Journal of Advances in Mathematics and Computer Science

Volume 37, Issue 12, Page 143-155, 2022; Article no.JAMCS.95283
ISSN: 2456-9968
(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

A Deep Learning Approach to Classify the Potato Leaf Disease

Md. Ashraful Islam [a++#] and Md. Hanif Sikder [b*]

[a] *School of Engineering, University of South Asia, Dhaka, Bangladesh.*
[b] *Department of Computer Science & Engineering, University of South Asia, Dhaka, Bangladesh.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMCS/2022/v37i121735

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/95283

Original Research Article

Received: 18/10/2022
Accepted: 26/12/2022
Published: 28/12/2022

Abstract

Purpose: This paper aims to classify potato disease using convolutional neural network in different epochs to observe the best performance of the model. The best model will help the farmers to make different decisions to prevent the loss of potato production.
Methodology: The paper implements a deep learning approach, the convolutional neural network, to explore potato disease classification. To accomplish the research objective, we collected 10000 images of potato leaves from different sources like google and raw data from potato fields. We collected a dataset of 2152 images from Kaggle and the other 7848 images from the above sources. The dataset belongs to a few classes. The classes are Potato Early Blight, Potato Late Blight, and Potato healthy leaf. The paper includes four main steps: data acquisition, data pre-processing, data augmentation, and image classification to find the output.
Findings: This study found that the model performed better when we applied 40 epochs for the 10000 images dataset & we achieved 100% accuracy as we applied a total of 3 different epochs and achieved an accuracy of 99.97% and 99.98% for 30 and 50 epochs, respectively.
Research Limitations: The study significantly contributed to the agriculture sector and farmers by providing suggestions to classify the Potato leaf Disease with the best output.

[++] *Professor;*
[#] *Dean;*
**Corresponding author: Email: hanifskdr@gmail.com;*

Besides, researchers need more raw data to build the model for better output, and they also should be concerned regarding the system when working with large volumes of data as it takes longer to run the code.
Originality/Value: This research paper contained high volume of the dataset, which is 10000 images of potato leaves. We collected a dataset of 2152 images from Kaggle and the rest, 7848 images from different sources like google, and raw data from potato filed. We showed different epochs to check the best performance and achieved 100% accuracy when 40 epochs were applied.

Keywords: AI; potato disease classification; deep learning; convolutional neural network.

1 Introduction

“Potatoes (Solanum tuberosum) are the world’s most important vegetable crop. Due to the vast diversity in types and high consumer consumption, potatoes are a good enterprise option for many growers. Around 130 and 95 developing nations grow potatoes, the world’s fourth most important staple food. The output of potatoes worldwide has been steadily increasing over the past year, including in emerging nations. However, it is also estimated that over 32% of potatoes are lost annually due to illnesses and pests” [1].

“Potato farmers in Bangladesh lose at least Tk 2,500 crore every year due to unsold surplus production and post-harvest losses” [2]. In order to help our framer, we have proposed a model based on convolutional neural network to classify potato leaf disease.

“Potato farming dominates as an occupation within the agriculture domain in additional than 125 countries. However, even these crops are subjected to infections and diseases, primarily categorized into two grades: (i) Early blight and (ii) Late blight. Moreover, these diseases cause damage the crop and reduce its production. In fact, potatoes have a more favorable overall nutrient-to-price ratio than many other fruits and vegetables and are an affordable source of nutrition worldwide” [3].

“Late blight damages leave, stems, and tubers. The leaves affected by this disease appear blistered and dry out. When drying out, leaves turn brown or black in color. The remedy to the matter is high humidity, cold, and leaf wetness. The primary blight could also be a specific disease occurring on the foliage at any stage of the expansion, causing characteristic leaf spots and blight. The primary blight is first observed on the plants as small black lesions totally on the older foliage. Lesions on the stems are quite like those on leaves, sometimes girdling the plant if they occur near the soil line. The remedy to the problem is warm, rainy, and wet weather” [4].

“Identifying diseases in potato plants quickly and accurately is important to chop back the impact of diseases on plants. Manual monitoring activities dispensed by farmers become difficult and impractical because it takes an extended time and in-depth knowledge. Identification of plants diseases types that are slow will trigger the spread of diseases in plants uncontrollably. Besides, farmers generally identify diseases in plants in some way that’s approximately and assumptions that allow inaccurate identification results because the symptoms on the leaves appear to possess similarities that are difficult to elucidate at a glance”. [5] “Farmers use the results of personal identification without expert advice within the sector of plant diseases as a reference for preventing plants infected with the disease. As a result, preventive measures taken by farmers could even be ineffective and will damage crops thanks to inadequate knowledge and misinterpretation of disease intensity, excessive dosage, or lack of dosage” [5]. This problem is that the inspiration of the proposed research is to facilitate farmers in identifying and classifying diseases in potato plants that are fast and accurate.

This paper presents a Convolutional Neural Network based approach to identify and classify two common potato infections. Farmers can easily detect the disease in potato crops using the proposed method with little computational effort.

2 Literature Review

Recent years have introduced new rice farming equipment that can automatically gauge the water temperature and level in the paddy field. A cultivation management system was also suggested to compile the expertise of farmers. However, there are other issues in biological information sensing, such as physiological and growing circumstances.

"Hokkaido Agricultural Research Center develops unmanned system Agricultural machinery (rice planters, tractors, etc.) equipped with remote monitoring functions" [6]. This system will be linked to Quasi-Zenith Satellites by applying object detection and straight-line tracking.

"Faria and other authors presented a framework for classifier fusion that can support automatic fruit and vegetable recognition in a supermarket setting. The authors demonstrate that the proposed framework outperforms several related works in the literature" [7].

"Geetharamani & the authors proposed a nine-layer CNN for leaf disease classification. They claimed that their model outperforms traditional approaches in terms of accuracy" [8].

Bouaziz et al. proposed "a deep learning-based approach that automates the process of classifying banana leaf diseases" [9].

The most cultivated and in-demand crop after rice and wheat is Potato. The potato is native to the Peruvian-Bolivian Andes and is one in every of the world's leading food crops. Potatoes are frequently served whole or mashed as a cooked vegetable and also are ground into potato flour, utilized in baking and as a thickener for sauces. The tubers are highly digestible and provide ascorbic acid, protein, thiamin, and niacin.

Shen and the authors proposed, since the majority of the existing plant disease grading is done by eye, a novel method based on computer image processing has been devised. Following an analysis of all relevant parameters, the leaf portion of the image was segmented using the Otsu method. In the HSI color system, the H component was chosen to divide the illness spot in order to lessen the disruption caused by changes in lighting and veins. The Sobel operator was then used to segment disease spot regions so that disease spot edges could be looked at. Finally, the ratio of disease spots and leaf areas is calculated to provide a grade to plant illnesses. According to studies, this method of grading plant leaf spot infections is quick and precise [10].

Appasaheb & the authors provide an overview of leaf parameter analysis, healthy, sick, or afflicted leaf region detection, and categorization of leaf diseases utilizing various plant types. The precise type of leaf disease must be seen with the naked eye, which is vital and challenging for human eyes. Each plant leaf displays a unique set of disease symptoms. With the leaf of another plant, the algorithm created for one plant does not function well. Along with the leaf parameter analyzer, specialized algorithms are needed to detect leaf diseases in custard apple plants [11].

Patnaree & the authors focus on developing a deep learning-based system to examine and categorize orchid disorders. In this study, we developed a deep learning-based model and evaluated it against three previously trained models: ResNet-50, VGG-16, and VGG-19. Using images of orchid illnesses as datasets, the model is trained and the parameters are adjusted [12].

Rabbia & the authors proposed "algorithm is a novel and first technique to address and report the successful implementation for the detection and classification of four diseases in potato leaves. The algorithm's performance was evaluated on the testing set and gave an accuracy of 97.2%" [13].

Chakraborty & the authors proposed "methodology finally achieved 97.89% accuracy for classification between late and early blight syndromes as compared to healthy potato leaf. This study showed the detailed architecture of the fine-tuned VGG16 model with validation accuracy and losses. Our proposed methodology has also been compared with the existing techniques" [14].

Chen & the authors proposed "procedure delivered a superior performance gain over other compared methods, and it realized an average identification accuracy of 97.73% on different potato disease types. Experimental findings present a competitive performance and prove the validity of the proposed procedure" [15].

3 Deep Learning

Deep Learning (DL) is characterized by a complex hierarchy that connects multiple internal layers for feature detection and representation learning. In the real world, representation learning is used to express the extraction of essential information from observation data. Artificial operations use trial and error to extract features.

However, DL uses the image's pixel level as an input value and acquires the characteristic that is best suited to identifying it [16,17].

A single-layer perceptron network, which has a single layer of output and feeds its inputs straight to its outputs, is the simplest type of neural network (NN). This makes it the most basic variety of feed-forward network.

Convolutional Neural Networks (CNN) use the backpropagation paradigm, similar to a traditional multi-layer perceptron, for their learning process. CNN employs stochastic gradient descent to update the coupling coefficient and weighting filter. In this manner, convolutional and pooling processes are used by CNN to identify the optimal feature [18,19].

Computational models with numerous processing layers can learn representations of data at various levels of abstraction thanks to deep learning. The state-of-the-art in many other fields, including drug discovery and genomics, object identification, visual object recognition, and speech recognition has been significantly enhanced by these techniques. By employing the backpropagation technique to suggest changes to a machine's internal parameters that are used to compute the representation in each layer from the representation in the previous layer, deep learning can uncover detailed structure in massive data sets. Recurrent nets have shed light on sequential data types like text and speech, whereas deep convolutional nets have made advancements in the processing of pictures, video, speech, and audio [20,21].

4 Methodology

As shown in Fig. 1. The proposed methodology in this paper includes the following four main steps: data acquisition, data pre-processing, data augmentation, and image classification [22].

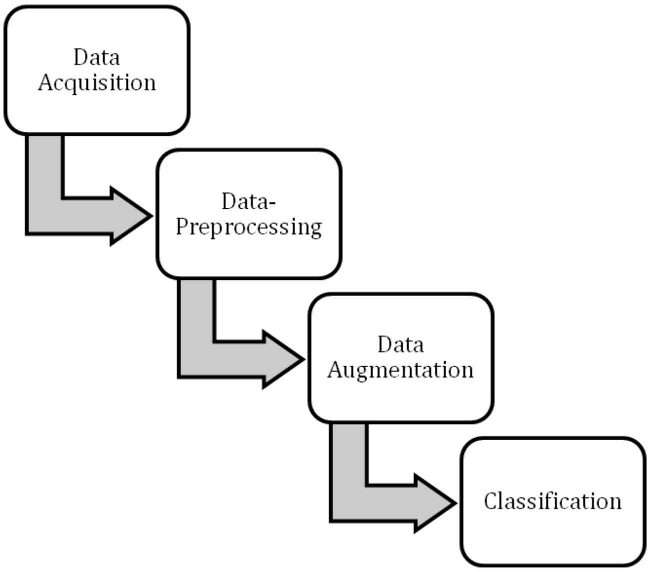

Fig. 1. Proposed methodology

4.1 Data acquisition

Different image resolutions and sizes were obtained from several sources, including those collected by authors from a potato plantation in Patuakhali. We also used an open-access image database from Kaggle. The author collected 7848 images directly from the field and 2152 from Kaggle. A total of 10000 images were used in the paper to perform the research. All the images are divided into three classes. These are Early Blight, Late Blight, and Healthy.

Early Blight: Early blight is a form of plant epidemic brought on by the bacterium Alternaria solani. Tiny black dots grow into massive, brown-to-black, round-to-ovoid lesions, which are sometimes constrained by leaf veins but may also be related to lenticels. The underside of the leaves then develops a black fungus. Tuber wilt in potatoes can be brought on by early blight. When temperatures exceed 26 C, this disease will start to spread. It frequently appears when potatoes' activity is decreased due to high-temperature drying or a lack of fertilizer.

Fig. 2. Early blight disease of potato leaf

Late Blight: Plant infections known as late blight are brought on by the bacterium Phytophthora infestans. A large amount of damage to potato output can be done by outbreaks in years with low temperatures and plenty of rain.

Fig. 3. Late blight disease of potato leaf

Healthy Leaf: Health leaf looks fresh and is not infected with any disease.

4.2 Data pre-processing

Pre-processing Data by removing the portion of the image that is not the region of interest, noise in the image can be reduced. The image will not be utilized if there is too much noise in it. For the dataset, input photos must be scaled to 256x256 pixels after being gathered from various sources and of varying sizes.

4.3 Data augmentation

Data augmentation is a method of modifying data without distorting its original meaning. This study needs to use data augmentation. The automatic application of straightforward geometric transformations, such as translations, rotations, scale changes, shearing, and vertical and horizontal flips, generates the augmentation parameters in this study.

Fig. 4. Healthy potato leaf

4.4 Image classification

Machine learning (ML), also referred to as deep learning (DL), deep neural learning, or deep neural network, is a component of artificial intelligence (AI). Deep learning contains more layers than machine learning, as indicated by the word "deep". Deep learning techniques have raised the bar in several fields, including object detection, speech recognition, object categorization, and image classification [19]. Convolutional Neural Network is one of the most well-liked classes in deep learning. Convolutional neural networks have been used in several research to identify plant illnesses based on the health of the leaves.

One or more convolutional layers that are organized into groups according to function make up convolutional neural networks in general. The subsampling layer is frequently followed by one or more fully linked layers that are typical of a neural network. A feature set contained in a limited area on the previous layer serves as input for each feature layer.

5 Implementation

In this study, we use the dataset from potato leaves to identify the significant diseases of potato leaves for categorization. The collection includes 10,000 photos of 3 classes that appear remarkably similar yet represent distinct diseases. We used python programming language to write the code.

Table 1 displays a list of CNN hyperparameters. The goal is to develop a training model. Hence we have used a maximum of 50 epochs for the training model, with a batch size of 32. The image has been scaled down to 256*256 pixels. The sequential model on which the network is built has four convolutional layers, four pooling layers, and four fully connected layers.

To enhance, we employ Rectified Linear Unit (ReLU) as the activation function. The model's depiction. To avoid overfitting, our CNN model takes into account the dropout layer. Softmax is utilized as the activation function in the output layer to divide the final result into various diseases.

Table 1. List of hyperparameters

Function	Values
Epoch	30, 40 & 50
Batch size	32
Filter sizes for convolution layer	3×3
Activation function	ReLU
Loss function	sparse Categorical cross-entropy
Optimizer	adam

5.1 Convolutional neural network (CNN)

Here, we used a convolutional neural network (CNN)-based approach, a type of deep learning (DL) technique that takes an image as input and prioritizes numerous other items in the image while also distinguishing between them.

Contrary to other classification algorithms, a CNN requires significantly less pre-processing than they do. CNN can learn these filters and properties with enough training, whereas simple techniques necessitate hand-engineering of filters [23].

Our architecture mainly contains the following layers:

- Input layer
- Convolution layer
- Pooling layer
- Fully connected layer
- Output layer

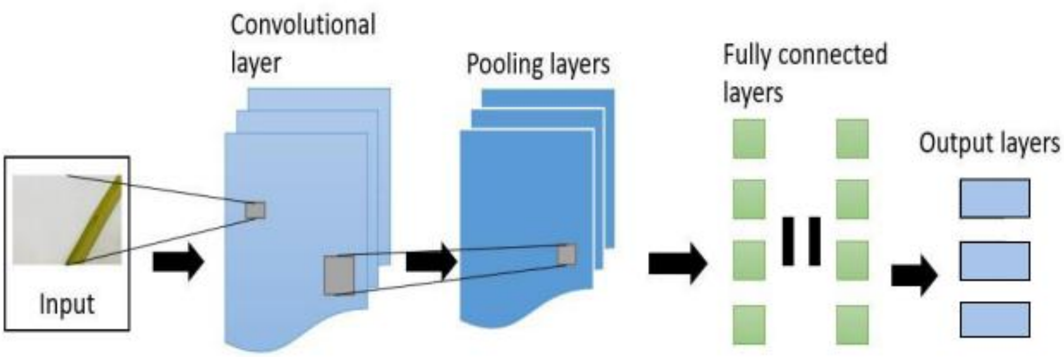

Fig. 5. Layers of CNN

The operational architecture of CNN is shown in the above diagram. After preprocessing the data and extracting the necessary features, the input in the form of an image is submitted to CNN, which processes it via three layers of CNN to accurately represent it. The final result is then shown [24].

● **Input Layer:** The input layer of CNN consists of the dataset. The input data will be represented as a 3X3 matrix.

● **Convolution Layer:** A layer that uses filters to learn from smaller sections of input data to obtain features from an image.

● **Pooling Layer:** This layer is used to shrink the image's dimensionality, lowering the processing power required for subsequent layers. There are two variations of pooling. They are:

● **Max pooling:** The pixel with the maximum value as input is selected and transferred to the output while parsing input. It is the most used approach compared to average pooling.

● **Fully Connected Layer (Dense):** This is one of CNN's last layers, and it can recognize features that are significantly linked with the output class. The result is a one-dimensional vector created by flattening the pooling layer results.

● **Dropout Layer:** Used to reduce model overfitting problem by removing a random set of neurons in that layer. It is connected with the FC layer.

● **Output Layer:** The output layer holds the final classification result.

Researchers needs to careful regarding the configuration of machine. In this study we have used intel core i5 processor, 8 GB Ram and Nvidia Geforce GPU. The used operating system was windows 10. It is highly recommended to use the GPU to train the model otherwise it will take long time. Training the model using supervised learning on the dataset allows for the analysis of the CNN's performance. Data annotations are used as references throughout the training process in supervised learning.

Table 2. Basic hardware & software

Hardware/Software Characteristic	Hardware/Software Characteristic
Processor (CPU)	Intel core-i5 (8th Gen)
RAM	8GB DDR4
Operating System	Windows 10
Graphics (GPU)	NVIDIA GeForce MX230
Environment	Tensorflow
Programming Language	Python

6 Results and Discussion

The dataset contains 10,000 images belonging to three classes of potato leaves. The results of training and validation accuracy and loss for epochs 30, 40 & 50 are given below. From the below 6,7 & 8, we identify the relationship between the number of epochs and learning outcomes.

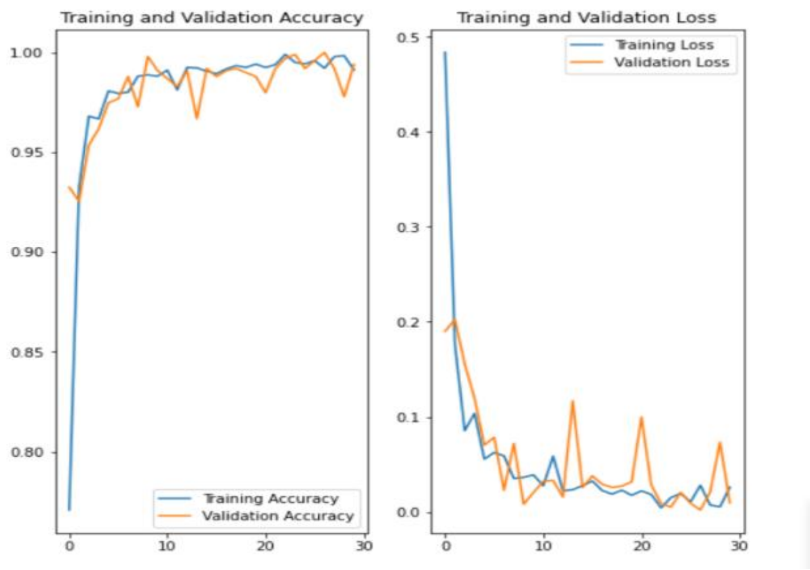

Fig. 6. Training, validation, and loss for 30 epochs

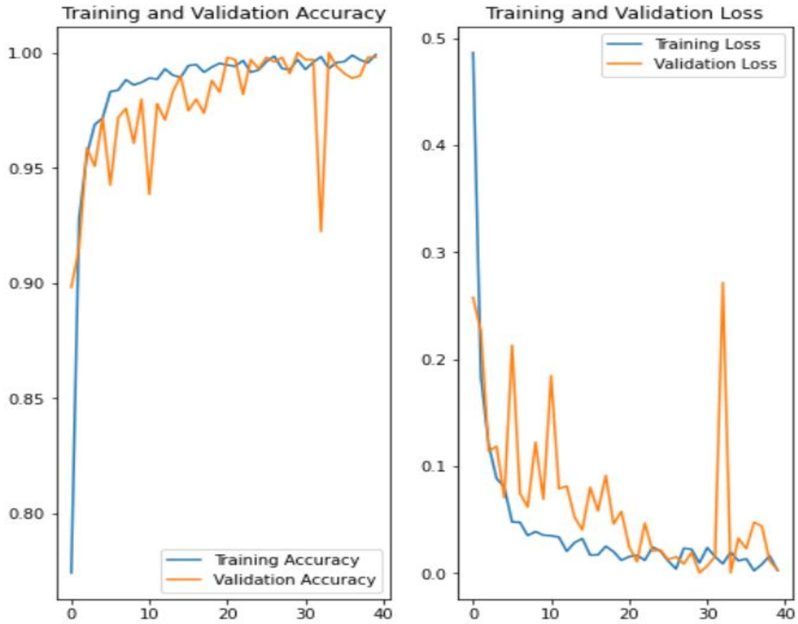

Fig. 7. Training, validation, and loss for 40 epochs

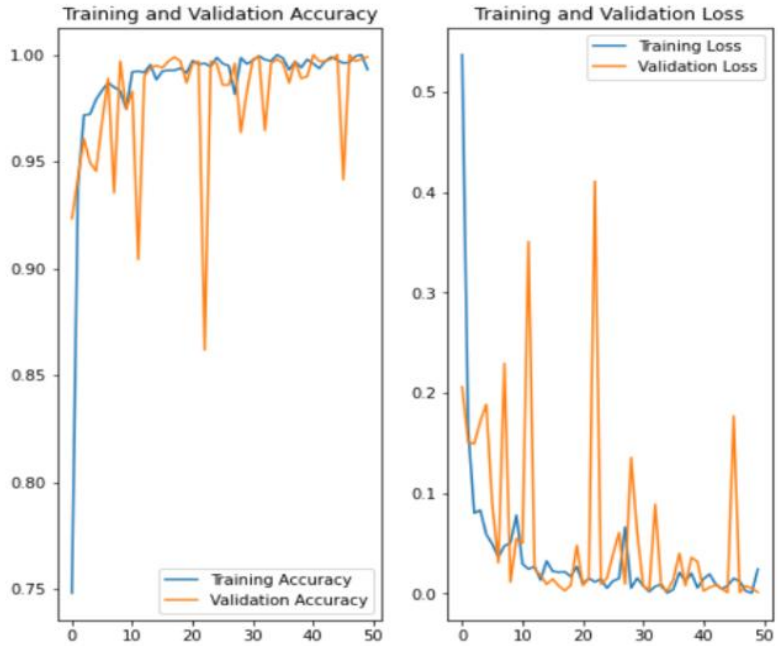

Fig. 8. Training, validation, and loss for 50 epochs

According to the figure, we can see that our model performed better when we applied 40 epochs.

We used a total of 20 images randomly to evaluate the proposed tool and investigate the classification accuracy. Table 3 is shown the results of potato disease classification for three different classes. The dataset has been split into three parts; training dataset, test data set and validation dataset. Neither the testing nor the validation data set were included in the training data set. The average model training time is more than 12 hours and testing time is more or less 3-5 hours.

Table 3. Validation results for classification

Epoch	Class	Quantity	Accuracy	Average Accuracy
30	Early Blight	7	99.98%	99.97%
	Healthy	8	99.97%	
	Late Blight	5	99.97%	
40	Early Blight	9	100%	100%
	Healthy	5	100%	
	Late Blight	6	100%	
50	Early Blight	5	100%	99.98%
	Healthy	7	100%	
	Late Blight	8	99.93%	

In Fig. 9 are shown the actual classes and predicted classes, including the confidence. 100% confidence means the accuracy is 100% of the predicted leaf. For every class, the accuracy of our model is 100%, which indicates that the model is working fine. The classification results show that the proposed model has good accuracy for 40 epochs.

Fig. 9. Training, validation, and loss for 40 epochs

Fig. 10 shows the code of actual classes and predicted classes, including the confidence for 40 epochs

```
In [63]: plt.figure(figsize=(20, 20))
         for images, labels in test_ds.take(1):
             for i in range(20):
                 ax = plt.subplot(4, 5, i + 1)
                 plt.imshow(images[i].numpy().astype("uint8"))

                 predicted_class, confidence = predict(model, images[i].numpy())
                 actual_class = class_names[labels[i]]

                 plt.title(f"Actual: {actual_class},\n Predicted: {predicted_class}.\n Cor

                 plt.axis("off")
```

Fig. 10. Code for actual classes and predicted classes, including the confidence for 40 epochs

7 Conclusion and Future Prospect

In this study, we proposed a convolutional neural network-based model for classifying potato leaf diseases. We assessed the performance's accuracy and loss while considering potato leaf diseases. Our suggested classification model can identify specific potato leaf diseases according to the evaluation results. The proposed model was observed to outperform with 99.91% accuracy during training, 99.80% accuracy during validation, and 100% accuracy in the testing phase. This model will help to classify the specific potato leaf disease and can take action according to that. In order to prevent the huge potato production loss every year in Bangladesh, we will continue our research.

Future research will analyze to launch of a web-based or android app for the benefit of potato farmers. We think our effort will have a wide range of positive effects on agricultural and global food security.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] UNICE, Food Loss And Waste; 2017.

[2] Editor; Mahfuz Anam, The Daily Star.
Available:https://www.thedailystar.net/business/economy/news/potato-farmers-lose-tk-2500cr-year-fao-2987086), access date 28/22/2022

[3] Drewnowski A, Rehm CD. Vegetable cost metrics show that potatoes and beans provide most nutrients per penny. PLoS One. 2013;8(5):e63277.

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

[5] TM P, Alla P, Ashirta KS, Chittaragi NB, Koolagudi SG. Tomato leaf disease detection using convolutional neural network. International Conference on Contemporary Computing (IC3); 2018.

[6] Hokkaido Agricultural Research Center, N.: HARC brochure; 2017.
Available:http://www.naro.affrc.go.jp/publicity report/publication/files/2017NARO english 1.pdf

[7] Faria FA, dos Santos JA, Rocha A, da S. Torres, R. Automatic classifier fusion for produce recognition. In: Proceedings of the 25th International Conference on Graphics, Patterns and Images (SIBGRAPI-2012). 2012;252–259.

[8] Geetharamani G, Arun Pandian J. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput. Electr. Eng. 2019;76:323–338.

[9] B. Mitschang et al. (Hrsg.): BTW– Workshopband, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 79 A Deep Learning-based Approach for Banana Leaf Diseases Classification Jihen Amara,1 Bassem Bouaziz, 2 and Alsayed Algergawy; 2017.

[10] Shen Weizhong, Wu Yachun [at.el]. Grading method of leaf spot disease based on image processing. IEEE. 2008;491-494.

[11] Appasaheb Gargade SA. Khandekar. A review: Custard apple leaf parameter analysis and leaf disease detection using digital image processing. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC). 2019;267-271.

[12] Srichocksittikul P, Nimsuk N. Design of Deep Learning Architecture for Classification of Orchid Diseases. 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). 2021;904-907.
DOI: 10.1109/ECTI-CON51831.2021.9454869

[13] Rabbia Mahum, Haris Munir, Zaib-Un-Nisa Mughal, Muhammad Awais, Falak Sher Khan, Muhammad Saqlain, Saipunidzam Mahamad & Iskander 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

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

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

[16] Hinton GE, Osindero S, The YW. A fast learning algorithm for deep belief nets. Neural Comput. 2006;18(7):1527–1554.

[17] Le QV. Building high-level features using large scale unsupervised learning. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. 2013;(ICASSP-2013):8595–8598.

[18] Kang L, Kumar J, Ye P, Li Y, Doermann D. Convolutional neural networks for document image classification. In: Proceedings of 22nd International Conference on Pattern Recognition 2014;(ICPR-2014):3168–3172.

[19] Lee H, Grosse R, Ranganath R, Ng AY. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: Proceedings of the 26th Annual International Conference on Machine Learning. 2009;609–616.

[20] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436.

[21] Potato By The Editors of Encyclopaedia Britannica
Available: https://www.britannica.com/plant/potato), access date 28/22/2022

[22] Ruedeeniraman N, Ikeda M, Barolli L. Performance evaluation of vegecare tool for potato disease classification. In: Barolli, L., Li, K., Enokido, T., Takizawa, M. (eds) Advances in Networked-Based

Information Systems. NBiS 2020. Advances in Intelligent Systems and Computing, Springer, Cham. 2021; 1264.
Available:https://doi.org/10.1007/978-3-030-57811-4_47

[23] Shrivastava VK, Pradhan MK. Rice plant disease classification using color features: a machine learning paradigm. Journal of Plant Pathology. 2021;103(1):17- 26.

[24] Chen Y, Jiang H, Li C, Jia X, Ghamisi P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing. 2016;54(10):6232-51.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/95283