Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021).
IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4

# Potato Disease Detection Using Machine Learning

Md. Marjanul Islam Tarik
Department of Computer Science and Engineering Daffodil International University Dhaka, Bangladesh
tarik15-6366@diu.edu.bd

Sadia Akter
Department of Computer Science and Engineering Daffodil International University Dhaka, Bangladesh
sadia15-986@diu.edu.bd

Abdullah Al Mamun
Department of Computer Science and Engineering Daffodil International University Dhaka, Bangladesh
abdullah15-8405@diu.edu.bd

Abdus Sattar
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
abdus.cse@diu.edu.bd

2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) | 978-1-6654-1960-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICICV50876.2021.9388606

***Abstract*— In Bangladesh potato is one of the major crops. Potato cultivation has been very popular in Bangladesh for the last few decades. But potato production is being hampered due to some diseases which are increasing the cost of farmers in potato production. However, some potato diseases are hampering potato production that is increasing the cost of farmers. Which is disrupting the life of the farmer. An automated and rapid disease detection process to increase potato production and digitize the system. Our main goal is to diagnose potato disease using leaf pictures that we are going to do through advanced machine learning technology. This paper offers a picture that is processing and machine learning based automated systems potato leaf diseases will be identified and classified. Image processing is the best solution for detecting and analyzing these diseases. In this analysis, picture division is done more than 2034 pictures of unhealthy potato and potato's leaf, which is taken from openly accessible plant town information base and a few pre-prepared models are utilized for acknowledgment and characterization of sick and sound leaves. Among them, the program predicts with an accuracy of 99.23% in testing with 25% test data and 75% train data. Our output has shown that machine learning exceeds all existing tasks in potato disease detection.**

***Keywords*— *Potato Disease, Image processing, Machine Learning, Disease Detection, Agriculture***

## I. INTRODUCTION

Potato is one of the major harvests in our country. The production of potatoes in Bangladesh is very lower than compared to the other developed countries in the world. The production of potatoes is hampered by Many kinds of Pests and diseases. So we can't export potatoes to our expectations in the other countries. Among them early blight, leaf roll virus, scab, Hollow heart etc are the most terrible disease of potatoes at previous present and times in Bangladesh. In Bangladesh, the major area's farmers faced many hampers on this disease every year. The farmers and businessmen of our country are facing many problems with those diseases particularly in the case of export to other countries as Russia, Indonesia, Malaysia, Sri Lanka, Thailand, Hong Kong, Turkey, Vietnam, Maldives etc. Due to COVID-19, the price of potatoes increases day by day. Even a decade ago, production was below half a million tons. Now it is moving towards billions of tons. This success has brought Bangladesh to the top ten potato producing countries in Qatar. The recognition was given by the Food and Agriculture Organization of the United Nations. According to a report of this organization, Bangladesh is in the eighth place with a production of 722 lakh 10 thousand tons. Not only is this a wonderful achievement in production, potato is now one of the most profitable crops in the country. It has also become a means of earning foreign currency. Potatoes worth 33 million were exported last year. But before that Bangladesh had to import potatoes from 20 countries of the world. We believe that, if we detect the disease of potato properly and provide the proper treatment, the production growth will increase to our expectation.

We have taken the help of image processing to diagnose potato disease and potato leaf disease. Here we have used these parameters to diagnose the disease which can be identified by looking at the characteristics of the disease and the type of disease and we have tried to give more antidotes here. Accuracy is very good with training data in the project and accuracy with sample data is above 99%. So our project will be able to accurately diagnose potato disease and leaf disease.

## II. LITERATURE REVIEW

Recently several works completed within the identification of leaf diseases. Here totally completely different strategies that are analyzed by different researcher's disease detection are reviewed.

This paper planned the magnitude relation of the wavelength of sunshine across that's absorbed the spectrum, that counting on the leaf structure that detect the potato mold victimization Segmentation of Images and Machine Learning. Once the right spectrum up from greenhouse research with associate accuracy of 84.6% diseases were analyzed severally, once models area unit healthy and might distinguish between bays and spectra. Leaves, additionally as 3 categories lately blight wound progression with 92% of accuracy. For trickster, the model distinguish between pre-marked leaves, healthy & black-crowned trees fictions 74.6% classification accuracy [1].

In [2] this work, they gift a strategies for act image process and machine learning to consent to diagnose from leaf illustrations. Trendy phenotyping and disease detection offer committed steps towards food security and property agriculture. This is often automatic methodology that classify diseases in potato from the publicly accessible plant picture information mentioned as 'Plant' Village '. Use our



Authorized licensed use limited to: Florida Institute of Technology. Downloaded on August 04,2021 at 08:56:24 UTC from IEEE Xplore. Restrictions apply.

segmentation approach and support Vector machines show a classification of 300 diseases pictures with the accuracy of 95%.

This paper [3] offers a picture that area unit process and machine learning-based automatic systems Potato leaf diseases are becoming to be known and classified. During this paper, figure the division completed over 450 pictures of healthy and sick Potato leaves, that area unit taken from publicly accessible plant villages Databases and 7 classified algorithms area unit utilized for acknowledgment and also the classification of sick and healthy leaves, Irregular Forest Classifier provides 97¬curacy.

This paper,[4] represent a Uses models that like pre-trained model VG19 for fine-tuning for lifting pertinent properties from the dataset. That's the results of many classifications were complete supplying regression has surpassed others significantly the margin of classification accuracy got over97.8% check dataset.

During this analysis paper they demonstrate Transfer Learning Technology which may be used for early detection of potato mold once it's troublesome to gather pictures of thousands of latest pages. Transfer learning uses already learned deep learning models to resolve new issues. The experiments enclosed pictures 152 healthy leaves, one thousand blight leaves and one thousand early The Blight Page program predicts with 99.43% accuracy. Tested with two hundredth check knowledge and eightieth train knowledge [5].

P.Badar [6] has used how of segmentation victimization K-Man agglomeration on totally different characteristics of potato leaves like color, texture, region, etc. & has applied the back propagation neural network formula for identification and classification of the diseases within the image of leaf with a classification of 92% accuracy.

M.R [7] depicted a straightforward infection discovery framework for the cotton utilizing the leaf picture. A picture of the sickness influenced leaf is taken. At that point utilizing picture handling and Artificial Neural Network distinction is altered over solid and unhealthy examples. What's more, ANN order of 80% precision.

In [8] dealt with the order of three basic paddy leaf sicknesses ( Leaf impact, Brown spot, and Bacterial curse) and taught manures or pesticides k-implies grouping is utilized for the division of the illness affected part and Visual substance are utilized as highlights during this framework. At that point Support, vector machine classifier does the characterization cycle. The general precision is 92.06%.

The system proposed by Yao [9] is aimed at identifying and classifying three types of rice crop damage. In his figure the diseased regions are isolated using the Otsu's method, then divided. The first picks up the tone, shape and surface features, and the second came from HSV shading space. And finally, the properties are submitted to the vector machine, which classifies in last word.

U. Kumari [10] has utilized a methodology of picture Segmentation during which they have extricated different highlights of a picture like difference, relationship, energy, homogeneity, mean, fluctuation and change, and so on. Subsequent to extricating highlights, Neural Network is applied as a classifier to distinguish and arrange sicknesses on the leaves of two plants for example Tomato and Cotton. Utilizing this methodology, they had the option to accomplish a classification accuracy of 92.5%.

The method proposed in this paper [11] was aimed at identifying and discriminating deficiencies of four classes of minerals (nitrogen, phosphorus, potassium and magnesium). Before the resolution, the photos are changed to the HSI and L*a*b* shading spaces. Those differentiations are assessed by Euclidean distances decided in both concealing spaces.

## III. PROPOSED METHODOLOGY

All through the proposed model, the CNN algorithm is utilized to recognize various kinds of potato infections, having 7 classes of potato sicknesses and accomplished 99% accuracy rate.

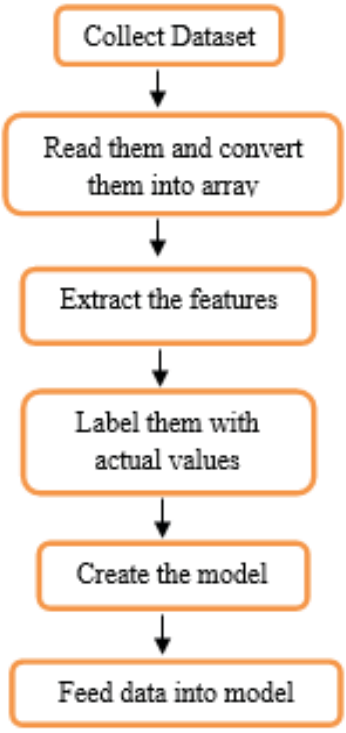


Figure 1. methodology

### *A. Dataset*

We collected this information by taking pictures directly from the potato fields of the village. We were able to collect data on about seven types of diseases of potatoes and potato leaves. We collect total 2034 potato and potato leaves images. We basically split the data between potato leaves and potato disease. While collecting data, we noticed that diseases is more contagious in potato leaves. The class that divides our data is mentioned below:

- potato leaf roll virus
- Hollow heart of potato
- scab of potato
- Soft rot of potato
- Sutali poka rog
- Virus jonito rog



Authorized licensed use limited to: Florida Institute of Technology. Downloaded on August 04,2021 at 08:56:24 UTC from IEEE Xplore. Restrictions apply.

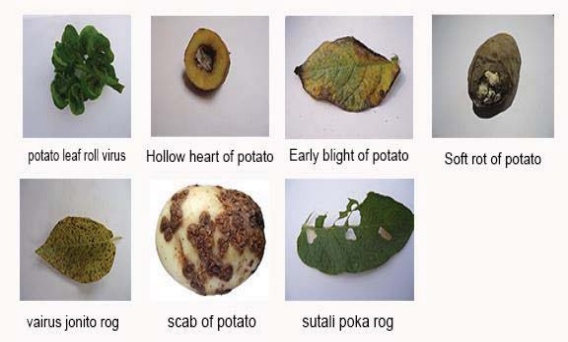


Figure 2. Dataset example

Table I: Dataset

| class | Disease name | Images (number) | Train data | Test data |
|---|---|---|---|---|
| 01 | early blight of potato | 428 | 329 | 100 |
| 02 | potato leaf roll virus | 394 | 274 | 75 |
| 03 | Hollow heart of potato | 221 | 218 | 80 |
| 04 | scab of potato | 154 | 150 | 98 |
| 05 | Soft rot of potato | 238 | 200 | 66 |
| 06 | Sutali poka rog | 305 | 295 | 95 |
| 07 | Virus jonito rog | 294 | 275 | 120 |
| total | | 2034 | | |

### B. *Preprocessing*

In this project we have to take four types of image processing steps to normalize the image, change the color of the image, and identify the properties, Image processing such as filtering and transformation of the image. We have used Python's Opencv Library for this purpose. The features of the OpenCV library are:

- Read & write images
- Capture and save the videos
- Image_processing such as filtering and transformation
- Detection the feature
- Video or picture object detection

The picture document is perused with the OpenCV work the request for colors is BGR. Then again, in Pillow, the request for colors is thought to be RGB.

Figure 3. Data Processing

### C. *Model Architect*

CNNs thinks about piece by piece of picture. The pieces that CNN looks for are called highlights. It finding the harsh element matches in two pictures in similar positions, CNNs improve at seeing closeness than entire picture coordinating plans. Each component resembles a smaller than normal picture, a little two-dimensional cluster of qualities.

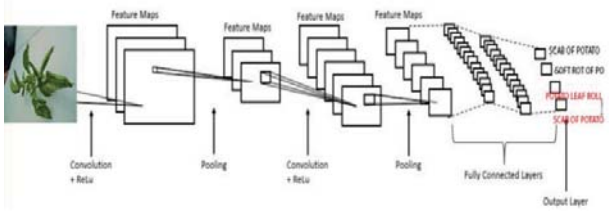


Figure-4: The architecture of Plant Disease recognition

Features match common aspects of the images.

- Give input photograph into convolution layer
- Choose boundaries, apply channels with steps, cushioning if requires. Perform
- Convolution on the picture and apply ReLU enactment to the grid.
- Execute pooling to decrease Dimension size.
- Add as numerous convolutional layers until satisfied.
- Flatten the yield and feed into a completely associated layer.
- Output the class utilizing enactment and order pictures.

Table: Sequential Model

| Layer (type) | Output shape | Param# |
|---|---|---|
| Conv2d (Conv2D) | (None, 98, 98, 32) | 320 |
| Max_pooling2d (MaxPiiling2D) | (None, 49, 49, 32) | 0 |
| Conv2d_1 (Conv2D) | (None, 47, 47, 64) | 18496 |
| Max_pooling2d_1 (MaxPooling2D) | (None, 23, 23, 64) | 0 |
| flatten (Flatten) | (None, 33856) | 0 |
| dense (Dense) | (None, 3) | 101571 |

Total params: 120, 387
Trainable params: 120, 387
Non-trainable prams: 0

## IV. MODEL PERFORMANCE

### A. *Training, Testing and the Validation*

We first create a sequel model of CNN with 7 levels. We used Adam Optimizer to measure performance error and tune cross entropy. We then use transfer learning to create the model. To do that we only use built in module keras, keras applications that provide pre-trained weights. We modified the Deep Learning model so that the pre-trained weights matched our desired output dimension by dropping the last few layers and adding few Junne lavens.



Authorized licensed use limited to: Florida Institute of Technology. Downloaded on August 04,2021 at 08:56:24 UTC from IEEE Xplore. Restrictions apply.

Figure 5. Validation data

```
[ ] training = model.fit(X, y, epochs=15)

Epoch 1/15
35/35 [==============================] - 12s 330ms/step - loss: 1.7085 - accuracy: 0.3339
Epoch 2/15
35/35 [==============================] - 12s 331ms/step - loss: 1.0872 - accuracy: 0.6155
Epoch 3/15
35/35 [==============================] - 12s 330ms/step - loss: 0.7692 - accuracy: 0.7453
Epoch 4/15
35/35 [==============================] - 11s 325ms/step - loss: 0.6121 - accuracy: 0.8178
Epoch 5/15
35/35 [==============================] - 11s 326ms/step - loss: 0.5113 - accuracy: 0.8161
Epoch 6/15
35/35 [==============================] - 12s 329ms/step - loss: 0.4134 - accuracy: 0.8471
Epoch 7/15
35/35 [==============================] - 12s 331ms/step - loss: 0.2743 - accuracy: 0.9249
Epoch 8/15
35/35 [==============================] - 11s 327ms/step - loss: 0.2451 - accuracy: 0.9284
Epoch 9/15
35/35 [==============================] - 11s 328ms/step - loss: 0.1661 - accuracy: 0.9589
Epoch 10/15
35/35 [==============================] - 11s 328ms/step - loss: 0.1361 - accuracy: 0.9624
Epoch 11/15
35/35 [==============================] - 11s 326ms/step - loss: 0.1027 - accuracy: 0.9807
Epoch 12/15
35/35 [==============================] - 11s 327ms/step - loss: 0.0743 - accuracy: 0.9877
Epoch 13/15
35/35 [==============================] - 11s 327ms/step - loss: 0.0605 - accuracy: 0.9908
Epoch 14/15
35/35 [==============================] - 12s 329ms/step - loss: 0.0432 - accuracy: 0.9957
Epoch 15/15
35/35 [==============================] - 12s 329ms/step - loss: 0.0408 - accuracy: 0.9973
```

Figure 6: Training model

### *B. Model performance*

We have done 15 epochs, This model has made the progress rate is 98.36% in train Dataset and 98.81% in preparing approval set that we make. Trial with arbitrary pictures the instructional meeting went easily after the end. It was best aftereffect of precision Overall. Subsequent to dissecting the result and disarray grid, it is perceptible that the exhibition of our model is satisfactory. The exhibition of our model is given beneath :

```
plt.plot(training.history['accuracy'])
[<matplotlib.lines.Line2D at 0x7f124ab38eb8>]
```

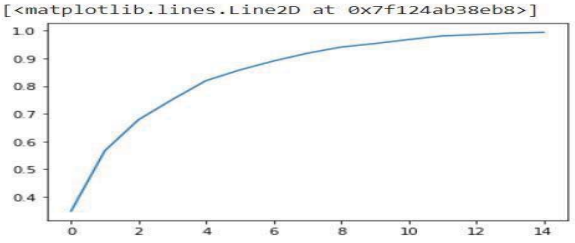


Figure-7. Training and validation accuracy

```
plt.plot(training.history['loss'])
[<matplotlib.lines.Line2D at 0x7f124c868860>]
```

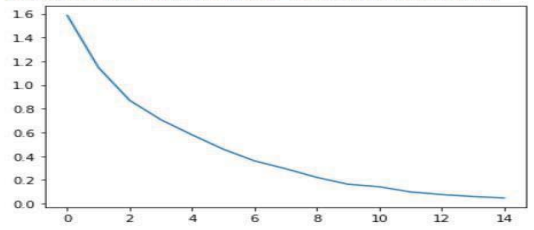


Figure-8: Training and validation loss

## V. CONCLUSION

This thesis mainly focuses on disease detection of potatoes from any surface by using machine learning (CNN). We found that CNN is the best way to perform this type of detection object. However, this model gains 99% of validation accuracy. We have a large amount of data set and to get the best accuracy, we have tried our best. We think this type of project will play a vital role in our agriculture sector. Most of the farmers of the village in Bangladesh are not literate and they can't know about the disease properly. They can't know the method of detecting disease. That's why the insect is destroying the potato and our farmers get to suffer from it. We think that, this work can change the situation of the potato grower in Bangladesh. In future, our aim is to create an android application that can detect the disease of every type of crop and can provide the proper solution for those diseases of the crop. In future by increasing our database, we will able to get better accuracy. By building an android app we will continue the development process. And we will create a system where the farmers of Bangladesh can easily get instant service and advice on their problem by detecting the disease.

## REFERENCES

[1]. Damian Bienkowskia, Matt J. Aitkenheadb, Alison K. Leesc, Christopher Gallaghera, Roy Neilsona "Detection and differentiation between potato diseases using calibration models trained with non-imaging spectrometry data" Computers and Electronics in Agriculture 167 (2019) 105056.

[2]. Monzurul Islam, Anh Dinh, Khan Wahid ,Pankaj Bhowmik "Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine" , 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE).

[3]. Md. Asif Iqbal and Kamrul Hasan Talukder,"Detection of Potato Disease Using Image Segmentation and Machine Learning"

[4]. Divyansh Tiwari, Mritunjay Ashish,Nitish Gangwar,Abhishek Sharma, Suhanshu Patel ,Dr. Suyash Bhardwaj."Potato Leaf Diseases Detection Using Deep Learning". 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.

[5]. Farabee Islam,Md Nazmul Hoq,Professor Dr. Chowdhury Mofizur Rahman." Application of Transfer Learning to Detect Potato Disease from Leaf Image".

[6]. Athanikar, Girish and Ms. Priti Badar. "Potato Leaf Diseases Detection and Classification System Mr." (2016).

[7]. Malvika Ranjan, Manasi Rajiv Weginwar, NehaJoshi, A.B. Ingole,"Detection and Classification of Leaf Disease Using Artificial Neural Network," International Journal of Technical Research and Applications, 2015, pp. 331–333.

[8]. F. T. Pinki, N. Khatun and S. M. M. Islam, "Content based paddy leaf disease recognition and remedy prediction using support vector machine," 2017 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, 2017, pp. 1–5.

[9]. Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B, "Application of support vector machine for detecting rice diseases using shape and color texture features," 2009 International Conference on Engineering Computation, IEEE, Hong Kong, 2009, pp. 79–83.

[10]. C. U. Kumari, S. Jeevan Prasad and G. Mounika, "Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 1095-1098.

[11]. Wiwart M, Fordonski G, Zuk-Golaszewska K, Suchowilska E "Early diagnostics of macronutrient deficiencies in three legume species by color image analysis," Compute Electron Agric 65, 20



Authorized licensed use limited to: Florida Institute of Technology. Downloaded on August 04,2021 at 08:56:24 UTC from IEEE Xplore. Restrictions apply.