Rice Disease Detection Based on Image Processing Technique

Md. Asfaqur Rahman[1], Md. Shahriar Nawal Shoumik[1], Md. Mahbubur Rahman[1] and Hasna Hena[2]

[1]Daffodil International University, Dhaka, Bangladesh
{asfaqur15-7086, nawal15-7156, mahbub15-7012}@diu.edu.bd
[2]lecturer, Daffodil International University, Dhaka, Bangladesh
hena@diu.edu.bd

Abstract. Rice plant diseases are a major problem in Bangladesh. Detection and monitoring of these rice plant disease is a critical issue. Rice plants are affected in various kind of disease like leaf blast, Leaf blight, brown spot and show the syndrome in the leaf of these diseases. If these diseases are detected early and take appropriate action, it will restrain extensive economic loss for the farmer. In this research paper, the proposed model will successfully classify and find out the rice leaf diseases based on image processing techniques. Machine learning algorithm CNN used to implement this model. Healthy and disease affected leaves are taken for the proposed method and separated healthy and unhealthy characteristics of rice plant leaves. Then these images are being processed with the proposed model and classified the rice plant leaf as either infected in disease or healthy. This proposed model provided as accuracy of 90%. This model successfully identifies the infected and healthy rice plant.

Keyword: Rice disease detection, image processing, neural network.

1 Introduction

Bangladesh is an agricultural country. Rice is the main cereal crops in our country and plays an important role on our agricultural economy. About 75% cultivated land are being used for rice cultivation. Rice farming plays a vital role on agricultural economy, it contributes one half of the agricultural GDP. As the population is increasing, demand of rice is also increasing and it is a major challenge to producing more crops to encounter the increasing demand. Because for the past few years its production level decreases due to some unexpected disease. But in our agriculture, rice plants are often affected with serious disease like Hipsa, Brown spot, leaf blast, bacterial leaf blight etc. which causes serious damage of rice plant and also decrease the production of rice [9]. When rice plants are affected in disease, some syndrome are seen on their leaves. There emerges some spot on the leaf of the plant. Leaf Blight, brown spots are viewed in many shapes on the leaf. From these spots disease can be identified and can take an appropriate action. This paper proposed a machine learning method to identify these diseases from rice plant leaves. Huge number of images of paddy's leaves are needed for this research because the research is based on images. Hipsa, healthy, brown spot and leaf blast these four types of images are collected. Divided the total image into two for training and testing purpose. Almost 70% of images are in training folder and the rest will be in the testing folder. The proposed model can identify the healthy and infected rice leaves. That’s why, farmer can easily identify the infected leaves and will take appropriate action and increase their productivity.

2 Literature Review

John William ORILLO, et al [1] proposed a model in Plant disease identification using back propagation where at first all the image is converted to binary-level image and differentiate the spot and then extracting all the spot using back propagation. But the limitation of using back propagation is once a network teach one sets of weights, any new learning causes catastrophic forgetting.

John William Orillo, et al [2] proposed a model on rice plant disease identification and detection using Fuzzy Neural Network. They describe a method of identification of rice disease using sound signal processing system. The classification is done by MATLAB function and Fuzzy neural network. But the disadvantages of fuzzy network are if the model is nonlinear with a disturbance term the testing error is very large.

S. Arivazhagan, et al [3] developed a model on unhealthy region of plant leaves and classification using texture features. For the inputted RGB images color transformation structure is create. Main disadvantages of using support vector machines that creating a color transformation structure for input RGB images is very difficult task.

Dheeb al Bashis, et al. [4] have proposed a method for detection and Classification of plant leaf based on K-means algorithm. The classification is done by neural network classifier based on statistical classification. But the limitation of using k-means algorithm is that to predict the value of k and clusters of different size and different density does not work well.

Phadikar, et al [5] proposed a method for rice disease using software prototype system based on pattern recognition technique. Their prototype works based on the infected rice leaf’s images. Classification of the infected parts done by neural network. It’s only identified the infected part from an infected image but it couldn’t differentiate the leaf into healthy and infected.

3 Methodology

3.1 Working diagram

For detecting rice leaf's disease, there are couple of successful steps have to finish. Block diagram of the model are given below:

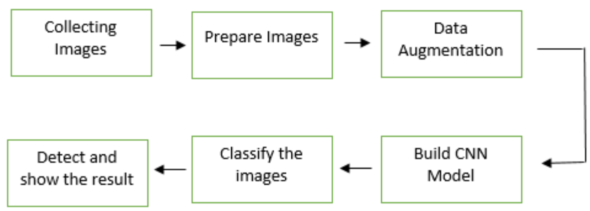


Fig 1. Block Diagram.

3.2 Dataset collection

Data is the main assets of this project. Images are collected from Kaggle, Dataquest and captured some images manually. Almost 3000 images are collected but all these weren't useful for us. Some pictures resolution is very poor, some are hardly classified as healthy and unhealthy. That's why, 300 images are selected for brown spot, healthy and leaf blight, hispa for the training purposes. Similarly, 110 images are selected per folder for the validation. The dataset which used for this research is not relatively clean. That's why we have to perform data processing to get ready our data for modeling.

3.3 Neural Network

CNN is both supervised and unsupervised learning architecture. They can used for predict something or classification. But mostly CNN used for supervised method. CNN classify images based on their attributes [12]. By using activation function CNN computes future maps. The function was:

$$y_j^l = f(z_j^l) \tag{1}$$

Where, y_j^l is the future map and $f(z_j^l)$ is the activation function. Datasets are stored in 2-dimentional convolution operation in convolutional neural network (CNN).

$$O = \frac{(W-F+2P)}{(S+1)} \quad (2)$$

Where O, W, K, P, S are expressed respectively output height/length, input height/length, filter size, padding, stride [15].

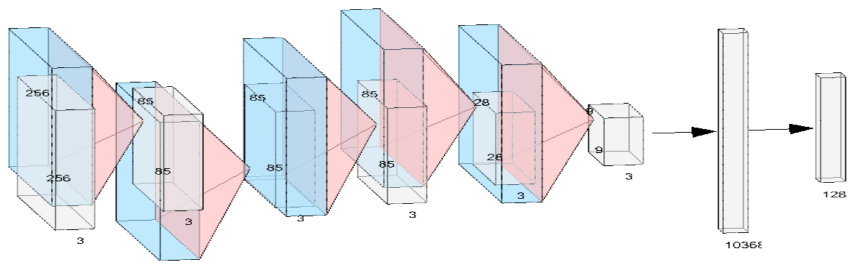


Fig 2. Convolutional Neural Network (CNN).

Fig 2. shows the basic structure of CNN of the model using the layer, padding, height, width, dense that used on the model [10].

3.4 Building the model

The sequential model is built up by using CNN algorithm [14]. Keras helps to build this model as layer by layer [13]. For adding layers, add () function used. convolutional2D layer used to deal with input images to seen as 2-dimensional matrices and dense layer used for output images. The kernel size was 3x3 as a filter matrix. ReLU or Rectified Linear Activation was used as an activation function for this model [6]. Images must be inputted in this model. The input size of the images is (256,256). That means height and weight of the images should be 256. Between the convolutional2D layer and dense layer there was a flatten layer which was serves as a connection between them. For prediction based on the highest probability SoftMax added as an activation in the model [16]. SoftMax function equation is given below:

$$P(y = j|x) = \frac{e^{X^T W^l}}{\sum_{k=1}^{k} e^{X^T W^l}} \quad (3)$$

Here, X^T W denotes the inner product of X and W.

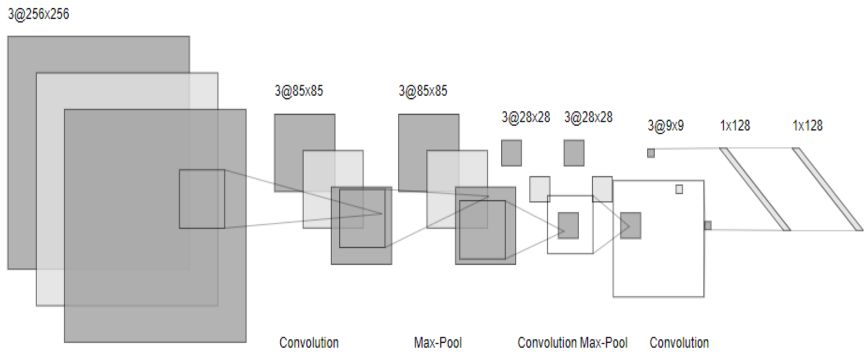


Fig 3. Pictorial view of the sequential model

Pictorial view of the sequential model is given in the Fig.3[10].

3.5 Compiling the Model

Optimizer, loss and metrics these parameters used for compiling the model. “adam” was used as an optimizer. It actually a good optimizer for adjusting the learning rate throughout training. ‘categorical cross entropy’ used for the loss function. To make things even easier to interpret, ‘accuracy’ metric used to see the accuracy score on the validation set during training the model.

3.6 Training and Testing the Model

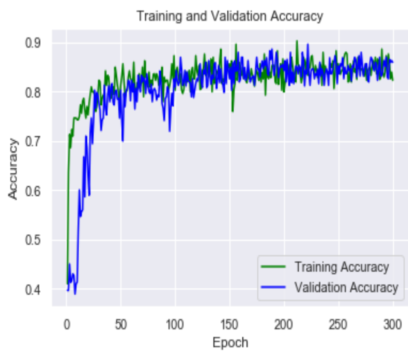
To train, ‘fit ()’ function used on the model. Testing dataset used as validation data. In the fit function we set the epochs number which will cycle the model through the data. After finishing the training process the testing process is processed. It verifies the effectiveness of the trained CNN.

4 Result and Outcome

4.1 Learning rate:

Learning rate shows a visual graph of adjusting the weights of our network with respect the loss gradient. Relationship between weight, gradient and learning given below:

$$New\ weight\ =\ existing\ weight - learning\ rate\ \times\ gradient \tag{5}$$



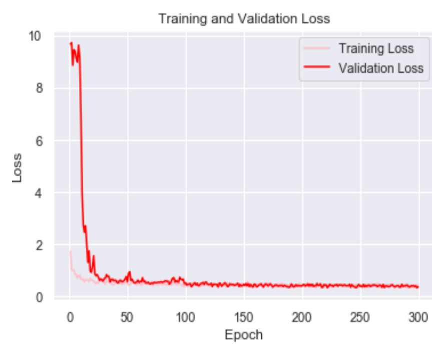


Fig 4. Accuracy graph

Fig 5. Loss graph

Fig 4. shows the accuracy function graph of our model while training datasets. At the beginning training accuracy is very low but after time being the accuracy is increased. Fig 5. shows the loss function graph. Initially the validation loss is very high but after time being validation loss reduced [7].

4.2 Confusion matrix

Confusion matrix visually shows the performance of an algorithm. It counts the values of correct and incorrect predictions [8]. Accuracy or classification rate formula and error of the model formula are given below:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{6}$$

$$Error = 1 - accuracy \tag{7}$$

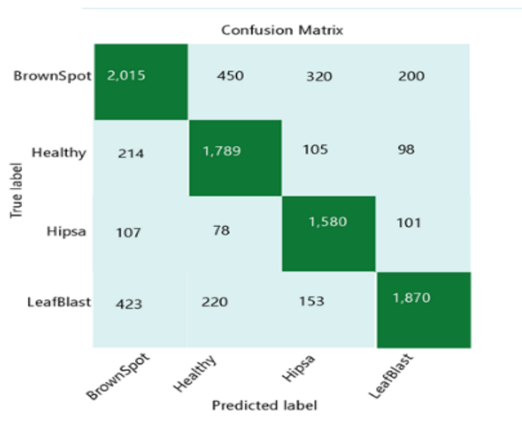


Fig 6. Confusion matrix

Fig 6. shows us the confusion matrix of our classification model. It shows that the diagonal value in each row is the highest value. So, our algorithm will give good accuracy.

4.3 Classification

In jupyter notebook the entire process is implemented. In the first phase different infected rice leaf image was inputted for classify those. The images are selected from the dataset folder. At the end expected classified outcome are founded.

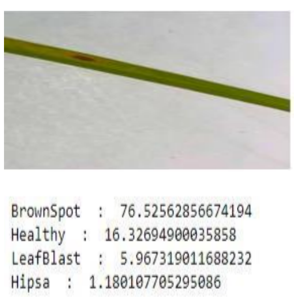


Fig. 7

Fig. 8

Fig. 9

Fig. 10

When brown spot images inputted then in the Fig. 7 shows that the brown spot disease percentage is over 76%. The model successfully detects brown spot disease image. Similarly, Fig. 8 shows that highest percentage is healthy that’s why it can predict that this image is a healthy leaf image. Fig. 9 predict the inputted leaf is affected in hispa because it gives highest percentage. In Fig. 10 predict the inputted leaf is affected in leaf blast because it gives highest percentage. So, our model successfully detects all the disease and healthy images.

5 CONCLUSIONS

For real time rice disease detection, we have proposed convolutional neural network (CNN) based on Classifier model named as sequential. For achieving desired accuracy two stage training have also introduced. comprehensive study on various kinds of rice disease are conducted. Lot of images of rice plants are collected from the rice field. Different types of algorithm based on the CNN architecture implemented for finding the best result. The model successfully been able to classify the disease affected plant and healthy plant. This model also help farmer to decision making while farming. It can give more accurate result when the dataset is larger. Before processing the datasets can valuable in the prediction and unprocessed data can also affect the efficiency of the model. Our validation accuracy and test accuracy are found to be very high, because our training, validation and test set are well classified. We plan to work with memory efficient non-sequential CNN models in order to achieve higher accuracy in plant disease classification in future.

Reference

1. John William ORILLO, Jennifer DELA CRUZ, Leobelle AGAPITO, Paul Jensen SATIMBRE Ira VALENZUELA “Identification of Diseases in Rice Plant (Oryza Sativa) using Back Propagation Artificial Neural Network” 2014 International Conference on Humanoid,

Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM).

2. John William Orillo, Timothy M. Amado, Nilo M. Arago, Edmon Fernandez. “Rice plant disease identification and detection technology through classification of microorganisms using fuzzy neural network” Jurnal Teknologi (Sciences & Engineering) 778: 5–8 (2016) 25–31.
3. S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini “Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features” Agricultural Engineering International : The CIGR e-journal 15(1):211-217 · January 2013.
4. Al-Bashish, D., M. Braik, and S. Bani-Ahmad. 2011. Detection and classification of leaf diseases using K-means-based segmentation and neural networks based classification. Information Technology Journal, 10(2): 267-275.
5. Phadikar, S. &Sil, J. (2008). Rice Disease Identification using Pattern Recognition Techniques. Paper presented at the 11th International Conference on Computer and Information Technology, Khulna, Bangladesh.
6. ReLU understanding, available at << https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>>, last accessed on 30-10-2019 at 10:27pm
7. Learn about learning rate, available at<< https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/>>, last accessed on 22-10-2019 at 10:30pm.
8. Confusion matrix, available at <<https://classeval.wordpress.com/introduction/basic-evaluation-measures/ >> last accessed on 29-10-2019 at 10:40pm.
9. Learn about Bangladesh agriculture, available at << https://en.wikipedia.org/wiki/Bangladesh _Agricultural_Research_Institute>> last accessed on 29-10-2019 at 10:40pm.
10. CNN architecture drawing, available at <<http://alexlenail.me/NN-SVG/AlexNet.html>>>>, last accessed on 27-10-2019 at 12:00pm.
11. Traing and testing, available at<< https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7**>>** last accessed on 30-10-2019 at 12:00pm.
12. CNN Architecture description, available at <<https://towardsdatascience.com/building-a-convolutional-neural-network-cnn-in-keras-329fbbadc5f5>>, last accessed on 20-10-2019 at 11:30pm.
13. Learn about Keras, available at << https://towardsdatascience.com/building-a-deep-learning-model-using-keras-1548ca149d37 >>, last accessed on 20-10-2019 at 12:30pm.
14. Learn about sequential model, available at << https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53 >>, last accessed on 22-10-2019 at 10:30pm.
15. Learn about CNN equation, available at << https://www.nature.com/articles/s41598-019-38966-0 >>, last accessed on 06-08-2019 at 10:00pm.
16. SoftMax function, available at << https://en.wikipedia.org/wiki/Softmax_function>>, last accessed on 30-10-2019 at 12:00pm.