*Original Research Article*

# A Deep Learning Approach to Classify the Potato Leaf Disease




## Abstract

### Purpose

The purpose of this paper is 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 take further decisions in order to prevent the loss of potato production.

### Design/Methodology/Approach

The paper implements a deep learning approach which is 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 filed. We collected a dataset of 2152 images from Kaggle and the rest of 7848 images from the above-mentioned 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 theoutput.

### Findings

This study found that the model performed better when we applied 40 epochs for the 10000 images dataset & we achieved 100% of accuracy. As we have applied a total of 3 different epochs and achieved the accuracy of 99.97% and 99.98% for 30 and 50 epochs respectively.

### Research Limitations/Implications

The study significantly contributed to the agriculture sector and farmers by providing suggestions to classify the Potato leaf Disease with the best output. Besides, Researches need more raw data to build the model for better output, and also should be concerned regarding the system when working with large volumes of data as it took long time to run the code.

### Originality/Value

This research paper contained high volume of dataset which is 10000 images of potato leaves. We collected a dataset of 2152 images from Kaggle and the rest of 7848 images from different sources like, google, and raw data from potato filed. We showed different epochs to check the best performance and we achieved 100% of 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 wide diversity in types and high consumer consumption, potatoes are a good enterprise option for many growers. Around 130 countries, 95 of which are developing nations, grow potatoes, the fourth most important staple food in the world. The output of potatoes worldwide has been steadily increasing over the past years, 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 the potato leaf disease.

Potato farming dominates as an occupation within the agriculture domain in ~~additional~~morethan 125 countries. However, even these crops are~~,~~ subjected to infections and diseases, mostly categorized into two grades: (i) Early blight and (ii) blight. Moreover, these diseases cause damages to 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 typical disease occurring on the foliage at any stage of the expansion and causes 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 theproblem is warm, rainy, and wet weather [4].

The identification of diseases in potato plants quickly and accurately is extremely essential 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. Farmers use the results of personal identification without expert advice within the sector of plant diseases as a reference for the prevention of 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.

In this paper, we present a Convolutional Neural Network based approach to identify and classify two common potato infections. Using the proposed method farmers can easily detect the disease in potato crops with little computational effort.

# 2. LITERATURE REVIEW

Recent years have seen the introduction of new equipment for rice farming 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 the realm of 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 [5]. 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 main 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. [1]

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 [16].

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 [17].

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[18].

# 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 [10,11].

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 [12,13]

# 4. PROPOSED 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.

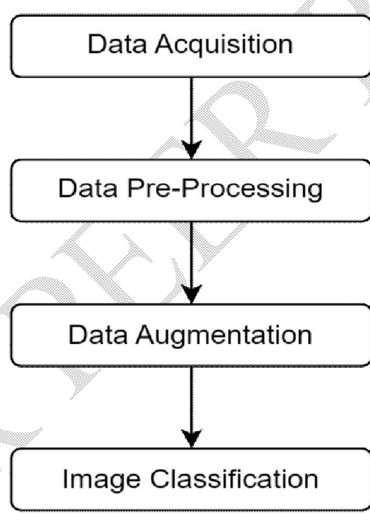


Fig. 1. Diagram Block Proposed Methodology.



## *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 has collected 7848 images directly from the field and 2152 images collected 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. Small black dots grow into huge, 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 are higher than 26 C, this disease will start to spread. It frequently appears when potatoes' activity is decreased as a result of 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.

Fig. 4. Healthy potato leaf

## *4.2 Data Pre-Processing*

Pre-processing of Data by removing the portion of the image that is not the region of interest, noise in the image can first be reduced. The image won't 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, vertical and horizontal flips, is employed to generate the augmentation parameters in this study.

## *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 [13]. 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 major diseases of potato leaves for categorization. The collection includes 10,000 photos of 3 classes that appear remarkably similar yet represent distinct diseases.

Table 1 displays a list of CNN hyper parameters. 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 hyper parameters.

| 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 |

Figure 5, 6 & 7 is showing the code for 30, 40 & 50 epochs respectively.

```
In [26]: history = model.fit(
             train_ds,
             batch_size=BATCH_SIZE,
             validation_data=val_ds,
             verbose=1,
             epochs=30,
         )
         racy: 0.9950 - val_loss: 0.0049 - val_accuracy: 0.9990
         Epoch 25/30
         250/250 [==============================] - 483s 2s/step - loss: 0.0193 - accu
         racy: 0.9942 - val_loss: 0.0204 - val_accuracy: 0.9919
         Epoch 26/30
         250/250 [==============================] - 436s 2s/step - loss: 0.0104 - accu
         racy: 0.9959 - val_loss: 0.0091 - val_accuracy: 0.9960
         Epoch 27/30
         250/250 [==============================] - 436s 2s/step - loss: 0.0280 - accu
         racy: 0.9921 - val_loss: 0.0019 - val_accuracy: 1.0000
         Epoch 28/30
         250/250 [==============================] - 435s 2s/step - loss: 0.0070 - accu
         racy: 0.9979 - val_loss: 0.0209 - val_accuracy: 0.9919
         Epoch 29/30
         250/250 [==============================] - 500s 2s/step - loss: 0.0052 - accu
         racy: 0.9984 - val_loss: 0.0733 - val_accuracy: 0.9778
         Epoch 30/30
         250/250 [==============================] - 474s 2s/step - loss: 0.0259 - accu
         racy: 0.9912 - val_loss: 0.0095 - val_accuracy: 0.9940

In [27]: scores = model.evaluate(test_ds)

         32/32 [==============================] - 42s 469ms/step - loss: 0.0084 - accura
         cy: 0.9951
```

Fig. 5. Code for 30 epochs

```
In [27]: history = model.fit(
             train_ds,
             batch_size=BATCH_SIZE,
             validation_data=val_ds,
             verbose=1,
             epochs=40,
         )

racy: 0.9931 - val_loss: 4.6254e-04 - val_accuracy: 1.0000
Epoch 35/40
250/250 [==============================] - 512s 2s/step - loss: 0.0118 - accu
racy: 0.9957 - val_loss: 0.0329 - val_accuracy: 0.9940
Epoch 36/40
250/250 [==============================] - 540s 2s/step - loss: 0.0137 - accu
racy: 0.9961 - val_loss: 0.0230 - val_accuracy: 0.9909
Epoch 37/40
250/250 [==============================] - 509s 2s/step - loss: 0.0025 - accu
racy: 0.9989 - val_loss: 0.0477 - val_accuracy: 0.9889
Epoch 38/40
250/250 [==============================] - 501s 2s/step - loss: 0.0085 - accu
racy: 0.9969 - val_loss: 0.0441 - val_accuracy: 0.9899
Epoch 39/40
250/250 [==============================] - 509s 2s/step - loss: 0.0165 - accu
racy: 0.9956 - val_loss: 0.0117 - val_accuracy: 0.9980
Epoch 40/40
250/250 [==============================] - 496s 2s/step - loss: 0.0028 - accu
racy: 0.9991 - val_loss: 0.0033 - val_accuracy: 0.9980

In [31]: scores = model.evaluate(test_ds)

32/32 [==============================] - 14s 438ms/step - loss: 7.3153e-05 - ac
curacy: 1.0000
```

Fig. 6. Code for 40 epochs

```
In [26]: history = model.fit(
             train_ds,
             batch_size=BATCH_SIZE,
             validation_data=val_ds,
             verbose=1,
             epochs=50,
         )

uracy: 0.9990 - val_loss: 0.0050 - val_accuracy: 0.9980
Epoch 45/50
250/250 [==============================] - 581s 2s/step - loss: 0.0077 - accu
racy: 0.9975 - val_loss: 8.3544e-04 - val_accuracy: 1.0000
Epoch 46/50
250/250 [==============================] - 567s 2s/step - loss: 0.0151 - accu
racy: 0.9961 - val_loss: 0.1770 - val_accuracy: 0.9415
Epoch 47/50
250/250 [==============================] - 586s 2s/step - loss: 0.0122 - accu
racy: 0.9965 - val_loss: 0.0011 - val_accuracy: 1.0000
Epoch 48/50
250/250 [==============================] - 586s 2s/step - loss: 0.0024 - accu
racy: 0.9994 - val_loss: 0.0076 - val_accuracy: 0.9970
Epoch 49/50
250/250 [==============================] - 587s 2s/step - loss: 4.0667e-04 -
accuracy: 1.0000 - val_loss: 0.0050 - val_accuracy: 0.9980
Epoch 50/50
250/250 [==============================] - 585s 2s/step - loss: 0.0242 - accu
racy: 0.9931 - val_loss: 0.0013 - val_accuracy: 0.9990

In [27]: scores = model.evaluate(test_ds)

32/32 [==============================] - 65s 689ms/step - loss: 0.0032 - accura
cy: 0.9971
```

Fig. 7. Code for 40 epochs

# 6. RESULT & DISCUSSIONS

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 figures 8, 9 &10, we identify the relationship between the number of epochs and learning outcomes.

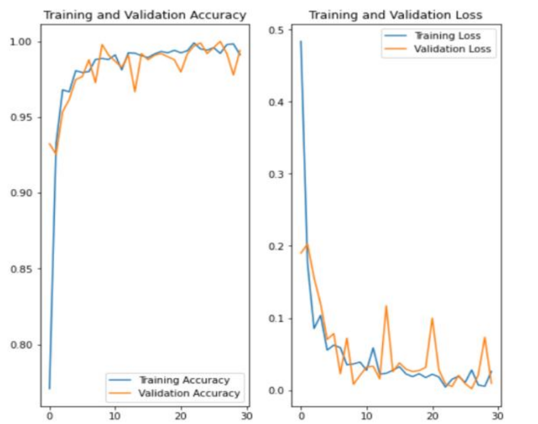


Fig. 8. Training, validation, and loss for 30 epochs

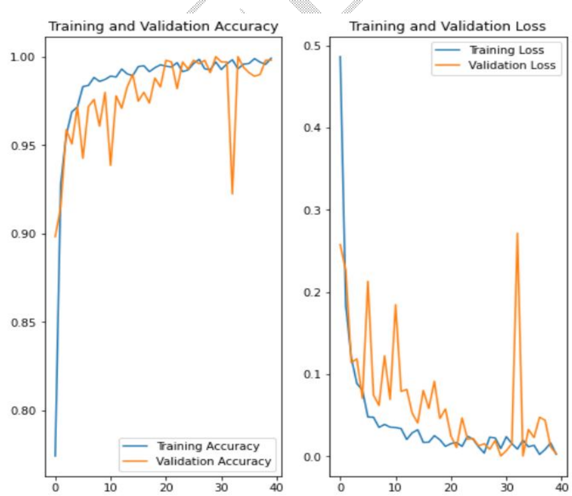


Fig. 9. Training, validation, and loss for 40 epochs

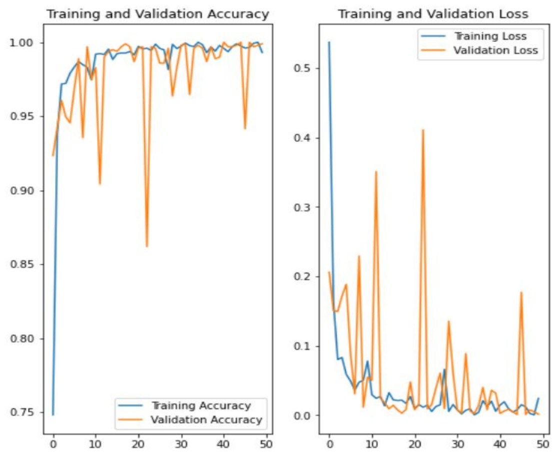


Fig. 10. Training, validation, and loss for 50 epochs

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 investigated the classification accuracy. Table 2 are shown the results of potato disease classification for three different classes.

**Table 2.** Testing 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% | |
| | Early Blight | 5 | 100% | |

| | | | | |
|---|---|---|---|---|
| 50 | Healthy | 7 | 100% | 99.98% |
| | Late Blight | 8 | 99.93% | |

In Fig 11. 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. 11. Training, validation, and loss for 40 epochs

Fig. 12 is showing 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. 12. actual classes, and predicted classes including the confidence for 40 epochs

# 7. CONCLUSION AND FUTURE WORK

In this study, we proposed a convolutional neural network-based model for classifying potato leaf diseases. We assessed the accuracy and loss of the performance while taking potato leaf diseases into account. Our suggested classification model can identify typical potato leaf diseases, according to the evaluation results. It was observed that the proposed model outperforms 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 app or android app for the benefit of potato farmers. We think that our effort will have a wide range of positive effects on agricultural and global food security.

.

# 8. References

1. UNICE, Food Loss And Waste, Nov2017.
2. Editor; MahfuzAnam, The Daily Star (https://www.thedailystar.net/business/economy/news/potato-farmers-lose-tk-2500cr-year-fao-2987086), access date 28/22/2022
3. Drewnowski, A., Rehm, C.D.: Vegetable cost metrics show that potatoes and beans provide most nutrients per penny. PLoS ONE 8(5), e63277 (2013)
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. 34, 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," International Conference on Contemporary Computing (IC3), 2018.
5. Hokkaido Agricultural Research Center, N.: HARC brochure. http://www.naro.affrc.go.jp/publicity report/publication/files/2017NARO english 1.pdf (2017)
7. Faria, F.A., dos Santos, J.A., 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), pp. 252–259 (2012)
8. Geetharamani, G., Arun Pandian, J.: Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput. Electr. Eng. 76, 323–338 (2019)
9. B. Mitschang et al. (Hrsg.): BTW 2017 – 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 AlsayedAlgergawy.
10. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. Neural Comput. 18(7), 1527–1554 (2006)
11. Le, Q.V.: Building high-level features using large scale unsupervised learning. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 2013 (ICASSP-2013), pp. 8595–8598, May 2013
12. 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), pp. 3168–3172, August 2014
13. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y.: Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 609–616, June 2009
14. Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," Nature, vol. 521 no. 7553, p. 436, 2015.
15. Potato By The Editors of Encyclopaedia Britannica (https://www.britannica.com/plant/potato), access date 28/22/2022
16. Shen Weizhong, Wu Yachun [at.el], "Grading method of leaf spot disease based on image processing,'' IEEE, 2008, pp.491-494.
17. AppasahebGargade, S.A. 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), pp.267-271, 2019.
18. P. Srichocksittikul and N. Nimsuk, "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, pp. 904-907, doi: 10.1109/ECTI-CON51831.2021.9454869.