A WIDE CONVOLUTION NETWORK BASED CLASSIFICATION PATH FOR RECOGNITION PAPAYA LEAF DISEASE

BY

Kaniz Fatima Ripa

ID: 221-25-100

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering

(Major in Data Science)

Supervised By

Dr. Sumit Kumar Banshal

Assistant Professor

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Thesis titled **“A Wide Convolution Network Based Classification Path for Recognizing Papaya Leaf Disease”**, submitted by **Kaniz Fatima Ripa**, ID No: **221-25-100** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17-01-2023.

BOARD OF EXAMINERS

Chairman

Dr. S M Aminul Haque, PhD
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Ms. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Md. Sadekur Rahman
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Mohammad Shorif Uddin, PhD
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

I hereby declare that this research has been done by me under the supervision of **Dr. Sumit Kumar Banshal** Assistant Professor. I also declare that neither this research nor any part of this research has been submitted elsewhere for the award of any degree.

Supervised by:

Banshal

Dr. Sumit Kumar Banshal
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Kaniz Fatima Ripa

Kaniz Fatima Ripa
ID: 221-25-100
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully. I really grateful and wish my profound my indebtedness to **Supervisor**, **Dr. Sumit Kumar Banshal, Assistant Professor,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project. I would like to express my heartiest gratitude to the Almighty Allah and Head**,** Department of CSE, for his kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.I would like to thank my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

The top papaya-producing countries are India, Mexico, Brazil, Indonesia, Nigeria and the Dominican Republic. But in papaya especially leaf curl, foot rot of papaya, papaya ring spot and many other diseases are causing great damage to papaya yield. If the diseases are not determined in the early stages, papaya production will decline. The main goal of this piece of research work is to develop and put into effect an algorithm for diagnosing papaya illnesses at the early stage, which are Papaya leaf curl virus (PaLCuV), Papaya ring spot disease, Anthracnose, Foot rot of papaya and Papaya mosaic disease. This paper presents a technique for papaya disease detection using photograph processing techniques, to understand papaya field diseases from images, primarily based on the color, texture and form of diseased papaya and provide appropriate options to farmers, so that papaya diseases can be prevented at an early stage and as a result higher production of papaya can be achieved. I used the cross-validation technique in this model and get the accuracy of 87% which is more than expectation.

TABLE OF CONTENTS

CONTENTS **PAGE**

CHAPTER **PAGE**

LIST OF TABLES PAGE

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Introduction

Bangladesh is an agriculture dependent country. About 80% of the people work in the agricultural industry. The agricultural sector plays a very important role in economic development. In agriculture Papaya is important products as it is regarded to be one of Bangladesh's popular vegetable and cereal crops and for this reason, many attempts have taken to ensure its safety, one of them is crop management of papaya plants. Papaya plants are affected by various fungal and bacterial diseases. This work focuses on recognizing three papaya plant diseases namely Papaya leaf curl disease, Powdery mildews and Papaya mosaic disease. The proper detection and recognition of disease is very important in applying required fertilizer. Powdery mildew is one of the most common fungal disease and most damaging papaya plant disease. In this disease, they are at first little, the disease occurs on leaves and pods. The infection first appears as small dark spots on the leaves, which later turn into white powdery spots. These spots enlarge and cover the entire leaf area. Severely affected leaves may become chlorotic and deformed before dropping. Affected fruits are, affected fruits are small in size and deformed. In more severe cases dieback symptoms develop. Papaya mosaic disease Causes leaf mosaic and stunting in papaya. Papaya leaf curl disease is caused by Papaya leaf curl virus (PaLCuV). Major symptoms of papaya leaf curl disease are inward or outward curling of plant leaves, main objective of this work is to develop a system for classifying the papaya plant diseases using image processing technique.

1.2 Motivation

Agriculture sector has a great impact on the economy of Bangladesh. But still the uses of IT in this sector is very low and our farmers are still lack behind to use the modern technology of agriculture. In today's digital age, it is important that the farmers get to use the contemporary technology for efficient management of their crops. The use of facts access through cellular phones among the farmers has expanded in latest years, which has made a wonderful have an impact on the output of the production. However, there is lacking a missing of know-how sharing between the farmers and the agriculture specialists whilst it comes on a theme of suited crop administration due to the challenges in training the farmers on a mass level on subjects like disorder identification and their management. As a result, in most of the cases the farmers count number on their trip and instinct for choice on figuring out crop illnesses and their treatments. The production might turn out no longer as predicted if the symptoms are now not handled in a suited manner, the usage of excellent amount of fertilizers guided by way of agriculture specialist. The motivation of this paper is to find out the diseases and solve the problem. Image processing methods have been utilized in this find out about for identification of 3 papaya illnesses named Papaya leaf curl disease, Powdery mildews and Papaya mosaic disease.

1.3 Research Questions

- Is it always possible to detect a disease by farmer?
- Can we collect the original papaya diseases data for deep learning research?
- How the prediction of papaya diseases will be more accurate by the approach?
- What is the impact of the existing system and agriculture related application on the detection of the papaya disease?

1.4 Expected Outcome

In agricultural sector of my country, our farmers are not yet able to detect Papaya diseases properly. Using this program, farmers can easily detect Papaya diseases. And properly identifying the illness can help them to take the right step. As the result farmer will be benefited.

- Detecting the disease in early stage can help the farmer to take necessary steps to do early remedy of Papaya disease.
- Farming will be easier than before to farmers which will put a strong impact on economic growth of the country.

“Happy Farmer, growth of production”

1.5 Layout of the Report

Chapter one have demonstrated an introduction to the project with its motivation, research questions, and expected outcome.

- Chapter two will have “Background” demonstrates introduction, related works, Papaya diseases and challenges.
- Chapter three will have Research Methodology.
- Chapter four will have Experimental Results and Discussion.
- Chapter five will have Conclusion and Future works.

CHAPTER 2

Background Study

2.1 Introduction

In this division, I will consult with associated works and obstacles about this research. In related works area, I will analyse other research paper and their works, their methods, and perfection which are related to my work. In research overview section I will give the recapitulation of our related works. In challenges section, I will discuss how we increased the perfection level.

2.2 Papaya Diseases

Disease damages to Papaya will significantly decrease production rate. They're mainly caused by bacteria, viruses, or fungi. In most of the instances the ailments produce visual symptoms, specifically growing spots on the leaf body of papaya. The foremost common diseases of papaya are Foot rot of papaya, Powdery mildews, Anthracnose, Papaya mosaic disease, Papaya ring spot disease, Papaya leaf curl disease, IPM for Papaya and lots of others. For applying inventive and judicious based completely disease awareness supported visible indications, this paper emphasizes on three diseases named Papaya leaf curl disease, Powdery mildews and Papaya mosaic disease. Normally, Disease to the vegetations that proceeds residence when a virus, bacteria influence a plant and disarrangements its natural production.

Papaya leaf curl disease

Papaya leaf curl virus (PaLCuV) is responsible for Papaya leaf curl disease, a begomovirus carried by whitefly. Major symptoms of papaya leaf curl disease are inward/outward curling of plant leaves, vein fatting, and plant buildup with scanty deteriorated fruits or no fruits. Peach leaf curl, also known as leaf curl, is a disease caused by the fungus Taphrina deformans. Papaya leaf curl virus is a great hazard for the production, and the virus has the ability to suit new plant hosts very promptly which assists in their host expansion that also has come into view as an expanding peril in papaya cultivation. Actually, there is no chemical treatment for Papaya leaf curl disease viral infections. There is no chemical treatment for viral infections. However, holding the population of whiteflies in check can reduce the severity of the infection. Soil application at the time of sowing and 4-5 foliar sprays of dimethoate or metasystox at an interval of 10 days can effectively control whitefly populations. Howbeit, holding the population of whiteflies in checkup can rebate the austerity of the infection. Clay utilization at the time of growing and 4-5 foliar sprays of dominate or metastasize at an intervolved 10 days can energetically restriction whitefly populations. So, Papaya leaf curl virus disease (PaLCuV) is an important idiopathic that is answerable for low-rent raw material, now that is substantially jolting the peculiarity production of leaves. It is indispensable to study the impingement of disease on physiological and biochemical domains of papaya leaves spectacles so that an advisable extenuation subtlety can be developed.

Powdery mildews

Powdery mildew affects papaya plants regardless of their maturity. The disease is caused by the fungus Oidium caricae-papayae. Development of powdery mildew is influenced by many environmental factors including temperature, RH, light, and air circulation. Because these optimum conditions usually occur in mid to late summer, powdery mildew outbreaks are most common at that time. The first sign of powdery mildew is pale yellow leaf spots. White powdery spots can form on both upper and lower leaf surfaces, and quickly expand into large blotches. The large blotches can cover the entire leaf, petiole and stem surfaces. Powdery mildew is a fungal disease that affects a wide range of plants. Powdery mildew diseases are caused by many different species of ascomycete fungi in the order Erysiphales. Powdery mildew is one of the easier plant diseases to identify, as its symptoms are quite distinctive. The fungus survives and reproduces on papaya plants only. The spores are dispersed from plant to plant and between fields by wind. Leaves at all growth stages can be affected, but older leaves are more susceptible. The disease constantly affects immature leaves but can also harm unripe fruits. The powdery mildew also develops on papaya petioles, pedicels, and peduncles.

Symptoms of powdery mildews are:

- White stripes of fungal augmentation progress on the under surface of the leaf.
- Leaf boundery curl upwards, disclosing the white, powdery fungal growth buildup.

- Mini, circular, dark fungal shapes (cleistothecia) may also appear on the beneath of the leaves.

Actually, Powdery mildew is acquired by a group of related fungi in the Erysiphe family. Powdery mildew spores are readily lugged by the wind to adjoining plants or to plants swarms of miles away. Moreover, once a spore area on a host plant, it will immediately generate and beginning a new infection.

Papaya mosaic disease

Papaya ring spot virus is responsible for Papaya mosaic disease, which arises in total aerial sections of host apart from the embryos. The disease come out as ample mottling and puckering of juvenile leaves and consecutively expanding ones. Symptoms more noticeable than mosaic are decreased size, chlorotic and malformed presence of young leaves. Defoliation of aged leaves results in presence of bunch of mini leaves at the uppermost. Affected plants display degeneration and marked reduction in production. As the virus is passed by aphids, spray of insecticides to retain the vectors away to abate the degree of disease. so, the virus is a monopartite strand of decisive sight, specific wrecked RNA encompassed by a capped built for a separate fervid encoded protein. Papaya mosaic diseases is mechanically transmissible viruses associated with other viral disease, from papaya mosaic virus in being aphid-borne and restricted in host range to papaya and cucurbits. The phenom has been painstakingly squeezed. It is bequeathed by stereotyped prevention. It is not bequeathed in ovum and no termite vector is convoluted. So, the mellow virials form blended sub Sumption that can be seen in leaf segment of contaminated plants when threadbare with either the Orange-Green proteron smudge or the Azure A nucleic acid stair. Mosaics viruses are irremediable- admitting sometimes they intelligibly create contrastingly repeated leaves without approximately abbreviating a plant’s vigor. Actually, Leaves are pitted or streaked due to mosaic virus damage. Another possible diagnostic includes anamorphic or diminutive growth, miniaturized yield, and low product peculiarity.

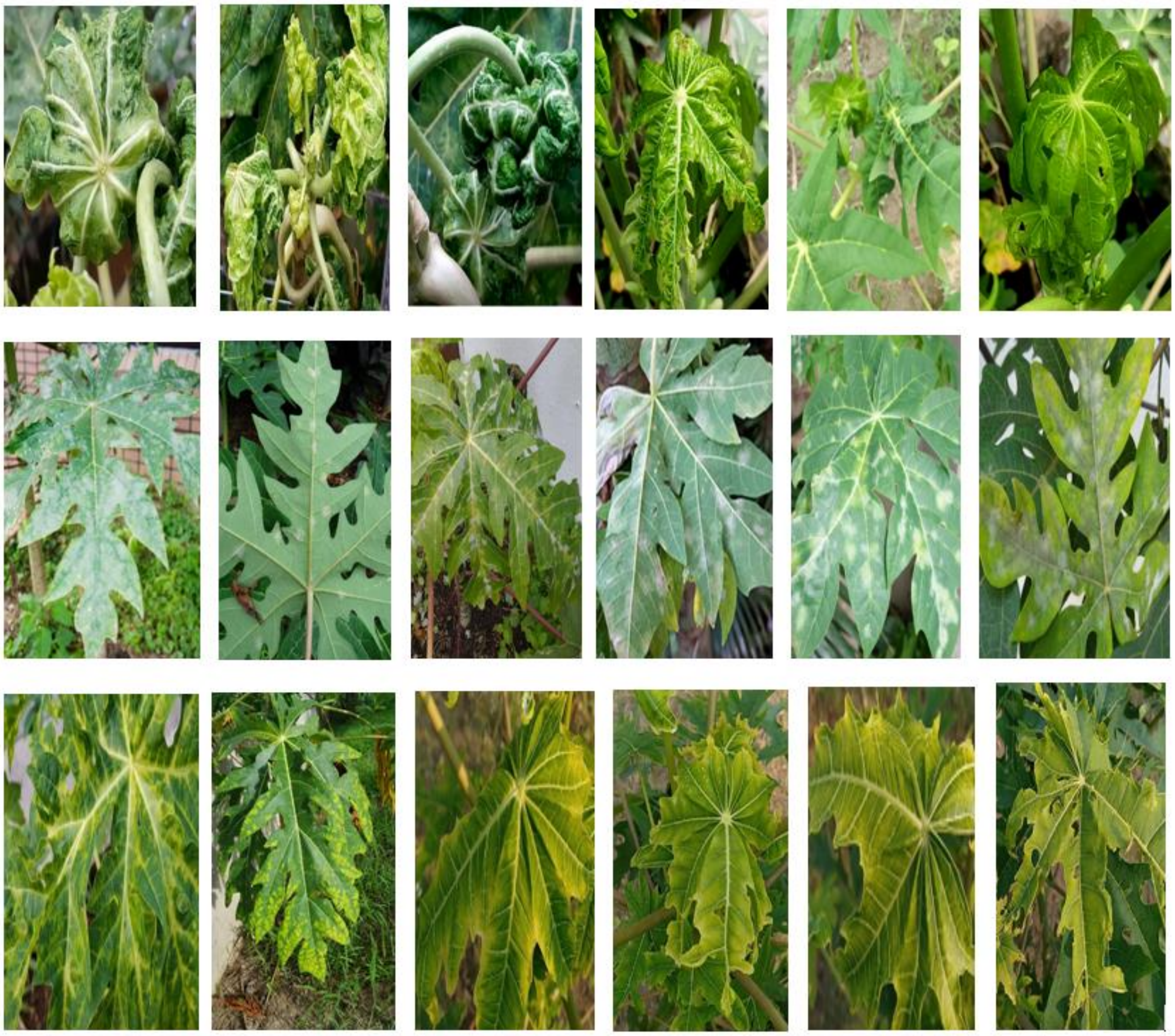


Fig 2.1: A small part of our dataset

2.3 Related Works

In last few years, there are lots of work have taken place in agricultural sector in Bangladesh especially in the plant protection field which ultimately leads to crops management. There are so many applications and research work on it. Described different features like shape, the color of a disease part of the leaf have been removed by developing an algorithm. To achieve scrutiny of leaf image, K-means unsupervised algorithm was used. To perform component subsequent counseling effectively clarification of accomplished Leaf images in a complicated backdrop with different encouraging condition. Shanwen Zhang [1]. Author using K-means unsupervised algorithm. Image demarcation technique is used to recognize contaminated parts of leaf by using

K-means clustering algorithm. In this depreciat, the clarifier is used as artificial neural network [2]. Papaya Leaf curl disease is explained by leaf curl. It can scheme by fungus or virus [3]. H.Al-Hiary, proposed revelation of plant diseases using industrialization and codify its diseases. Here constituent is concerted on set of lineaments into utter k classes. That there is more assemblage cause whereas leaf has more than one disease then disease. ANN is recycled to apprehension and collocation of disease. Farther research requirements to expansion the accuracy of apprehension [4]. The author has used a reference object to measure leaf area Zhou et.al has introduced a novel strategy for vigorous and early Cercosporin leaf spot identification in sugar beet utilizing hybrid extraction Gabor wavelet conversion was used. Author of this paper experimented with homemade dataset. In algorithm of template matching support vector machine. This technique is schedule identification and persistent quantization under normal light conditions [5]. Malti recommended a computerized disease apprehension system for plant leaf of Phaseolus vulgaris (Beans) and Camellia assamica (Tea). So, this study sophisticated image accretion, image pre-processing, image distribution, feature eradication and distribution. Then, Image reduction was depleted by using k-means clustering. Color Co-occurrence method was used for catheterize color and intermixture lineaments of an image [6]. Sanjay, this work concentrates on brown spot and which is a fungal disease construct on sugarcane. Other hand the Images of brown spot indisposed leaves were appropriated for this study. Two category of image dissolution were worn. Transparent Threshold vivisection was used for discovery out range of leaf and dalliance thresholding methods were pre-owned for recommendation out abrasion suburb area. This experimentation determined by average accuracy of 98.60% [7].

2.4 Challenges

The accuracy of detecting Papaya diseases was not high. I tried to increase the accuracy of Papaya diseases by using traditional machine learning like transfer learning, convolutional neural network, Inception-V3 etc.

CHAPTER 3

Research Methodology

3.1 Data Collection Procedure

In my model I use images as input to classify the papaya diseases. My input dataset contains 1500 images from which I use 1200 images to train the model and another 300 images to test the model. The trained images contain three classes: papaya leaf curl disease which has 400 images, powdery mildews which has 400 images and papaya mosaic disease which has 400 images. The images have different dimensions and shapes which are collected day by day from field by using our camera. I select the most affected papaya leaf. This is our real data collection. I don't download any image from google or any other sources. I used Inception-V3 model to build run the model, train it and test its performance. My model produces a rating for each of the possible results then each of them is transformed to a probability through softmax. The dimensions of images is:

Table 3.1: Dimension of images

Number of pictures	Dimensions
03	320×175
49	320×163
107	320×119
135	320×192
146	320×143
170	320×142
205	320×185
322	320×110
401	320×151
499	320×139

3.2 Statistical Analysis

I select 1200 images to train and retrain Inception-V3's final layer. Then I test the retrained model. I test the model with 300 images and check each image if it is a Papaya leaf curl disease or Powdery mildews or Papaya mosaic disease.

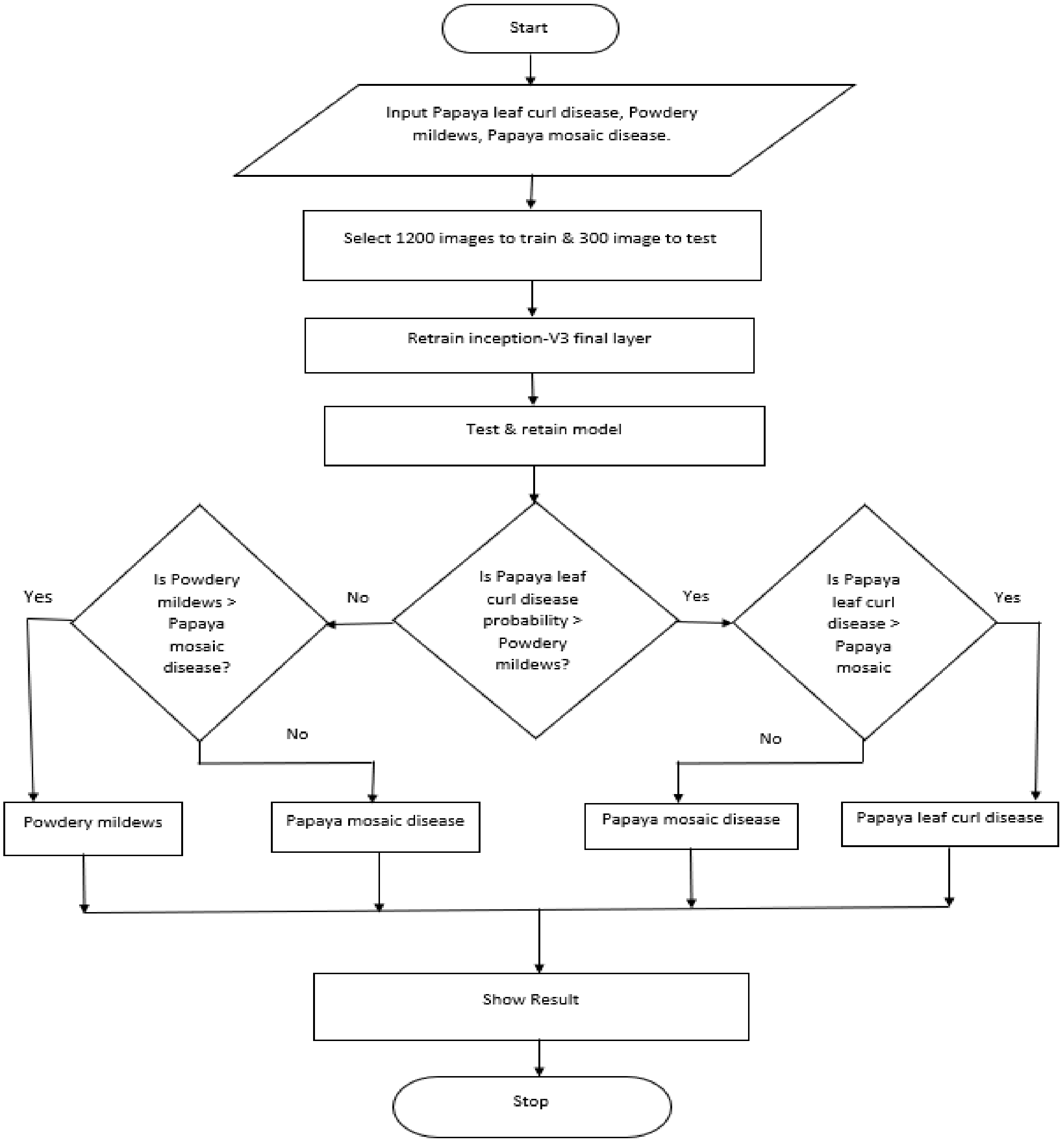


Fig 3.1: Flow Chart of Proposed Model

3.2.1 Inception-V3

Inception-V3 is a model which is trained to classify an image with an error rate that proposed human performance. An important trick factorization is introduced in Inception-V3 which factorizes big kernels into small kernels such as one 7x7 kernel= two 5x5 kernels with strides 2 = three 3x3 kernels with stride 1.

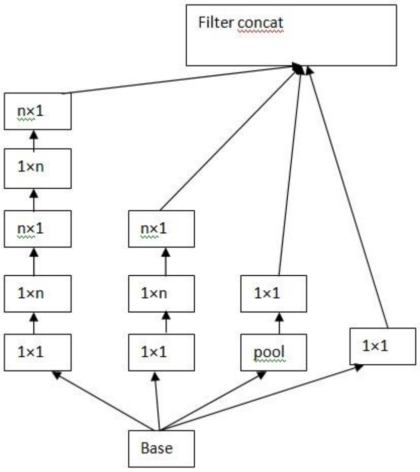


Fig 3.2: Inception module after the factorization the n×n convolutions.

In above picture, after factorizing n * n into 1 * n and n * 1, the inception module was chosen n=7. A 17x17 grid looks like this:

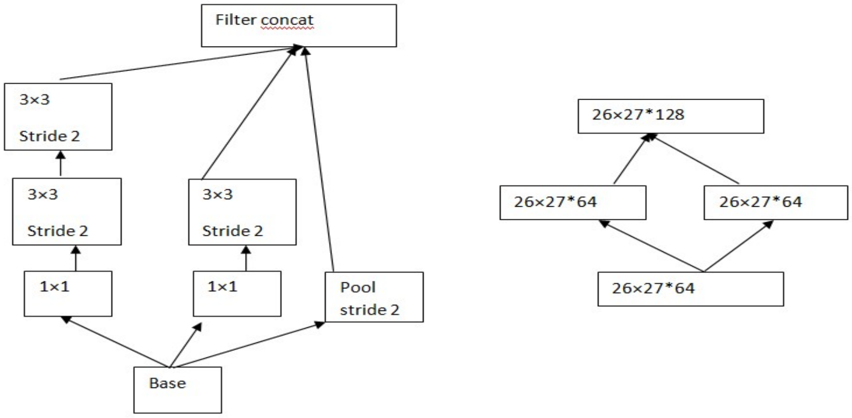


Fig 3.3: Inception module that reduce the grid size while expands the filter banks.

In this figure inception module reduces the grid size and expands the filter banks which is at the same time cheap and avoids representational bottlenecks. I use Inception-V3 to extract the feature from my input image using my CNN model and then classify both disease and not disease part by using fully-connected layer and soft ax layer.

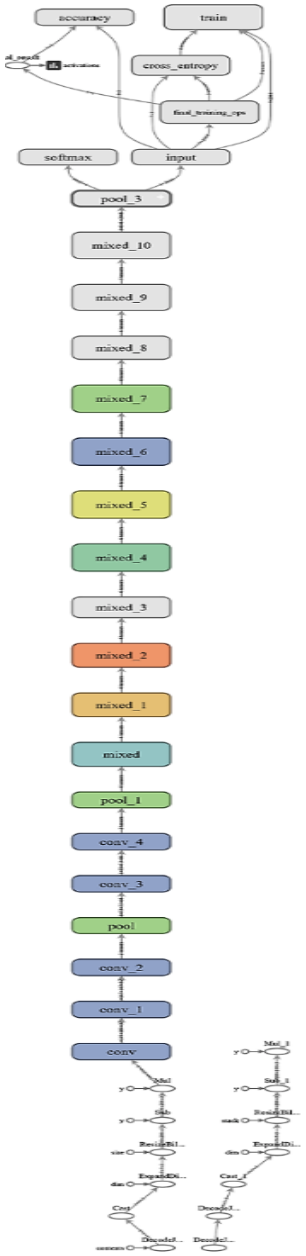


Fig 3.4: Final architecture of inception v3

3.2.2 Transfer Learning

Transfer learning uses the knowledge gained from solving one problem and apply it to another related problem. It use to transfer learning to relate my data set. Transfer learning is the application of knowledge gained from completing one task to help solve a different, but related, problem. The basic premise of transfer learning is simple: take a model trained on a large dataset and transfer its knowledge to a smaller dataset. Transfer learning allows developers to circumvent the need for lots of new data. A model that has already been trained on a task for which labeled training data is plentiful will be able to handle a new but similar task with far less data. There are other benefits of transfer learning through deep learning as well. Transfer learning can be effectively used for carrying out image recognition tasks. Facing the hassle of amassing sufficient education information to rebuild models, switch gaining knowledge of objectives to transfer know-how from a giant dataset recognized as source area to a smaller dataset named target domain. Either the characteristic spaces between area data are unique or the supply duties and the target duties center of attention on distinctive topics, boosting the performance of the goal task. Transfer studying using CNNs is often used in distinctive fields. Three fundamental situations of transfer learning are, ConvNet as constant feature extractor, whereby taking a ConvNet pre-trained on ImageNet the ultimate fully-connected layer is eliminated and the relaxation of the ConvNet is handled as a constant feature extractor for the new dataset two. The second one is Fine-tuning the ConvNet, used to fine-tune the weights of the pre-trained network by continuing the back propagation. It’s possible to keep some of the earlier layers fixed and only fine-tune some higher-level portion of the network or to fine-tune all the layers of the ConvNet. In pre-trained models, the network is trained on large dataset like ImageNet and also trained all the parameters of the neural network and it takes hours on GPU. I use the pre-trained Inception-V3 for our analysis by loading all layers of the Inception-V3 with pre-trained weights. I used to transfer learning to retrain my dataset which is different from Inception-V3 images. To classify my dataset, I build a new model, then reuse the feature extraction part and re-train the classification part with our dataset.

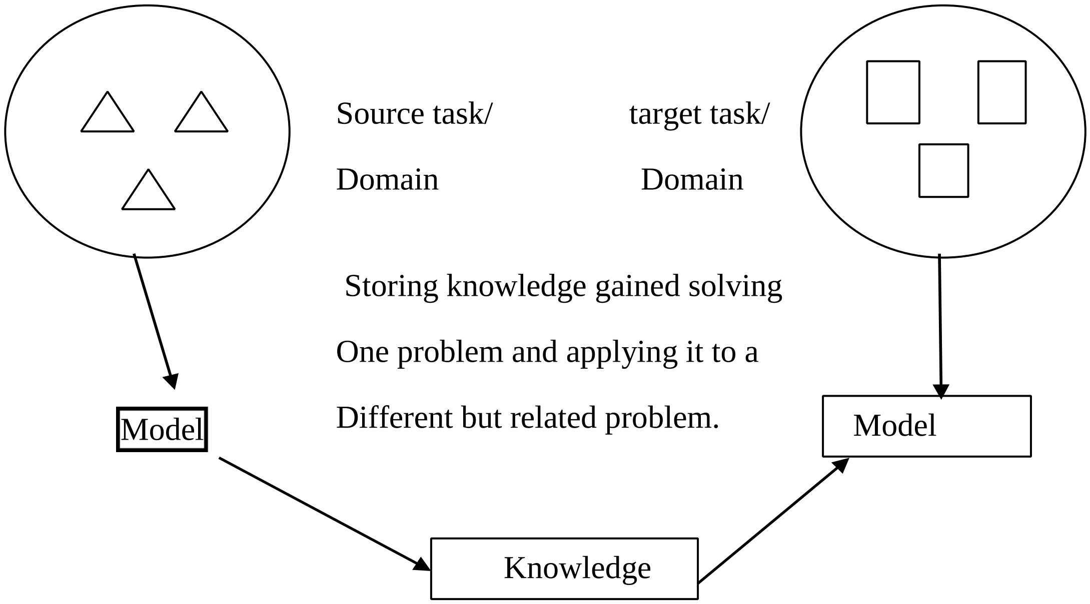


Fig 3.5:Transfer Learning

3.2.3 Convolutional layer

Convolution layer extract feature from an input image. A convolutional operation is performed to the input and then passes the result to the next layer. Using small squares of input data, convolution learns image features and preserves the spatial relationship between pixels. CONV layer's parameters are made of a set of learnable filters. Every filter is small spatially (along width and height), but extends through the full depth of the input volume. An example of convolution layer is:

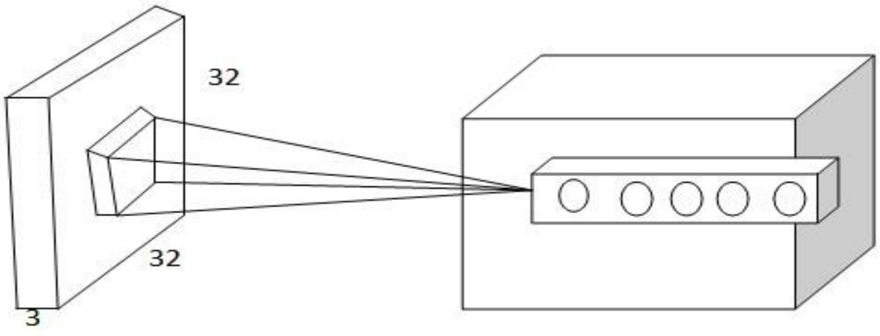


Fig 3.6: Convolutional layer

It is an input volume in red (e.g. a 32x32x3 CIFAR-10 image), and an example volume of neurons in the first Convolutional layer. Each neuron in the convolutional layer is connected only to a local region in the input volume spatially, but to the full depth and there are multiple neurons (5 in this example) along the depth, all looking at the same region in the input.

3.2.4 Rectified linear unit (ReLU)

Rectified linear unit or ReLU is the activation function which is commonly used in deep learning networks for hidden layers. The function returns 0 if the input is less than 0 and if the input is greater than 0 then the output is equal to the input. Its derivative is either 0 or 1. When the input is positive the derivate is just 1 so there is no squeezing effect on back propagated errors. It can be written as,

$$(x) = \max(0, x)$$

Where x is the input to a neuron. ReLUs result in much faster training for large networks. ReLU does not saturate and the gradient is always high if the neuron activates.

3.2.5 Max pooling

Pooling or down sampling reduces the dimensionality of each sub-region but saves the most significant information. Pooling layer is inserted in-between every consecutive Conv layers in convolutional neural network architecture. I used max pooling to reduce the dimensionality and get the highest element of each sub-region. Max pooling uses the maximum value from each sub-region of every node at the previous layer. Max pooling is discarded 75% of the activations and controlling over fitting.

3.2.6 Fully Connected Layer

Fully Connected layers are not defined by the number of nodes, just by how they are connected to adjacent layer`s nodes. The fully connected layer also introduced by dense layers used in classification adding previous layer neurons to every neuron on the next layer. Different types of function like softmax activation function, SVM, and many others are used here for high-level reasoning in the neural network. But in my model, I stick used softmax for classification. After several convolution and pooling layers, I get some high-level features as input. These input images features are used as classifying to explore various classes. But when we combine convolution layer`s features and pooling layer`s features it gives the better result of classifications. In Fully Connected layers summation of output probabilities is 1.One Conv layer share weights with other Conv layers.It is very difficult to attach all nodes with a softmax layer that’s why we use a fully-connected layer to increase the efficiency of classification in our model.

3.2.7 Softmax

Let us consider a classification model to classify with n classes. This model takes input datasets and an algorithm and produces a score of each class. The softmax activation function converts from score to the probability between 0 to 1.the summation of all probabilities is 1.I used this function to the final layer of convolutional neural networks to classify the classes. This function is produced multiple class from an input array. The probability distribution of softmax function is:

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_{j=1}^{n} e^{x_i}}$$

Where i=1,2,3,…….n and j=1,2,3,…….n

3.2.8 Cross Entropy

Cross-entropy loss increases as the estimated chance diverge from the true label. Cross-entropy loss, or log loss, measures the overall performance of a classification model whose output is a chance value between zero and 1. In our classification tasks to classify illnesses based totally on images of disease, a very common kind of loss function to use is Cross Entropy loss. It is defined as

$$(p, q) = [-\log^{q}] = (p) + D_k\ (p||q)$$

Where H (p) is the entropy of p.

$(p||q)$ is the kullback-leibler divergence of q from p.

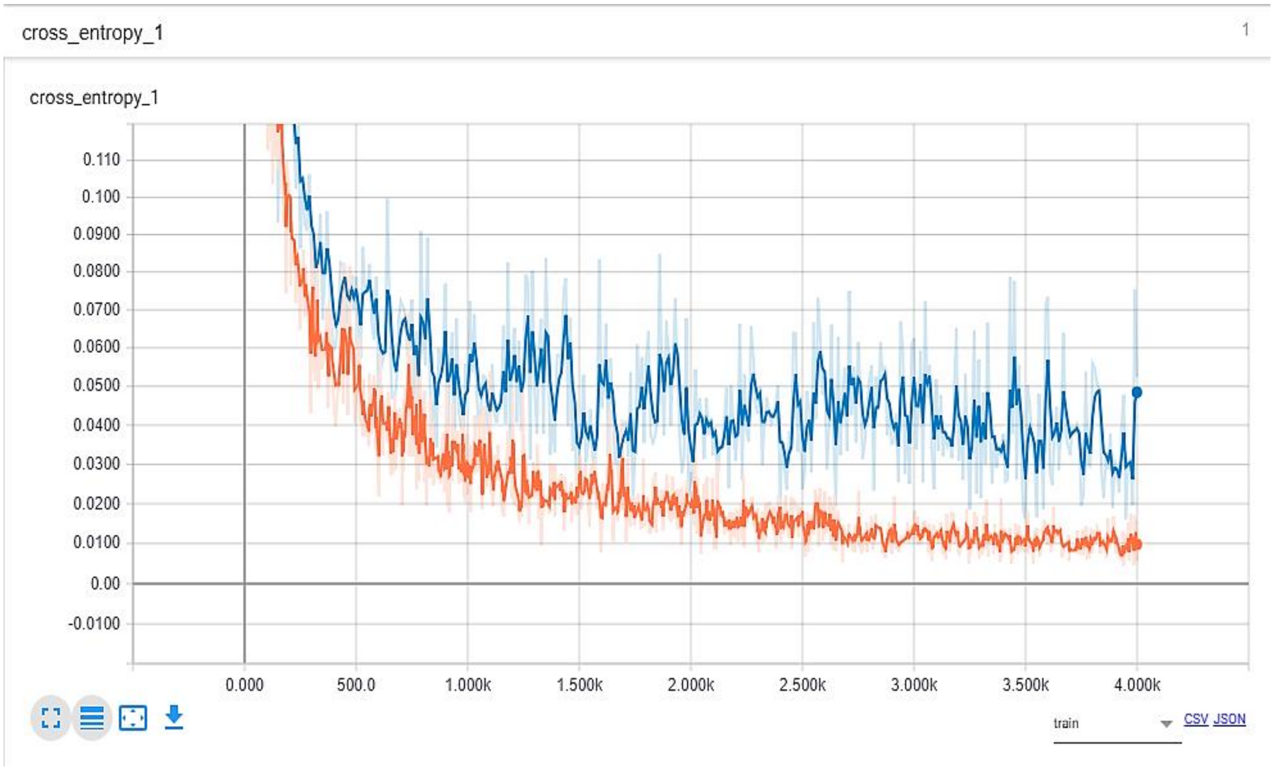


Fig 3.7:Cross entropy error function

3.3 Proposed Methodology

In my proposed method I use the convolutional neural network. I use main four type of layers to build my architecture: convolution layer, ReLu layer,max-pooling and fully connected layer and use softmax function to classify the probability of all the outputs.

I measure the dimension from the previous activation in each convolution layer by the following equation is given by

$$width \quad N_w^{[l]} = \frac{N_w^{[l-1]}+2p^{[l]}-F_w^{[l]}}{S^{[l]}}+1 \quad \text{------------------- (1)}$$

$$height \quad N_h^{[l]} = \frac{N_h^{[l-1]}+2p^{[l]}-F_h^{[l]}}{S^{[l]}}+1 \quad \text{----------------------- (2)}$$

Where p = padding size

s = number of stride

F_w=width of the convolution filter

F_h=height of the convolution filter

$N_w^{[l-1]}$=width of the previous layer

$N_h{}^{[l-1]}$ =height of the previous layer

And measure the max pool layer by the given equation:

$$maxpool = \frac{N_w^{[l-1]} - F_w^{[l]}}{S^{[l]}} + 1 \times \frac{N_h^{[l-1]} - F_h^{[l]}}{s^{[l]}} + 1 \times N_c \text{ ------------------------ } (3)$$

Where,

$N_w^{[l-1]}$=width of the previous layer

$N_h{}^{[l-1]}$ =height of the previous layer

F_w=width of the convolution filter

F_h=height of the convolution filter

Nc =number of channel

Measure the ReLU activation by the equation:

$$(x) = \log\,(1 + e^x)$$

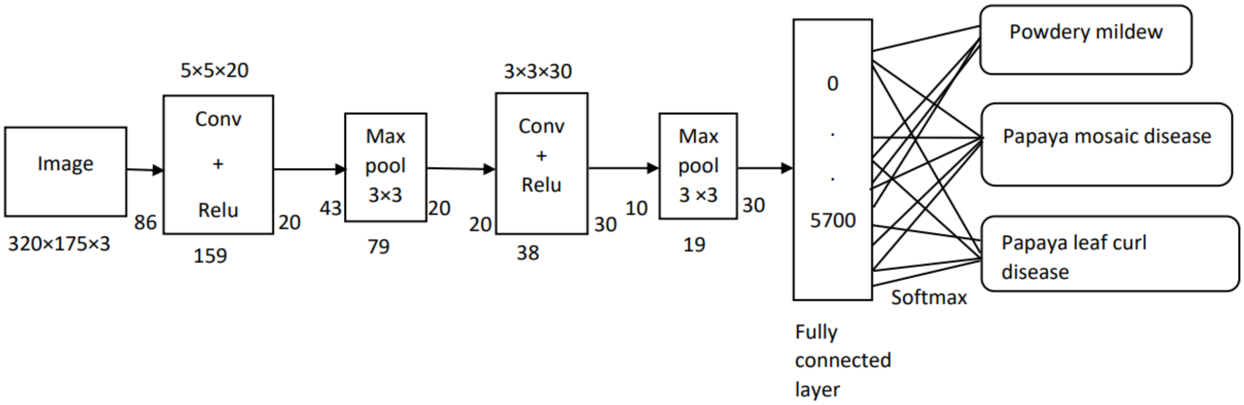


Fig 3.8: Convolutional neural network

My model take input as raw pixels value of the image with dimension 320×175×3.the dimension is described as width: 320, height: 175 and the number of channels are 3.channel 3 means it contains the color of Red, Green, and Blue.

I apply 5×5 convolution layer and ReLu with padding p=0and stride s=2 for out input image to compute the output of neurons. This layer is multiplied by the input image and I decided to use 20 filters.apply the 5×5 convolution layer to get the neurons as output is 159×86×20.

For this output, I apply 3×3 max pool with stride s=2 to compute the output of neurons. The output is 79×43×20. Both conv+Relu and max pool are represented layer one.

For the next step, I apply 3×3 c convolution layer+ ReLu with padding p=0 and stride s=2 and decided the number of filters is 30. Again, this layer is multiplied by the previous activation output and produce the output neurons is 38×20×30.

Again, I apply 3×3 max pool with stride s=2 and number of filters is same as the previous convolution layer. The output of the layer is 19×10×30.

ReLu layer is applied an element wise activation function and the output volume is unchanged.

The fully connected layer is made by the all activations of in the previous layer and computes the class score. The volume of the resulting size is 5700.

I apply softmax function in the fully connected layer to classify the probability of all the outputs and output classes are papaya mosaic disease, powdery mildews and papaya leaf curl disease.

When I train and test my data in our model its takes more time to train my datasets. The computational cost and hardware requirement is high. And the accuracy of my model is 30%.

I want to increase the accuracy of my proposed model that's why I use inception v3 in the last layer in my model. Inception v3 is a pre-trained model by GooLeNet.

We will describe a small portion of inception v3.the module contain a different type of convolution layer: 1×1,3×3,5×5 and 3×3 max pooling . Each convolution working in parallel and concatenating the resulting feature maps before going to the next layer.

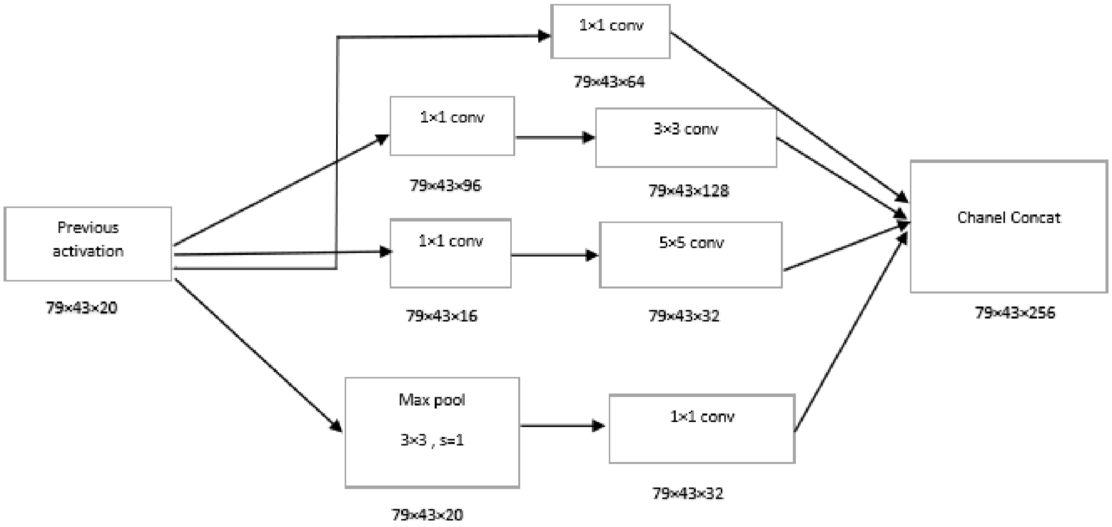


Fig 3.9: Inception-v3 module

1×1 convolution is key because it will be used to reduce the dimensionality of its feature map.out previous activations are 79×43×20.here width: 79, height:43 and number of channel is 20.apply 1×1 convolution with a number of channels 64 and padding p=0 and stride s=1.the output is 79×43×64.

Applying another 1×1 convolution followed by 3×3 convolution with padding p=1 and stride s=1 and number of channels 128.the output is 79×43×128.

Applying another 1×1 convolution followed by 5×5 convolution with padding p=2 and stride s=1 and number of channels 32.the output is 79×43×32.

Apply 3×3 max pooling with stride s = 1 followed by the 1×1 convolution with padding p=0 and stride s=1 and the output is 79×43×32. The channel concatenation is the summation of all the convolution layer and max pool layer is 79×43×256.

Applying inception v3 in our proposed method to increase the accuracy. The accuracy of our proposed method is 87%.

Chapter 4

Experimental Results and Discussion

4. Experimental Results

To measure the performance of my proposed system I use 300 images as test data in ten different datasets to test the accuracy. I dataset contains 1200 images. I applied Cross Validation Technique and portioned our final dataset into 05 equal subsamples to get a higher accuracy. I get the final accuracy 87% which is much higher than we expected. Using these 05 subsample accuracy values I draw a number of iterations versus accuracy graph. I also use the (3*3) confusion matrix to calculate precision, recall, False Positive Rate, f-measure, and accuracy of the model.

The confusion matrix is a table to describe the performance of a classification model on a set of test data. Confusion matrix can define four terms:

True Positive (TP): we predicted result as disease which are actually disease.

True Negative (TN): we predicted result as not disease which are actually not disease.

False Positive (FP): we predicted disease, but these are not actually disease.

False Negative (FN): we predicted not disease, but these are actually disease.

Precision: precision is the piece of related instances among the retrieved instances. High precision means that an algorithm returned substantially more relevant results than irrelevant ones.

$$precision = \frac{tp}{tp + fp}$$

Recall: Recall is the piece of relevant instances that have been retrieved over the total amount of relevant instances. High recall means that an algorithm returned most of the relevant result.

$$Recall = \frac{tp}{tp + fn}$$

F-measure: f-score is a measure of test's accuracy by considering both precision and recall. It is a harmonic average of precision and recall.

$$F - score = 2 * \frac{precision * recall}{precision + recall}$$

Accuracy: accuracy refers to the familiarity of the measured value to a known value.

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

False Positive Rate: Calculate the false positive rate by the given equation:

$$False\ positive\ rate = \frac{FP}{TN+Fp}$$

Table 4: Measure Accuracy based on confusion matrix.

Number of iteration	Recall (%)	False positive rate (%)	Precision (%)	F-measure (%)	Accuracy (%)
1	84	08	84	84	84
2	85	08	85	85	85
3	86	07	86	86	86
4	88	06	88	87	88
5	90	05	89	90	90
Average	87	7	86	86	87

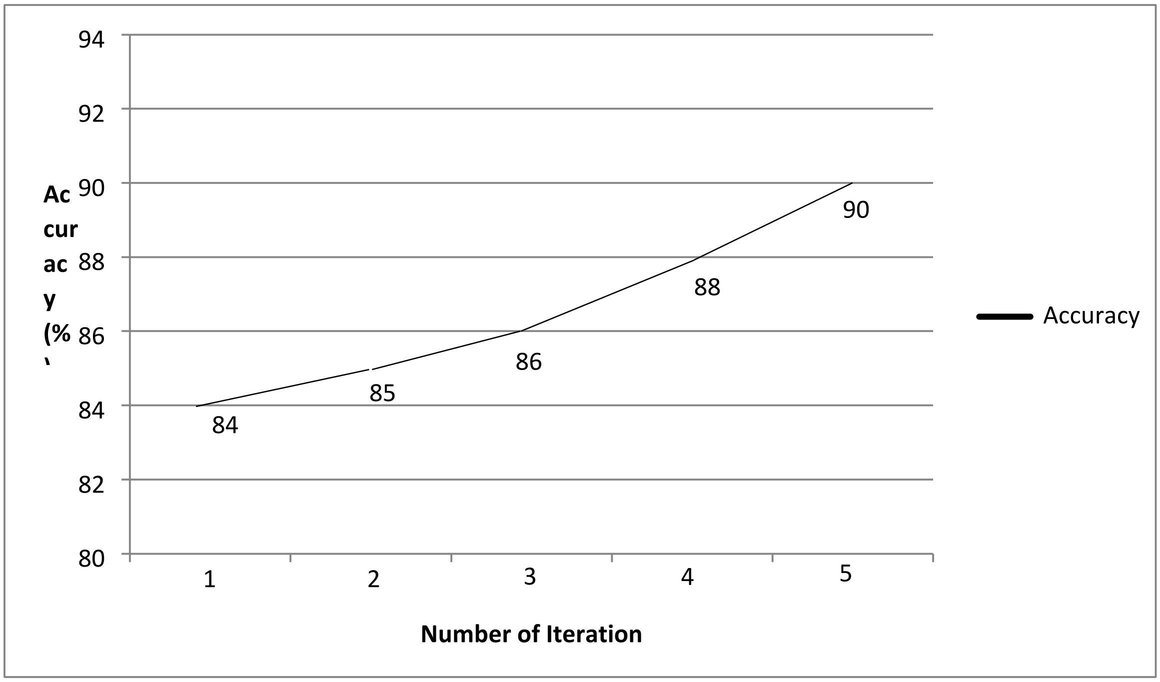


Fig 4: Number of iteration versus Accuracy

Chapter 5

Conclusion and Future works

5.1 Conclusion

In this paper, confusion matrix and neural network is used for segmentation and classification of diseases which affected by different papaya diseases. I measure the probability of an image either it is a disease or not using softmax. Training a Convolutional Neural Network using pre-trained Inception-V3 can show great outcome to classify papaya images. I use 1200 images to train our model and retrained Inception-V3's final layer. Then I test the retrained model using an image which gives the probability of disease or not disease. I used the cross-validation technique in this model and get the accuracy of 87% which is more than expectation. Disobedient variation may be accessible for convinced crops. Discipline weeds, afterward many provide as container for viruses. To devoid growing tobacco mosaic virus, don't pollution over tomato-family vegetation and laundering our hands before stunning them.

5.2 Future Works

I want to proceed updating this system for implementation of the project in real life. The main focus will be to overcome the limitations of the developed system. I also want to work in different stage of diseases. Background cancellation algorithms need to be applied for the image processing to work for any image. In addition, I also want to upgrade this system that can detect any crop disease using only mobile phone application.

REFERENCES

[1] Shanwen Zhang, Zhuhong You, Xiaowei Wu, " Plant disease leaf image segmentation based on superpixel clustering and EM algorithm, " Springer, June 2017.

[2] S. Santhosh Kumar, B. K. Raghavendra, "Diseases Detection of Various Pant Leaf Using Image Processing Techniques: A Review",5th International Conference on Advanced Computing &Communication Systems (ICACCS),2019.

[3] V. Kothari, and S. Seshadri, "Antioxidant activity of seed extracts of Annona squamosa and Carica papaya", *Nutrition & Food Science*, Vol. 40 No. 4, pp. 403-408, 2010.

[4] H.Al-Hiary, S. Bani-Ahmad, M.Reyalat, M.Braik & Z.AlRahamneh, “Fast and Accurate Detection and Classification of Plant Diseases”, International Journal of Computer Applications, Vol.17,No.1, pp.31-38.March 2011.

[5] Rong Zhou, Shunichi Kaneko, Fumio Tanaka, Miyuki Kayamori, Motoshige Shimizu, “Early Detection And Quantization Of Plant Disease Using Template Matching And Vector Machine Algorithms”, First International Symposium On Computing And Networking, 2013.

[6] Malti K. Singh, Subrat Chetia,” Detection and Classification of Plant Leaf Diseases in Image Processing using MATLAB” International Journal of Life Sciences Research ISSN 2348-3148 (online) Vol. 5, Issue 4, pp: (120-124), Month: October - December 2017.

[7] Sanjay B. Patil, Dr. Shrikant K. Bodhe,” Leaf disease Severity Measurement Using Image processing”, International Journal of Engineering and Technology. October 2011.

[8] B. W. Porter, D. A. Christopher and Y. J. Zhu, "Genomics of Papaya Disease Resistance", *Genetics and Genomics of Papaya*, pp. 277-307, 2014.

[9] S. K. Behera, A. K. Rath and P. K. Sethy, "Maturity Status Classification of Papaya Fruits based on Machine Learning and Transfer Learning Approach", *Information Processing in Agriculture*, 2020.

[10] R. K. Veeraballi, M. S. Nagugari, C. S. R. Annavarapu and E. V. Gownipuram, "Deep Learning Based Approach for Classification and Detection of Papaya Leaf Diseases", *International Conference on Intelligent Systems Design and Applications*, pp. 291-302, 2018, December.

[11] I. Hamim, M. Alam, M. Ali and M. Ashrafuzzaman, "Incidence of post-harvest fungal diseases of ripe papaya in Mymensingh", *Journal of the Bangladesh Agricultural University*, vol. 12, no. 1, pp. 25-28, 2014.

[12] J. A. Ventura, H. Costa and J. da Silva Tatagiba, "Papaya diseases and integrated control", Diseases of Fruits and Vegetables: Volume II, pp. 201-268, 2004.

[13] Tarek Habib, Anup Majumdar, A.Z.M. Zakaria, Morium Akter, Mohammad Shorif Uddin and Farruq Ahmed, "Machme vision Based Papaya Diseases Recognition", *Journal of King Saud University- Computer and information Sciences*, vol. 32, no. 3, March 2020.

[14] Shiv Ram Dubey and Anand Singh Jalal, "Application of Image Processing in Fruit and Vegetable Analysis: A Review", J. Intell. Syst, vol. 24, no. 4, pp. 405-424, 2015.

[15] M. Malathi, R. Vishnu Priya and V. Kavidha, "Detection and Classification of Plant Leaf Disease in Agriculture Environment", Asian Journal of Applied Science and Technology (AJAST), vol. 2, no. 1, pp. 212-220, 2018.

[16] Kawaljit kaur and Chetan Marwaha, "Analysis of Diseases in Fruits using Image Processing Techniques", International Conference on Trends in Electronics and Informatics ICEI, 2017.

[17] M. E. Pothen and M. L. Pai, "Detection of Rice Leaf Diseases Using Image Processing", 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), pp. 424-430, March, 2020.

[18] A. Nadeem, T. Mehmood, M. Tahir, S. Khalid and Z. Xiong, "First report of papaya leaf curl disease in Pakistan", Plant Disease, vol. 81, no. 11, pp. 1333-1333, 1997.

[19] R. Ming, S. Hou, Y. Feng et al., "The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus)", Nature, vol. 452, pp. 991-996, 2008.

[20] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic,” Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” Hindawi Publishing Corporation Computational Intelligence and Neuroscience,Vol 2016.

[21] Amar Kumar Dey, Manisha Sharma, M.R.Meshram,” Image Processing Based Leaf Rot Disease, Detection of Betel Vine (Piper BetleL.),” Procedia Computer Science 748 – 754, science direct, 2016.

[22] Vijai Singh, A.K. Misra,” Detection of plant leaf diseases using image segmentation and soft computing Techniques,”Information Processing In Agriculture 4 (2017) 41–49 , science direct, 2017.

[23] Yuanyuan Shao, Guantao Xuan, Yangyan Zhu, Yanling Zhang, Hongxing Peng, Zhongzheng Liu & Jialin Hou,” Research on automatic identification system of tobacco diseases”, vol. 65, no. 4, 252–259, Taylor & Francis, 2017.

[24] Shanwen Zhang, Zhuhong You, Xiaowei Wu,” Plant disease leaf image segmentation based on superpixel clustering and EM algorithm,” Springer, June 2017.

[25] M. Islam, A. Dinh, K. Wahid and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine", 2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE), pp. 1-4, April,2017.

[26] D. Gonsalves, "Control of papaya ringspot virus in papaya: A case study", Annual Review of Phytopathology, vol. 36, pp. 415437, 1998.

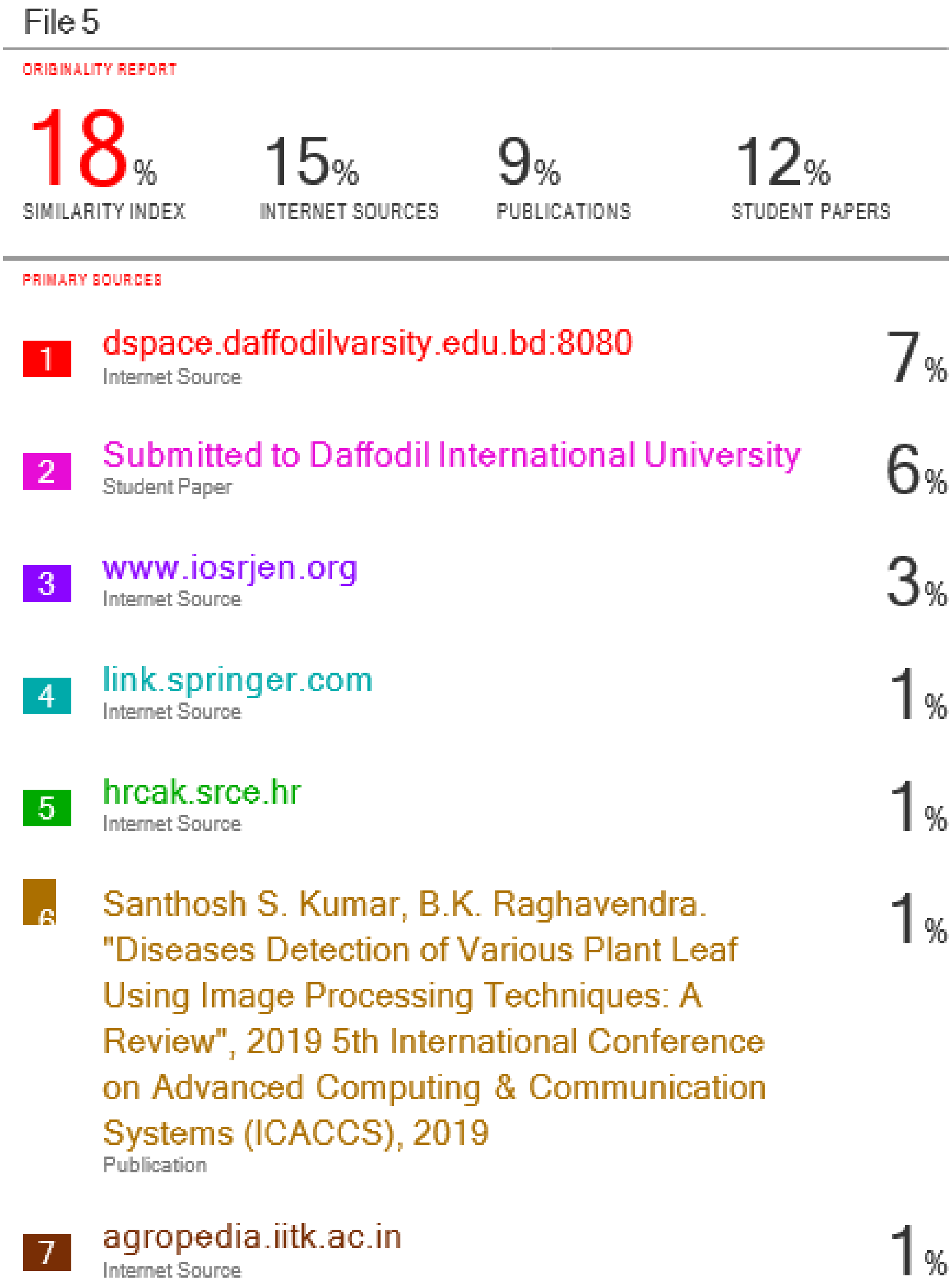
[27] A. Camargo and J. S. Smith, "Image pattern classification for the identification of disease causing agents in plants", Computers and Electronics in Agriculture, vol. 66, no. 2, pp. 121-125, 2009.

[28] Q. Yu, E. Tong, R.L. Skelton et al., "A physical map of the papaya genome with integrated genetic map and genome sequence", BMC Genomics, vol. 10, no. 371, 2009.

[29] S. P. Mohanty, D. P. Hughes and M. Salathé, "Using deep learning for image-based plant disease detection", Front. Plant Sci, vol. 7, pp. 1419, 2016.

[30] S. Tripathi, J. Y. Suzuki, S. A. Ferreira and D. Gonsalves, "Papaya ringspot virusP: characteristics pathogenicity sequence variability and control", Molecular Plant Pathology, vol. 9, pp. 269280, 2008.

Plagiarism



File 5

ORIGINALITY REPORT

18%	15%	9%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	7%
2	Submitted to Daffodil International University Student Paper	6%
3	www.iosrjen.org Internet Source	3%
4	link.springer.com Internet Source	1%
5	hrcak.srce.hr Internet Source	1%
6	Santhosh S. Kumar, B.K. Raghavendra. "Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review", 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), 2019 Publication	1%
7	agropedia.iitk.ac.in Internet Source	1%