International Journal of Agronomy and Agricultural Research (IJAAR)
ISSN: 2223-7054 (Print) 2225-3610 (Online)
http://www.innspub.net
Vol. 24, No. 6, p. 15-19, 2024

RESEARCH PAPER OPEN ACCESS

Potato leaf disease detection using image processing

Rokon Uz Zaman[*1], KU Ahmmad[2], Debasis Sarkar[3], MAA Mumin[4], N Salahin[5]

[1]*Agricultural Engineering Division, Regional Agricultural Research Station, Jashore, Bangladesh*

[2]*Regional Agricultural Research Station, Jashore, Bangladesh*

[3]*Plant Pathology Division, Pulse Research Center, Ishurdi, Pabna, Bangladesh*

[4]*Entomology Division, Regional Agricultural Research Station, Jashore, Bangladesh*

[5]*Soil Science Division, Regional Agricultural Research Station, Jashore, Bangladesh*

Article published on June 05, 2024

Key words: Potato disease detection, Image processing, CNN, Smart farming, One hot encoding

Abstract

Agriculture is one of the most important pillars of Bangladesh's economy. However, due to some factors such as plant diseases, pests, climate change, the yield of the farming industry decreases, and the productivity decreases as well. The detection of plant diseases is crucial to avert the losses in the productivity and in the yield. It is not obvious to monitor the plant diseases manually as the act of disease detection is very critical. It needs a huge effort, along with knowledge of plant diseases and extensive processing times. Therefore, image processing technology is used to detect the plant disease, this is done by capturing the input image that undergoes the process and is compared with the dataset. This dataset is composed of diverse diseases of potato leaves in the image format. This study aims to build a web application to predict the diseases of potato plants that will help farmers to identify the diseases so that they can use appropriate fungicide to get more yields. The purpose of this study is to assist and provide efficient support to the potato farmers. In this study, we propose a system that will use the techniques of image process to both analyze and detect the plant diseases using machine learning Conventional Neural Networks (CNN) with Tensorflow framework 2. The results of the implementation show that the designed system could give a successful result by detecting and classifying the potato leaf diseases and healthy plant.

* **Corresponding Author:** Rokon Uz Zaman ✉ rokonae99@gmail.com

Introduction

Agricultural industry is the backbone of our economy that contributes about 11.63% of GDP (BBS 2021). Potato is an important and leading crop in Bangladesh. Bangladesh is the seventh potato producing country in the world and ranks second after rice in terms of production and are the third most important food crop after rice and wheat in terms of human consumption in Bangladesh (FAOSTAT, 2020). According to the DAE statistics, about 9.61 million MT of potatoes have been produced in 2020 against the annual demand of about 6.82 million MT, bringing a surplus of 3.40 million MT despite some amounts is being exported (DAM, 2020).

But in Bangladesh late blight is the most common and highly destructive, fungal disease in potato and annual potato yield losses due to late blight have been estimated at 25-57% (GEOPOTATO project report, 2016-2019). Plants are sensitive to diseases especially the plant leaves as symptoms of the disease appear first on the leaves. Due to the bad impacts of plant diseases on the both the economy and environment, the farmers should consider monitoring the crops in such a way that they may mitigate losses. It exists a way that is used by experts to monitor the crops which is the naked eye observation. This is a traditional method that has many constraints related to time consuming as the operation of monitoring is done manually, and it requires the presence of experts. However, lately, crop monitoring is being developed to be digital and semi-automatic, meaning that only from the symptoms that are shown on the leaf, the disease could be detected in an easier, quicker, cheaper way. Therefore, this digitalized method will be beneficial for the farmers as well since it will facilitate for them the detection of the diseases because most of the farmers do not have a sufficient background and knowledge about monitoring the crops and dealing with the variety of diseases that could affect them. There are many researcher reported, leaf dieses classification and detect is successfully possible by using image processing techniques of deep learning as well as machine learning. Different methods for machine learning and deep learning include the Support Vector Machines (SVM), Random.

Forests (RF), K-nearest Neighbor (KNN), Artificial Neural Network (ANN), and Convolutional Neural Network (CNN), along with models such as AlexNet, GoogleNet, and Caffe are used to classify and detect to leaf dieses (Knaak *et al.*, 2021). The report presented a machine-learning model including canny edge detection technique for edge feature extraction, grid color movement for extracting color features and local binary pattern (LBP) for texture analysis. Where the features were extracted combined to create a combined feature vector which was used for training the artificial neural network (ANN). The convolutional model is also capable of differentiating the plant leaves and recognizing rice plants and their diseases (Shrivastava *et al.*, 2022). Potato leaf diseases were detected by using random forest classifiers where image pre-processing was done in two steps like image normalization and color space conversion where segmentation was done using thresholding HSV images in RGB color space and global feature descriptor (GFD), gray level co-occurrence matrix (GLCM), color histogram were used for extracting features. Finally, classification was done using random forest (RF) classifiers (Iqbal *et al.*, 2020). The proposed system that we are suggesting in this paper could be used by the farmers to increase the yield with no need to consult experts. The core purpose of this proposed system is not aiming only at detecting the plant diseases using the image processing technology, but it aims also at directing the user farmer to use a mobile application in which he will upload the image and receive the type of disease infection along with a suggestion of needed pesticides. The digitalization of the agriculture field has known the intervention of the latest technologies namely the image processing. As a result, our system that is designed to be automated system is implemented using image processing technique using machine learning Convolutional Neural Networks (CNN) with Tensorflow Framework 2.

Materials and methods

Proposed model

The proposed methodology for prepare model consists of various steps like image acquisition, data augmentation, Gaussian smoothing, image normalization, dimensionality reduction, one hot encoding, and classification by using CNN with Tensorflow Framework 2 (Fig. 1). CNN is a deep learning algorithm specifically designed for any task where object recognition is crucial such as image classification, detection, and segmentation. Tensorflow Framework is open source mainly deal with high dimensional data when building machine learning and deep learning models. Tensors are multidimensional arrays with a uniform type used to represent different features of the data. Finally the deployment of web application is completed by using HTML, CSS and bootstrap where flask framework are used for back-end.

Description of dataset

The images like dieses and healthy leaf of potato are collected from the different potato field of Bangladesh. The dataset consists of both healthy and disease images of about 2500 where about 2350 images are disease affected leaves and 150 are healthy leaves (Fig. 2). There was not good amount of image in data set. The potato leaf images were categorized into three types like early blight, late blight, and healthy state. The data set was scaled down to have a constant size of 500 by 500 pixels. Training and testing data set were created from the collecting data set where 80% are used for training and remaining 20% are used for validation. Training was done carefully to ensure the reliability. Python was used to implement the framework.

Image pre-processing

We have used CNN classifiers for image preprocessing. Image preprocessing was performed by data augmentation, Gaussian smoothing, image normalization and dimensionality reduction.

Data augmentation

Data augmentation refers to the process of minor modifications to current data in order to increase its diversity without collecting new data that help to increase sample size. Standard data augmentation techniques include vertical and horizontal flipping, cropping, rotation and shearing the data. Data augmentation supported to preventing neural networks from acquiring irrelevant properties (Fig. 3). In this study we were applying data augmentation techniques for the potato healthy leaves class only and the operations were done by applying flipping horizontally, flipping vertically, and slight rotation clockwise, slight rotation anti-clockwise. After the data augmentation healthy potato leaves were rises to 800 and total images would be 3150.

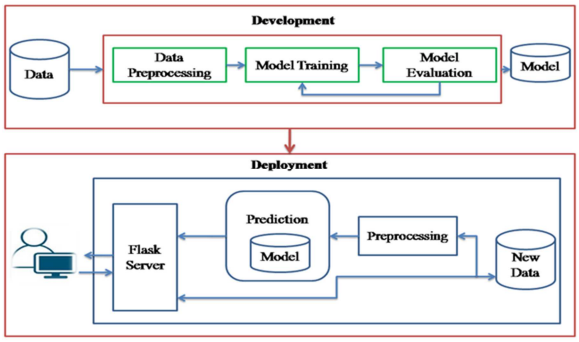


Fig. 1. Overview of the proposed methodology

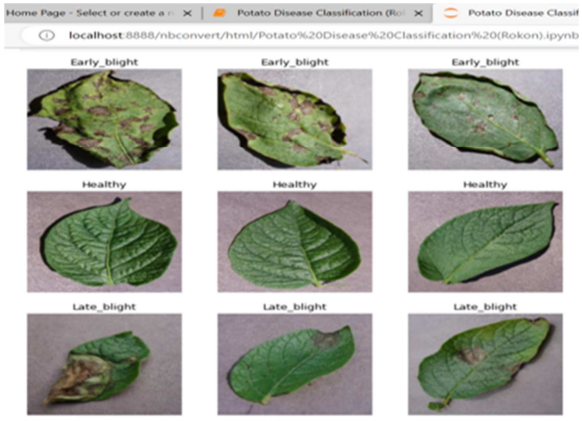


Fig. 2. Sample images from the dataset during operation

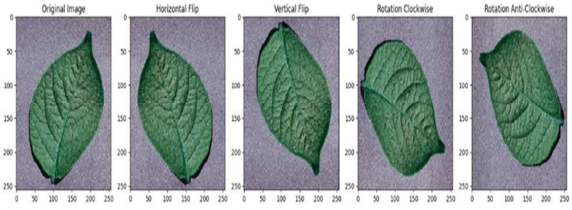


Fig. 3. Augmented image

Gaussian smoothing

The collected images having very sharp and noisy. For the better and accurate result, we need more smooth and sharpness images. We used Gaussian smoothing to reduce the sharpness, noise and blur the images.

Image normalization

Data rescaling defined, the projection pixel value of image (intensity) to a predetermined range (0, 1). This is a common preprocessing step for images to ensure scale invariance and faster convergence during the training. We need to convert images into a numerical format so that they can be easily processed by the neural network. We must standardize data so that we may use the same procedures to them. Generally, normalization is used to convert the pixels of the image into a more standard or pleasant appearance. The advantages of the image normalization are to ensure the fairness all photos like, when all the images are scaled to the same range then photos may contribute equally to the total loss. On the other hand, if these are not and photos having high and low resolution ranges that contribute disproportionately to the loss. Rescaling of images help to provide a consistent learning rate for all photographs. Generally, high-pixel pictures required a smaller learning rates and low pixel picture need a high learning rate.

Model architecture

We used CNN classifier with Tensorflow framework 2 in this study. The final step of the pre-processing of this model is one hot encoding. We are using Softmax layer at the end to output the probabilities of three classes. First, we define the model using the Sequential () class, and each layer is added to the model with add () function. We get the comprehensive summary of the model's architecture with information about each layer, its type, output shape and the total number of trainable parameters by using summary () function.

Results and discussion

Model evaluation

We have everything to train our model by using Adam algorithm for optimization and early stopping callback to avoid over fitting. The optimizer is responsible for updating the model's weights and biases. The loss function is used to measure the misclassification errors and the metrics is used to measure the performance of the model, and accuracy, precision, and recall will be displayed in our use case.

We will train our model used a total of 20 epochs for training and store the training history in the history variable. After 20 epochs, the result found 0.99 training accuracy and 0.97 validation accuracy that is good performance and no over fitting (Fig. 4).

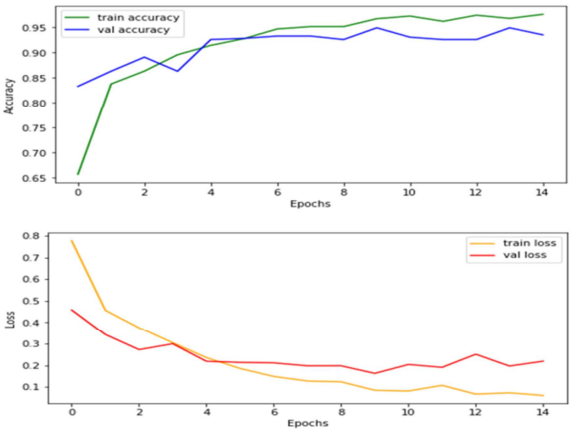
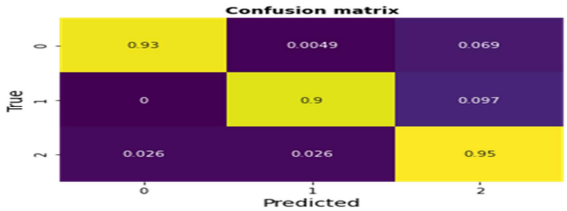


Fig. 4. Model evaluation



0=Early blight, 1=healthy, 2=Late blight

Fig. 5. Confusion matrix

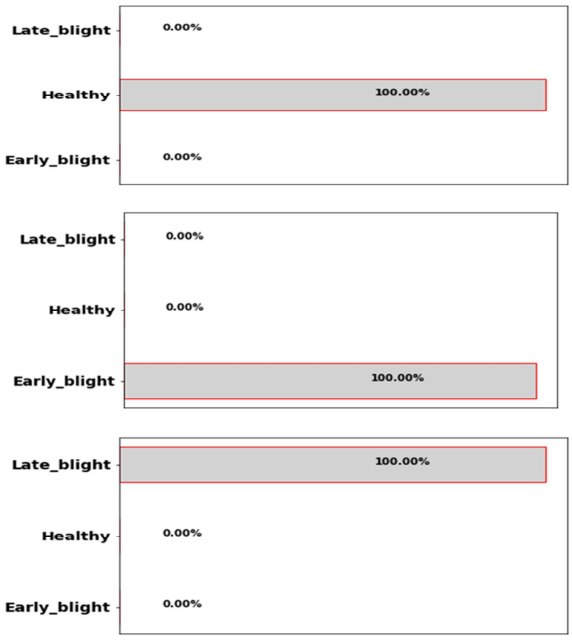


Fig. 6. Prediction on single image healthy or not by using model

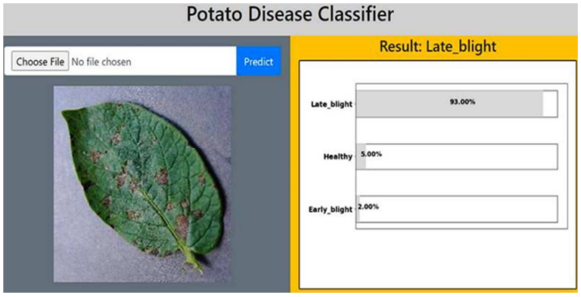


Fig. 7. Prediction on single Image healthy or not by using web application

Confusion matrix

We can determine the accuracy of the model by observing the diagonal values for measuring the number of accurate classification by visualizing the confusion matrix. The model is performing better on all classes. Only 10% of healthy leaves are classified as late blight. If we added more images for healthy class, we can overcome this. The proposed model gave the best result in terms of precision, recall, f1 score, and accuracy (Fig. 5).

Tested the model

The model build up is completed successfully and save the model, and tested to detect potato leaves. The model can predict successfully potato leaves are healthy or not (Fig. 6).

Deployment of web application

We have successfully trained our model and saved it for future use. Besides, we build our web application where HTML, CSS and bootstrap are used for front end part and used flask framework for back-end part. This web application can to predict the disease in an infected potato plant using its leaf (Fig. 7).

Conclusion

In this study, leaf diseases of potato plants and healthy leaves are detected and classified using machine learning classifiers with the use of image processing. However, developed web application can predict to the disease in an infected potato plant using its leaf successfully.

References

BBS. 2021. Report of the Share of economic sectors in the GDP in Bangladesh.

DAM. 2020. Report of the Department of Agricultural Marketing, Khamarbari, Dhaka, Bangladesh.

FAOSTAT. 2020. New food balance sheet for Bangladesh. Food and Agriculture Organization of the United Nations.
http://www.fao.org/faostat/en/#data/FBS.

GEOPOTATO. 2016-2019. Report of the GEOPOTATO.
https://www.wur.nl/en/project/geopotato-control-fungal-disease-in-potato-in-bangladesh.htm

Iqbal MA, Talukder KH. 2020. Detection of Potato Disease Using Image Segmentation and Machine Learning. International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 43–47.
DOI: 10.1109/WiSPNET48689.2020.9198563

Knaak C, Von Eßen J, Kröger M, Schulze F, Abels P, Gillner A. 2021. A Spatio-Temporal Ensemble Deep Learning Architecture for Real-Time Defect Detection during Laser Welding on Low Power Embedded Computing Boards. Sensors **21**(12), 4205, DOI: 10.3390/s21124205.

Shrivastava G, Patidar H. 2022. Rice Plant Disease Identification Decision Support Model Using Machine Learning. Ictact Journal on Soft Computing **12**(3), 2619-2627.
DOI: 10.21917/ijsc.2022.0375