed techniques like data augmentation and transfer learning during the training phase. The experimental outcomes demonstrate the efficacy of our proposed customized CNN model in accurately identifying and classifying potato leaf diseases. Our model's overall accuracy was an astounding 99.22%, surpassing the performance of existing methods by a significant margin. Furthermore, we evaluated precision, recall, and F1-score to evaluate the model's effectiveness on individual disease classes. To give an additional understanding of the model's behavior and its capacity to distinguish between various disease types, we utilized visualization techniques such as confusion matrices and sample output images. The results of this study have implications for managing potato diseases by offering an automated and reliable solution for early detection and diagnosis. Future research directions may include expanding the dataset, exploring different CNN architectures, and investigating the generalizability of the model across different potato varieties and growing conditions.

Index Terms: Potato leaf diseases, leaf disease detection, Convolutional Neural Networks (CNNs), image classification, crop protection.

1. Introduction

Potatoes hold immense agricultural importance globally, serving as a vital vegetable crop and a staple food in more than 130 developing countries. Unfortunately, potato farmers encounter substantial challenges, experiencing substantial losses of over 32% annually due to diseases and pests. In Bangladesh alone, these losses amount to Tk 2,500 crore every year. To mitigate this problem, a classification model utilizing convolutional neural networks (CNNs) has been proposed for the accurate identification of potato leaf diseases.

Potato farming is a prevalent occupation in over 125 countries, playing a vital role in their agricultural landscape. However, potato crops are highly vulnerable to infections, especially from diseases such as late blight and early blight. These diseases pose a significant threat to the crop, resulting in reduced production and compromising the overall quality of potatoes. [1]. Many potato types produced in Colombia are susceptible to late blight [2,3], and the disease is difficult to control without using significant amounts of pesticide. Late blight affects leaves, stems, and tubers, causing blistering and browning or blackening of the leaves. On the other hand, early blight manifests as characteristic leaf spots and blight [4]. Controlling these diseases requires specific environmental conditions such as humidity, temperature, and leaf wetness.

Efficient and accurate identification of diseases in potato plants is crucial to mitigate their impact. Manual monitoring by farmers is time-consuming and impractical, and relying on visual observation alone can lead to inaccurate identification due to the similarities in symptoms between different diseases [5]. Farmers often lack the necessary expertise, leading to ineffective preventive measures and potential crop damage [6].

Here, a Convolutional Neural Network (CNN) strategy is suggested to deal with these difficulties. The CNN model enables fast and accurate identification and classification of two common potato infections. By using this method, farmers can easily detect diseases in potato crops with minimal computational effort [7].

The research introduces a CNN-based approach with significant potential to assist farmers, agronomists, and researchers in managing diseases and optimizing crop yields within agriculture. The primary objective of this study is to address the challenge of precisely detecting and categorizing diseases affecting potato leaves, including late blight and early blight. Typically, the severity of late blight is assessed through a visual inspection process, which involves determining the proportion of a crop's foliage that is affected by the disease. However, evaluating disease severity through visual means is demanding in terms of labor and time, and it lacks consistency because it relies on subjective assessments by professionals. Furthermore, relying solely on diagnoses made by farmers or specialists may lead to inaccuracies, potentially resulting in inappropriate use of chemicals that can adversely affect crop quality and harm the environment over time.

Our study utilized several CNN architectures, including VGG16, VGG19, ResNet, MobileNet, and our customized CNN model. Among these architectures, MobileNet achieved the highest accuracy. However, our model stands out as the most efficient in terms of both training time and trainable parameters, while still maintaining accuracy levels comparable to MobileNet.

2. Related Research Work

Tiwari et al. [8] utilized the Kaggle dataset comprising potato plant leaf images to conduct classification tasks, distinguishing between healthy leaves, early blight, and late blight. They employed the pre-trained VGG19 model and had a 97.8% accuracy rate. In their research, they addressed the crucial issue of potato leaf disease detection which significantly impact potato quality and quantity. Their work is particularly relevant to our research objectives as it demonstrates the potential of pre-trained deep learning models, such as VGG19, in the context of potato leaf disease detection.

Mosleh et al [9] introduce a customized convolutional neural network (CNN) for potato blight recognition. They proposed a model that outperforms other machine and deep learning algorithms, achieving an impressive 99% accuracy with only 839,203 trainable parameters and a training time of 183 seconds. They custom-designed a convolutional neural network (CNN) tailored to optimize accuracy while significantly reducing trainable parameters and computation time. Their work not only stands out for its high accuracy but also for its innovative approach to mitigating computational complexities, aligning with our research objective of developing efficient solutions for potato blight detection.

Divyansh et al [10] present a model that utilizes pre-trained models and transfer learning, specifically employing VGG19 and logistic regression. They proposed a model that achieves a remarkable 97.8% categorization accuracy on the testing dataset.

Deep et al [11] focus on the use of deep learning methods, specifically the GoogleNet, Resnet50, and VGG16 models, for the classification of potato diseases based on leaf conditions. By achieving a high accuracy of 97% within the initial training epochs, the study demonstrates the feasibility and potential of deep neural networks in accurately identifying potato diseases. Their study's noteworthy achievement includes attaining a commendable 97% accuracy within the initial 40 epochs of training, showcasing the feasibility and potential of deep neural networks in addressing the challenges associated with potato disease detection. This research aligns well with our objective of exploring advanced techniques for precise disease identification in agriculture.

Pranay etl.[12] worked with an open-source dataset of a corn leaf. CNN model trained by using 4000 images and tested using 2000 images. This model optioned 98% accuracy.

Mangal et al. [13] addressed the pressing issue of early plant disease detection in India's agriculture sector. Conventional methods proved inefficient, causing significant crop losses. To tackle this challenge, they applied modern image processing techniques like Laplacian filter, Unsharp masking, and Canny edge detection. Their approach centered on a convolutional neural network (CNN), known for image classification tasks. The results were impressive, achieving a 97.82% accuracy in distinguishing healthy and diseased plant leaves. This work showcases the potential of technology to transform agriculture by enhancing early disease detection and reducing crop losses.

Varsha et al [14] use a personalized convolutional neural network (CNN) that was focused on identifying healthy, early blight, and late blight illnesses in potato leaves. The CNN model had a 96.0% accuracy rate, indicating its effectiveness in disease classification. This research showcases the potential of automated techniques, such as deep learning, to enhance disease monitoring in large farms, thereby contributing to improved crop health and productivity in India.

In their research, Rabbia and colleagues [15] introduced a novel algorithm that effectively detected and classified four illnesses in potato leaves. A testing set was used to gauge the algorithm's performance, achieving an impressive accuracy of 97.2%. The proposed algorithm utilized a pre-trained Efficient DenseNet model with an additional transition layer in DenseNet-201. The introduction of a reweighted cross-entropy loss function addressed the issue of imbalanced training data, while dense connections with regularization contributed to reducing overfitting during training on smaller datasets of potato leaf samples. Notably, this research represents the first and novel technique for successfully detecting and classifying four diseases in potato leaves.

In a separate study, Chakraborty et al. [16] developed a methodology that demonstrated a remarkable accuracy of 97.89% for distinguishing between healthy potato leaves and those with late and early blight syndromes. The study displayed validation accuracy and losses, highlighting the fine-tuned VGG16 model's complex architecture. Furthermore, the proposed methodology was rigorously compared with existing techniques, reinforcing its efficacy in addressing the challenges of detecting and classifying late and early blight diseases in potato leaves. Additionally, Chen et al. [17] proposed a procedure that outperformed other existing methods, achieving an impressive 97.73% identification accuracy on average across various potato disease categories. The experimental findings confirmed the validity of the proposed method by demonstrating its superior competitive efficacy.

3. Research Methodology

3.1 Dataset Description

We got our dataset from Kaggle.com.[18]. Kaggle is an online platform renowned for fostering a vibrant community and offering valuable resources for individuals passionate about data science and machine learning. It hosts a wide range of datasets, competitions, and notebooks that allow users to explore, analyze, and develop solutions for various real-world problems.

3.1.1 Data Acquisition

We used the Plant Village dataset, which is a collaborative effort between Penn State University and EPFL. It consists of 256x256 pixel JPG color images of infected and healthy leaves from 14 different plant species. Our study focuses on the potato crop, and we chose a subset of 152 images of healthy leaves, 1000 photographs of leaves with late blight, and 1000 photos of leaves with early blight. This subset from the Plant Village dataset allows us to effectively address our research objectives.

Table 1. Details of Plant Village Dataset

Label	Category	Number	Training Sample	Test Sample
1	Early Blight	1000	790	210
2	Late Blight	1000	790	210
3	Healthy	152	124	28
Total		2152	1704	448

Early Blight: Alternaria solani, a bacterium, is the source of the plant disease known as early blight. It appears as little black specks that grow into huge, brown to black, ovoid to spherical lesions. These lesions may be restricted by leaf veins or connected to lenticels. A black fungus can also grow on the undersides of plant leaves. Tuber wilt in potatoes can result from early blight. The disease tends to spread when temperatures exceed 26°C. It frequently happens when the potato plants' vigor is diminished because of elements like intense heat drying or inadequate fertilizer.

Fig. 1. Early Blight Disease of Potato Leaf

Late Blight: The bacterium Phytophthora infections is responsible for the plant disease known as late blight. Outbreaks of this disease can significantly harm potato crops, especially in years characterized by low temperatures and abundant rainfall. Late blight poses a severe threat to potato production and can lead to substantial yield losses. The disease is known for its rapid spread and destructive impact on the foliage and tubers of potato plants.

Fig. 2. Late Blight Disease of Potato Leaf

3.1.2 Data Pre-Processing

To enhance the quality of the data, pre-processing techniques were applied. This involves cropping out areas of the photograph that were unnecessary or did not accurately depict the area of concern, thus reducing noise within the image. Images with excessive noise were excluded from the dataset, as they could adversely affect the accuracy of the analysis. Additionally, to ensure consistency, input photos from various sources with different sizes were scaled to a standardized resolution of 256×256 pixels.

3.1.3 Data Augmentation

Data augmentation is a valuable technique employed in this study to expand the dataset without distorting the original meaning of the data. It involves applying geometric transformations to the existing data, for instance, translations, rotations, scale adjustments, and vertical and horizontal flips. By automatically generating parameters for augmentation, this approach enhances the robustness of the dataset and enhances the model's capacity to generalize characteristics and patterns.

3.1.4 Image Classification

Three subsets of the Plant Village dataset were created: training, validation, and testing. These subsets served various functions in evaluating our CNN model's performance. The validation and test datasets were used to determine the model's accuracy after it had been trained using the training dataset. To split the dataset, 80% of the images were utilized for training, 10% for validation, and the remaining 10% were used for testing. From the Plant Village dataset, we took 1640 photos for training, 256 for validation, and 256 for testing. To ensure the model's generalization and mitigate overfitting, we applied various data augmentation techniques exclusively to the training set. These techniques included rescaling, rotation, contrast adjustment, horizontal flipping, and vertical flipping.

Using the training samples, the CNN model was trained, starting at the input layer, and moving through the successive levels until it reached the output layer. Predictions were made, and any errors or discrepancies were determined. In cases where incorrect predictions occurred, we employed the back-propagation algorithm in reverse order to modify the model's weights for improved accuracy. Each complete iteration of this forward and backward process, known as an epoch, contributed to the model's learning.

Machine learning (ML), also known as deep learning (DL), or deep neural learning, is a crucial component of artificial intelligence (AI) [19]. Deep learning, as denoted by the term "deep," involves architectures with multiple layers. This advancement in ML has significantly impacted various domains, including object detection, speech recognition, object categorization, and image classification. Among the popular classes of deep learning, Convolutional Neural Networks (CNNs) have gained considerable popularity. Numerous studies have utilized CNNs to identify plant diseases based on leaf health [20]. In line with this approach, our aim was to develop a robust CNN model capable of accurately classifying and predicting image class labels within the Plant Village dataset. Specifically, we focused on identifying early blight, healthy, and late blight categories.

3.2 Proposed CNN Model

A custom Convolutional Neural Network (CNN) is an architecture specifically designed or modified for a particular task or dataset. Unlike pre-trained models, a custom CNN is built from scratch, allowing for customization of the network's layers, sizes, and configurations.

This tailored approach offers several advantages. Firstly, the architecture can be optimized for the specific characteristics of the dataset, resulting in improved performance. Secondly, by reducing unnecessary parameters, custom CNNs can be computationally efficient, leading to faster training and inference times. Additionally, the interpretability of custom CNNs allows for a better understanding of the learned features and decision-making processes.

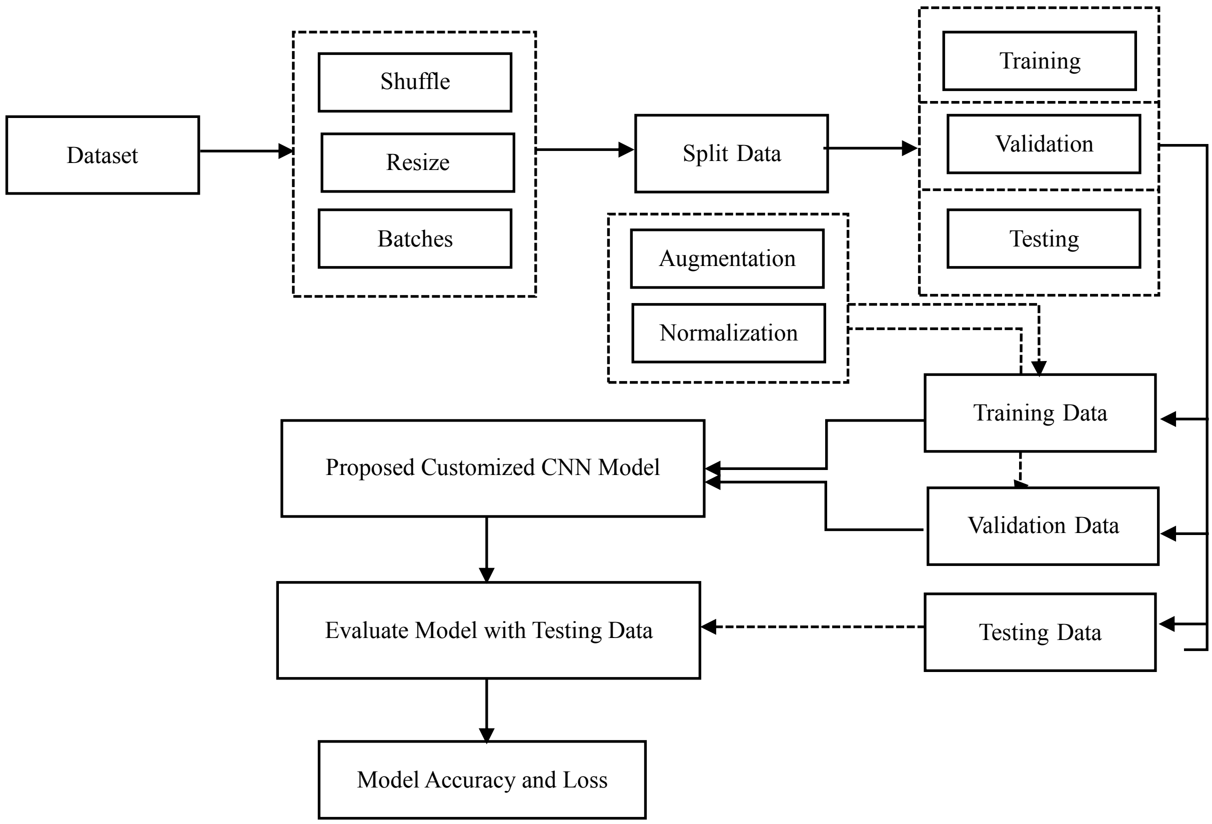

Fig. 3. Workflow Diagram of Proposed CNN Model

However, when developing a custom CNN, factors such as dataset size, problem complexity, and available computational resources need to be considered, alongside proper hyperparameter tuning for optimal performance. Ultimately, custom CNNs provide flexibility and control, enabling the creation of neural network architectures that effectively address specific tasks and datasets, leading to enhanced performance and insights compared to generic pre-trained models [19]. Our study utilized a customized CNN (Convolutional Neural Network) to accurately identify potato blight on leaves. The entire classification procedure for potato blight consisted of several steps: data normalization, data augmentation, data splitting, training, validation, and testing. A flowchart depicting the working method can be found in Fig. 3.

The sequential model is designed to extract significant features from the input image through a series of layers. It starts with a convolutional layer that takes an input image with dimensions 256 × 256 × 3. This layer applies 32 filters with a 3 x 3 size and makes use of the ReLU activation function. The convolution is performed with a stride of 1 and padding to maintain the output size. After the first convolutional layer, there is a max-pooling layer with a pool size of (2,2) that is utilized to decrease the image dimensions. The subsequent three convolutional layers each employ 64 filters with a 3 × 3 size and ReLU activation. These layers have a stride of 1 and do not use padding. After these convolutional layers, another max-pooling layer with (2,2) as the pool size is applied to further down-sample the feature maps. The fifth and final convolutional layer consists of 128 filters with a size of 3 × 3. Here the ReLU activation function is utilized and applies padding to maintain the output dimensions. Like previous layers, the stride is set to 1. To prepare the data for the fully connected layers, the flattened layer converts the convolved feature maps into a 1D vector.

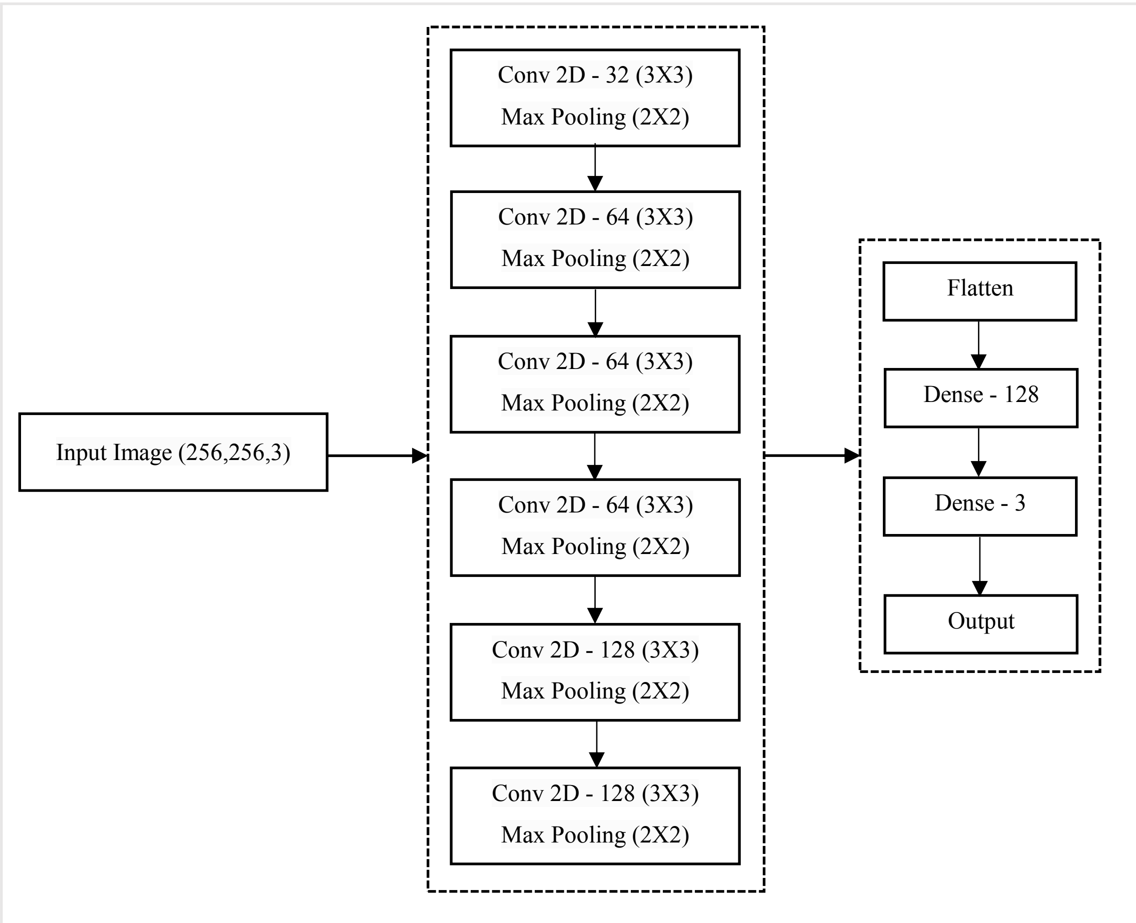

Fig. 4. The Architecture of Our Proposed Customized CNN Model

Two fully connected layers are used for classification based on the extracted features. The first hidden layer consists of 128 neurons with ReLU activation. The quantity of classes in the multi-classification problem dictates how many neurons are in the output layer. In this case, with three classes, the output layer has three neurons. The softmax activation function is applied to produce class probabilities. Class labels are predicted in the output layer and compared to the true labels to evaluate the model's overall accuracy. In summary, the proposed CNN architecture employs a sequential model with convolutional and max-pooling layers to extract essential features from input images. These features are then processed through fully connected layers for classification. The output layer's class label prediction is used to determine the model's accuracy. The details of our proposed CNN architecture can be found in Fig 4.

3.3 Evaluation Measures

3.3.1 Confusion Matrix

A confusion matrix serves as a valuable tool for measuring a classification model's effectiveness by providing a concise summary of its predictions. It presents a tabular representation that showcases the counts or percentages of true negatives (TN), true positives (TP), false negatives (FN), and false positives (FP) for every class in the dataset. In our dataset, there are three possible classifications. So, 3x3 matrixes are generated for each machine-learning technique. In the confusion matrix, TP (True Positive) describes the number accurately predicted for a certain class. TN (True Negative) represents how many samples there are that don't fit into a particular category and were accurately identified as not falling within that class. In other words, TN is the count of samples that are accurately classified as negative for a particular class. FN (False Negative) reflects how many samples there are that should belong to one class but were mistakenly projected to belong to another. FP (False Positive) represents the number of samples that do not associated with a specific class but were wrongly predicted as that class. Every cell in the matrix displays the number of samples based on the true class and the anticipated class that fits into a specific category. The samples that were successfully identified are represented by the diagonal elements (TP), whereas the incorrectly classified samples are shown by the off-diagonal elements (FN and FP).

3.3.2 Accuracy

Accuracy is a crucial performance metric that can be calculated based on the information provided by the confusion matrix. It serves as a measure of the overall correctness of a model's predictions, encompassing all classes present in the dataset.

The ratio of the sum of true negative (TN) and true positive (TP) to the total number of samples is used to calculate accuracy:

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

In the context of a confusion matrix, accuracy evaluates how effectively the model distinguishes between healthy and unhealthy potato leaves. A high accuracy value indicates that the model has made correct predictions for most of the samples, while a lower accuracy suggests that the model has a higher rate of misclassifications. Keep in mind that accuracy may not always provide a whole view of the model's performance, especially when working with datasets that are unbalanced or when the costs of different types of errors are significantly different. In these circumstances, it is advised to consider additional evaluation measures for a more thorough analysis, such as F1-score, recall, precision, or area under the ROC curve (AUC). While analyzing the results, it is beneficial to discuss the accuracy along with other evaluation metrics to give a further nuanced comprehension of the model's output in potato leaf disease detection. To accuracy measurement, we used the following performance parameters:

3.3.3 Precision

Precision represents the fraction of accurate positive predictions (true positives) within all positive predictions (true positives and false positives). It focuses on how well the model correctly predicted the positive outcomes.

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

A greater precision score means that the model is more accurate at identifying diseased potato leaves because it has a lower rate of false positives and minimizes misclassifications of healthy leaves as diseased. In the context of potato leaf disease detection, precision represents the accuracy of the model in correctly identifying diseased leaves. It is an important metric as misclassifying healthy leaves as diseased can have practical implications in terms of unnecessary treatments or resource allocation.

3.3.4 Recall

Recall quantifies the percentage of precisely predicted positive samples (true positives) among all the positive samples (true positives and false negatives). It focuses on the model's capacity to find all positive instances.

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

A greater recall value suggests that the model has a reduced rate of false negatives, meaning it is better at correctly identifying diseased potato leaves and minimizing the instances where it fails to detect actual diseased leaves. In the context of potato leaf disease detection, recall represents the model's ability to correctly identify and capture diseased leaves. It is an important metric as missing diseased leaves can lead to potential crop damage or the spread of diseases.

3.3.5 F1 Score

The F1 Score is a measurement that merges Precision and Recall forming a unified value, incorporating the occurrences of both false positives and false negatives. While accuracy is straightforward, the F1 Score is often more beneficial, especially when faced with disproportionate class distributions. Accuracy performs well when false positive and false negative rates are similar. However, when there is a substantial difference in the occurrence of false positives and false negatives, it becomes more informative to evaluate both Precision and Recall simultaneously.

$$F1\ Score = \frac{2 \times (Recall \times Precision)}{(Recall + Precision)}$$

3.3.6 ROC Curve

The ROC curve is a graphical representation that displays the performance characteristics of a binary classification model in a unique and distinct manner. It demonstrates the balance between the true positive rate (sensitivity/recall) and the false positive rate. By plotting the TPR against the FPR, the ROC curve offers a holistic assessment of how well the model's proficiency in discerning between the positive and negative classes is assessed across various classification thresholds.

TPR, also known as True Positive Rate, is an equivalent term for recall and is defined accordingly:

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

The following is the definition of FPR:

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

A model's performance is shown graphically by the ROC curve, which has a well-performing model with a curve near the top-left corner, suggesting a high true positive rate (TPR) and a low false positive rate (FPR) across different thresholds. The proximity of the ROC curve to the top-left corner serves as an indicator of the model's superior performance.

3.3.7 AUC Curve

The AUC curve, short for Area Under the Curve, is a measure created from the ROC curve, utilized for gauging the effectiveness of binary classification models. It quantifies the overall discriminatory power of the model in differentiating between positive and negative classes across all classification thresholds. The AUC value varies between 0 and 1, with a higher value indicating superior performance. An AUC of 1 represents a perfect classifier with 100% sensitivity and no false positives, while an AUC of 0.5 signifies a random classifier that doesn't perform any better than random. The AUC metric offers a succinct assessment of a model's proficiency in precisely ordering positive and negative instances, facilitating the comparison and selection of classification models.

A high AUC value suggests that the model has good classification performance across various thresholds, meaning it can effectively differentiate between positive and negative instances. In contrast, a low AUC value signifies inadequate discriminative capability of the model between the classes. Furthermore, The ROC curve and AUC provide valuable insights into the discrimination capability of a model, allowing for a nuanced evaluation of its performance, especially when the classification threshold is variable or uncertain.

4. Results and Discussion

4.1 Proposed Customized CNN Model Performance

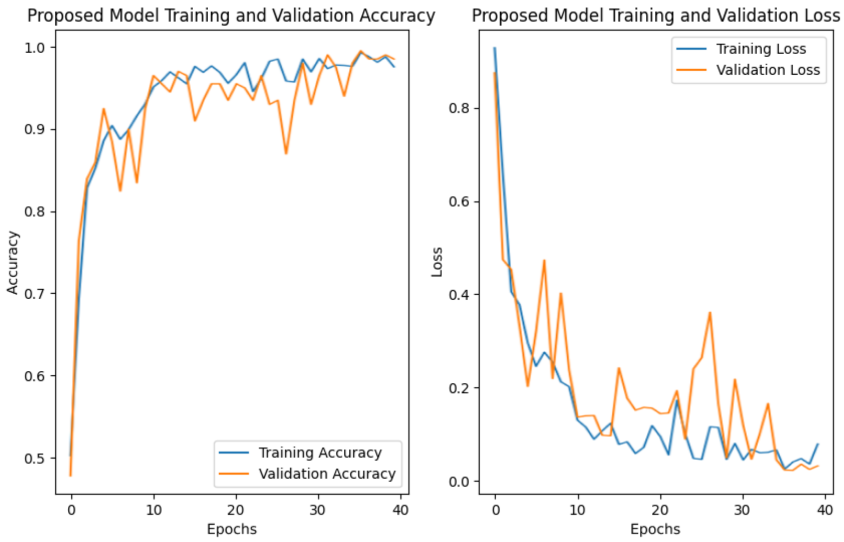

Fig. 5. Accuracy and Loss Curves of Proposed Customized CNN Model.

Our customized CNN model achieved high accuracy, precision, recall, and F1 score for potato leaf disease detection. With 99.22% accuracy, it performs similarly to established architectures like VGG16 and MobileNet, demonstrating its effectiveness in accurately identifying and classifying potato leaf diseases. By training on a comprehensive dataset and fine-tuning parameters, we optimized its performance and generalization to unseen data. The graphical representation of our proposed model's accuracy and loss is shown in Figure 5. The accuracy curve displays the accuracy values at each epoch on the y-axis, while the epochs are represented by the x-axis. Similarly, the loss curve illustrates the connection between the number of epochs and the corresponding loss values, with the y-axis indicating the loss and the x-axis denoting the epochs. As observed in the accuracy curve, the model's accuracy on the training data starts at 48.38% during the first epoch and steadily improves over the subsequent epochs, reaching 97.45%at the end of 40 epochs. The accuracy on the validation data shows a similar trend, starting at 45.83% and progressing to 98.44% by the final epoch.

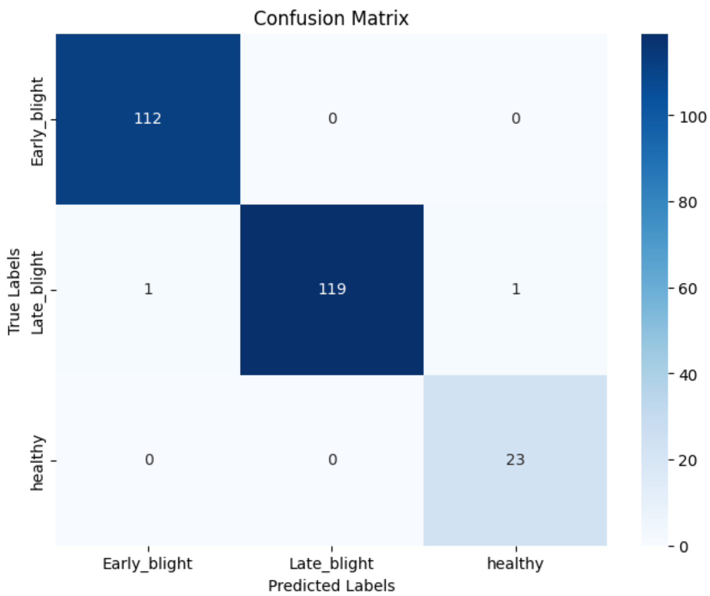

Fig. 6. Confusion Matrix of Our Proposed Customized CNN Model

In terms of loss, the model's training loss begins at 0.9321 and gradually decreases with each epoch, reaching 0.0843 by the 40th epoch. The validation loss follows a similar pattern, starting at 0.8785 and converging to 0.0375. These curves demonstrate our customized CNN model's effectiveness in learning from the training data and generalizing well to unseen validation data, capturing relevant patterns, and making accurate predictions. Now we will discuss the results along with the performance measurement techniques. Confusion matrix evaluates the performance of our proposed model for each class, providing insights into accuracy, precision, recall, and other metrics. It helps us analyze the model's performance, and identify error types, and areas for improvement. The confusion matrix of our proposed model is shown in Figure 6.

The confusion matrix, displayed as a heatmap, represents the number of samples predicted for every class label. In the test dataset of 256 samples, distributed among early blight (112 samples), late blight (121 samples), and healthy (23 samples), our model accurately detects 112 cases of early blight with no false detections. For late blight, it accurately detects 119 cases but has one false detection for early blight and one for healthy. The model accurately detects all 23 healthy samples with no false detections.

The confusion matrix analysis enables the calculation of precision, recall, and F1-score for every class, offering valuable information about the accuracy and misclassification patterns of the model. Please refer to Table 3. for the precision, recall, and F1-score metrics of our proposed customized CNN model.

Table 2. Precision, Recall, F1-Score of our Proposed Customized CNN Model

Performance Measures	Early Blight	Late Blight	Healthy	Average
Precision	99.11%	100%	95.83%	98.31%
Recall	100%	98.34%	100%	99.44%
F1-Score	99.55%	99.16%	97.87%	98.86%

The Customized CNN model that we proposed demonstrated exceptional performance with an accuracy of 99.22%. Moreover, it exhibited impressive precision, recall, and F1 score values of 98.31%, 99.44%, and 98.86% respectively. These findings emphasize the efficacy and dependability of our model in precisely categorizing and forecasting the target classes.

Precision, recall, and F1-score are essential metrics employed for evaluating the performance of a classification model. Our model achieved a precision of 98.31%, indicating a high accuracy in positive predictions. The recall value of 99.44% demonstrates the model's potentiality to point out positive samples effectively. With an F1-score of 98.86%, the model showcases an overall accurate classification performance. These metrics are particularly important when dealing with imbalanced datasets.

The ROC curve represents the efficiency of a classification model in a graphical format, specifically for binary classification tasks. It demonstrates how, at different categorization criteria, the true positive rate (TPR) and the false positive rate (FPR), trade off against one another. Fig. 7 displays the ROC-AUC curve of our proposed customized CNN model, revealing details about its capacity to distinguish between positive and negative occurrences across various threshold settings. The ROC curve is computed and plotted for each class in the classification problem. It demonstrates the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) in a graphical manner. The AUC values calculated from The ROC curve show how well the model can differentiate between classes. For Class 0 (early blight potato leaf), the AUC value is 0.73, suggesting a moderate ability to differentiate from other classes. Class 1 (late blight potato leaf) has an AUC value of 0.76, indicating slightly better performance. Class 2 (healthy potato leaf) has an AUC value of 0.70, indicating a lower ability to differentiate from other classes. Additionally, our model's precision value of 98.31% signifies high accuracy in predicting positive samples and minimizing false positives.

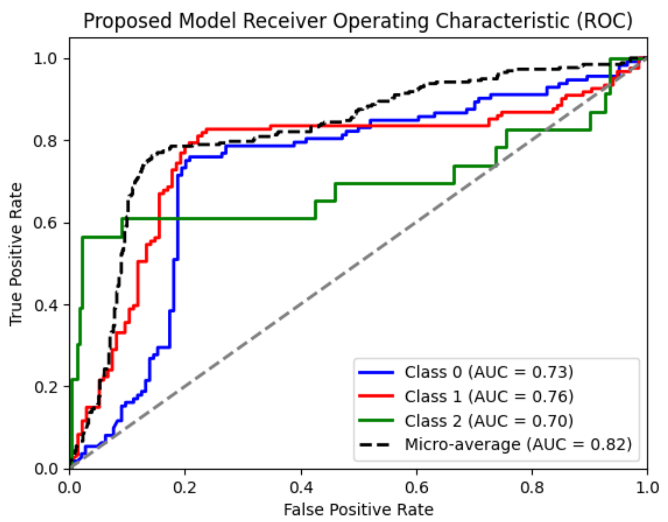

Fig. 7. ROC-AUC Curve of Our Proposed Customized CNN Model

The micro-average AUC value is 0.82, which provides an aggregate measure across all classes. This value indicates the overall discriminative ability of the model across the entire classification task. A higher micro-average AUC value suggests better overall performance in classifying instances from different classes. By considering these AUC values along with the ROC curve, we can acquire a valuable understanding of how well the model is able to classify instances for each class and the overall classification performance. The goal is to have higher AUC values for each class and the micro-average, indicating better discrimination between classes and improved overall classification accuracy.

4.2 Comparison with Other Pre-trained CNN Models

Performance metrics were evaluated for five different methodologies or classifiers: VGG16, VGG19, ResNet50, MobileNet, and our proposed customized CNN model. The VGG16 model achieved an accuracy of 99.22% with high precision and recall. VGG19 achieved an accuracy of 97.27% with balanced precision and recall. ResNet50 achieved 98.83% accuracy and demonstrated high recall. MobileNet had the highest accuracy of 99.61% with impressive precision and recall.

Table 3. Performance Measurement using Performance Matrixes

Classifier	Accuracy	Precision	Recall	F1 Score
VGG16	99.22%	99.45%	97.10%	98.21%
VGG19	97.27%	98.28%	97.89%	98.03%
ResNet50	98.83%	94.73%	99.09%	96.68%
MobileNet	99.61%	98.03%	99.69%	98.83%
Proposed Customized CNN Model	99.22%	98.31%	99.44%	98.86%

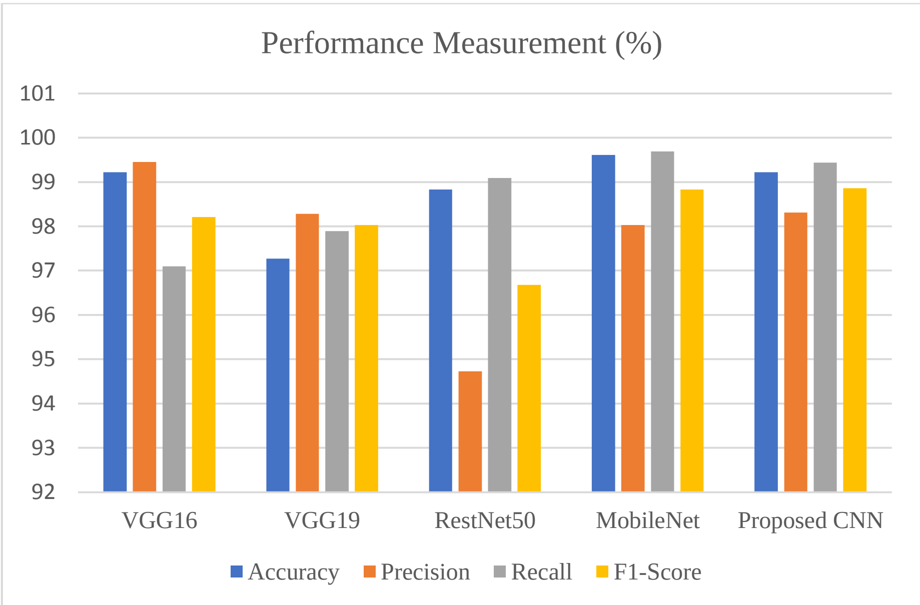

Fig. 8. Performance Measurement using Performance Matrixes

Our proposed customized CNN model achieved 99.22% accuracy, like VGG16, with balanced precision and recall. The MobileNet model had the highest accuracy, but our customized CNN model performed exceptionally well, showcasing high precision, recall, and F1 scores. These performance measurements provide insights into the effectiveness of each methodology or classifier. Performance comparison using Performance Matrixes is shown in table 3 and its graphical representation is shown in fig. 8.

4.3 Result Discussion

Deep learning models have been shown to be effective in detecting potato leaf diseases. In our study, we conducted a comprehensive evaluation of various deep learning models, such as VGG16, VGG19, ResNet50, MobileNet, and a customized CNN model. These models underwent training and evaluation using a dataset comprising potato leaf images with predefined disease labels. The outcomes revealed that MobileNet has the greatest accuracy and recall scores, indicating its effectiveness in accurately detecting potato leaf diseases. However, the customized CNN model also demonstrated competitive performance, achieving similar accuracy and F1 score values compared to VGG16 and MobileNet. This suggests that the customized model can be considered as a viable alternative, offering a good balance between accuracy and computational efficiency.

The precision score of 98.31% and recall score of 99.44% for the customized model demonstrate its effectiveness in correctly identifying positive instances of potato leaf diseases while minimizing false positives. The F1 score of 98.86% further confirms the balanced performance of the model in terms of precision and recall. These results highlight the model's capability to accurately detect diseased plants and make it a reliable choice for disease management. The successful application of deep learning models, such as the proposed customized CNN model, validates their efficacy in potato leaf disease detection. Moreover, the model's ability to strike a better balance between accuracy and computational efficiency makes it a practical option for real-world applications. Additionally, the robust performance of our customized CNN model underlines its adaptability to diverse agricultural settings and provides an opportunity for its integration into precision agriculture systems. The model's ability to automate disease detection processes can significantly reduce the workload on farmers and agronomists, allowing for timely intervention and more effective disease control strategies.

5. Conclusion

To summarize, our research has leveraged sophisticated deep-learning techniques to develop a robust system capable of automatically detecting and classifying potato leaf diseases. The implementation of our customized CNN model resulted in a remarkable accuracy rate of 99.22%, effectively distinguishing between early blight, late blight, and healthy leaf conditions. Comparative analysis, as depicted in Fig 8, demonstrates the superior performance of our system compared to existing methods. These findings validate the efficacy and potential impact of our approach in potato disease detection and management. Furthermore, our work contributes significantly to the field of agricultural technology by addressing persistent challenges related to the accuracy and speed of disease detection in potato crops. By tailoring a CNN model to enhance accuracy while reducing computational demands, we have paved the way for more efficient and reliable disease management practices. The application of our automated system on smartphones, as part of our future work, holds great promise in providing accessible and timely disease detection support to farmers. This technological integration can revolutionize potato farming by empowering farmers with real-time information and recommendations, ultimately leading to enhanced agricultural productivity and sustainability. In conclusion, our research stands at the forefront of agricultural innovation, offering not only a highly accurate disease detection solution but also a pathway towards practical implementation for the benefit of farmers and global food security.

Acknowledgment

We would like to thank our honorable supervisor, Dr. Nasrin Sultana, Associate Professor, Department of Electronics and Communication Engineering (ECE), and honorable co-supervisor, Md. Selim Hossain, Lecturer, Department of Electronics and Communication Engineering (ECE) for their leadership, inspiration, and helpful criticism throughout the course of this project. Even though they are constantly busy with other things, they provided us with more than enough time to complete the work. They not only offered us their time but also appropriate direction and helpful suggestions whenever we ran into problems. Their suggestions and comments aided us in completing this project.

References

[1] A. Drewnowski and C. D. Rehm, "Vegetable cost metrics show that potatoes and beans provide most nutrients per penny," PLoS One, vol. 8, no. 5, p. e63277, 2013.

[2] A. M. Vargas, L. M. Q. Ocampo, M. C. Céspedes, N. Carreño, A. González, A. Rojas, A. P. Zuluaga, K. Myers, W. E. Fry, P. Jiménez, et al., "Characterization of Phytophthora infestans populations in Colombia: First report of the A2 mating type," Phytopathology, vol. 99, pp. 82–88, 2009.

[3] W. E. Fry, "Phytophthora infestans: New Tools (and Old Ones) Lead to New Understanding and Precision Management," Annu. Rev. Phytopathol., vol. 54, pp. 529–547, 2016. DOI: 10.1146/annurev-phyto-080615-095835. [CrossRef] [PubMed]

[4] B. Ilbery, D. Maye, and R. Little, "Plant disease risk and grower–agronomist perceptions and relationships: an analysis of the UK potato and wheat sectors," Applied Geography, vol. 34, pp. 306–315, 2012.

[5] P. TM, P. Alla, K. S. Ashirta, N. B. Chittaragi, and S. G. Koolagudi, "Tomato leaf disease detection using convolutional neural network," in International Conference on Contemporary Computing (IC3), 2018.

[6] M. A. Khan, T. Akram, M. Sharif, K. Javed, M. Raza, and T. Saba, "An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection," Multimed. Tools Appl., vol. 79, pp. 18627–18656, 2020. DOI: 10.1007/s11042-019-08099-9. [CrossRef]

[7] Azrin Zahan, Md. Selim Hossain, Ziaur Rahman and SK. A. Shezan, "Smart home IoT use case with elliptic curve based digital signature: an evaluation on security and performance analysis", International Journal of Advanced Technology and Engineering Exploration, Vol 7(62),2020, http://dx.doi.org/10.19101/IJATEE.2019.650070

[8] D. Tiwari and M. Ashish, "Potato Leaf Diseases Detection Using Deep Learning," in Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020), IEEE Xplore Part Number: CFP20K74-ART, ISBN: 978-1-7281-4876-2.

[9] M.H. Al-Adhaileh, A. Verma, T.H.H. Aldhyani, and D. Koundal, "Potato Blight Detection Using Fine-Tuned CNN Architecture," Mathematics, vol. 11, no. 6, p. 1516, Mar. 2023. [Online]. Available: https://doi.org/10.3390/math11061516

[10] D. Tiwari, M. Ashish, and N. Gangwar, "Potato Leaf Diseases Detection Using Deep Learning," in Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020), ISBN: 978-1-7281-4876-2, IEEE Xplore Part Number: CFP20K74-ART.

[11] Hossain, Md. S., Md. H. Rahman, Md. S. Rahman, A. S.M.S. Hosen, Changho Seo, and Gi H. Cho 2021. "Intellectual Property Theft Protection in IoT Based Precision Agriculture Using SDN" Electronics 10, no. 16: 1987. https://doi.org/10.3390/electronics10161987

[12] P. Garlapati, A. Kuruba, et al., "Detection of Disease and Damage Control for Crops using Convolution Neural Networks," Journal of Xi'an University of Architecture & Technology, ISSN: 1006-7930, 2020.

[13] S. Mangal and P. Meshram, "PLANT DISEASE IDENTIFICATION USING DEEP LEARNING CLASSIFICATION MODEL: CNN," Journal of University of Shanghai for Science and Technology, vol. 23, no. 1, pp. January, 2021. ISSN: 1007-6735.

[14] V. P. Gaikwad and V. Musande, "Potato Plant Leaf Disease Detection Using CNN Model," European Chemical Bulletin, vol. 12, no. 1, pp. 516-527, 2023.

[15] R. Mahum, H. Munir, Z. U. N. Mughal, M. Awais, F. S. 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, 2022. DOI: 10.1080/10807039.2022.2064814.

[16] K. K. Chakraborty, R. Mukherjee, C. Chakroborty, and K. Bora, "Automated recognition of optical image based potato leaf blight diseases using deep learning," Physiological and Molecular Plant Pathology, vol. 117, p. 101781, 2022.

[17] W. Chen, J. Chen, A. Zeb, S. Yang, and D. Zhang, "Mobile convolution neural network for the recognition of potato leaf disease images," Multimedia Tools and Applications, vol. 2022, pp. 1-20.

[18] Kaggle, "Plant Disease Datasets," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease.

[19] P. Shrivastav, V. Avsthi, et al., "Plant Disease Detection using Convolutional Neural Network," International Journal of Advanced Research (IJAR), January 2021.

[20] F. Islam, Md. H. Rahman, N. Nurjahan, Md. S. Hossain, and S. Ahmed, "A Novel Method for Diagnosing Alzheimer's Disease from MRI Scans using the ResNet50 Feature Extractor and the SVM Classifier," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 14, no. 6, 2023. http://dx.doi.org/10.14569/IJACSA.2023.01406131.

Authors' Profiles

Abdullah Walid is an undergraduate student studying Electronics and Communication Engineering (ECE) at Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur-5200. His academic passions encompass Artificial Intelligence (AI) and Machine Learning, Data Science and Big Data, Computer Graphics and Visualization, Web Performance Optimization, Database Management Systems (DBMS), Data Privacy and Security, Photonic crystal fiber. He's really excited about using his studies to help make new discoveries and innovations in these areas through hard work and research.

Md. Mehedi Hasan is an undergraduate student at the Department of Electronics and Communication Engineering (ECE) at Hajee Mohammad Danesh Science And Technology University (HSTU), Dinajpur-5200. His academic interests lie in the fields of Machine Learning, the Internet of Things, Artificial Intelligence, Cybersecurity, Software Defined Wireless Networks, Bioinformatics, and Biomedical Image Processing. He is passionate about contributing to the advancement of these areas through academic research and innovation.

Tonmoy Roy is an undergraduate student at the Department of Electronics and Communication Engineering (ECE) at Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur-5200. His academic interests encompass cutting-edge areas such as Cybersecurity, Machine Learning, Software-Defined Networking, Computational Biology, and Medical Imaging Analysis. He is fervently dedicated to advancing knowledge in these domains through academic research and innovative initiatives.

Md. Selim Hossain is currently working as a Lecturer in the Department of Electronics and Communication Engineering (ECE), Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur-5200, Bangladesh. He was an Assistant Network Engineer in the Department of ICT (DoICT), under the Ministry of Posts, Telecommunications and Information Technology, Government of the People's Republic of Bangladesh. He served as a Senior Lecturer at the Department of Computing and Information Systems at Daffodil International University (DIU) and was awarded the best research award in 2022. He was also a Lecturer in the Department of Computer Science and Engineering at Khwaja Yunus Ali University, Sirajganj, Bangladesh from October 24th, 2016, to February 1st, 2021. He completed his B.Sc. degree in Telecommunication and Electronic Engineering from Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh with 2nd position in 2014 and his M.Sc. (Engg.) in Information and Communication Technology from Mawlana Bhashani Science and Technology University, Tangail, Bangladesh with 3rd position in 2017. He has published more than 40 research papers in different international journals and conferences. His main research interests are based on IoT, Blockchain, Photonics, Artificial Intelligence, Cryptography, and Network Security.

Dr. Nasrin Sultana received a B. Sc. from Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur-5200, Bangladesh in 2009 and a master’s from Bangladesh University of Engineering and Technology (BUET), Dhaka-1205, Bangladesh in 2014. She was awarded the Doctor of Philosophy (Ph.D.) degree from the Graduate school of Science and Engineering, Saitama University, Japan in 2020. Now she is working as an Associate Professor in the Department of Electronics and Communication Engineering, HSTU, Dinajpur. Her research interest is high-speed optical measurement and sensing in ultrafast photonic science.

How to cite this paper: Abdullah Walid, Md. Mehedi Hasan, Tonmoy Roy, Md. Selim Hossain, Nasrin Sultana, "Deep Learning-Based Potato Leaf Disease Detection Using CNN in the Agricultural System", International Journal of Engineering and Manufacturing (IJEM), Vol.13, No.6, pp. 9-22, 2023. DOI:10.5815/ijem.2023.06.02