APPROVAL

This Project titled "**Disease Detection of Potato Using Deep Learning**", submitted by Abu Bakar Siddique, ID: 142-15-3628 to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **5th May 2018.**

BOARD OF EXAMINERS

____

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Md Zahid Hasan — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science & Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Md. Riazur Rahman, Senior Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree.

Supervised by:

mrrahman

Md. Riazur Rahman
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Abu Bakar Siddique
ID: 142-15-3628
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First of all we express our heartiest thanks and gratefulness to almighty Allah the most merciful, the most beneficent to give me the capability to complete this project report successfully.

We really grateful and wish our profound our indebtedness supervisor **Md. Riazur Rahman, Senior Lecturer,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of “Computer Science” 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.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain, Professor and Department Head,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Computational analysis of plant disease is a challenging and interesting task in now a day. In this paper, I study on how to detect plant disease (potato late-blight) more accurately and give them solutions. Here I concerned about 3 different classes for classification. Several experiments are conducted on these collected dataset and extract feature of the classes. In this research we used deep learning based model (CNN – inception v3) for classification. First of all we train and fine tune our model. Then validate our model according to dataset. After 8000th iteration our classifier is able to classify different class accurately. In our experiment we measured up to 80% for all of the 3 classes. For first class late-blight disease, it gives accuracy level 87%, second class fresh potato detection is 86% and last if there have no potato in image then it also can detected, its accuracy is 87%.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Plant disease detection and analysis consists of inferring the disease of plant by extracting and analyzing the features from the plant leaf or fruits. Every disease has its own symptoms and color combination. Human eyes can easily separate these diseases from any plant leaf or fruit. From the beginning of the digital computing, there have many technique invented to identify disease or disease area. Some technique are really did good job to detect disease area of plant leaf or fruit. But when we come to find out exact disease and try to give them solution then it becomes harder than previous. Deep CNN [7] actually change our thinking and analyzing life. Now a day's almost every harder task can be solved using Deep CNN [7]. Object identification becomes more reliable and accurate. So we decided to solve plant diagnosis problem using Deep CNN [7] to get more accurate result and give farmer exact solution for disease prevention.

1.2 Motivation

Bangladesh is developing country. Many growing sectors are emerging to survive in the new digital world. Agricultural sector also improved more according to digitalization but there has some lacking which are:

1. Potatoes are one of the largest world's cultivated vegetable and Bangladesh is 7^{th} in worldwide potato production.
2. Huge losses because of potato late-blight [18] disease.
3. Lack of farmer's productive education on agriculture.
4. Farmers are not well-known to know their crop's disease.
5. They are not getting instant solution for these diseases.
6. Though internet is available but there has no instant solution guide to suggest them about their problem.
7. There have no mechanism to detect their problem and give them instant solution.
8. Sometimes they faced huge loss for being late to solve disease.

Potato is the popular cultivating vegetable in our country and its disease is causes of lots of loss in our farmer and so I decided to solve these issues digitally.

1.3 Rationale of the study

Plant disease detection is a much harder subject in research area but interesting. In this research solution mainly helps potato farmer to identify potato disease and get primary and permanent solution for disease. In related research there have some device and technique to identify proper disease of pant but the implementation for rural people is hard. Because they have limited education, they don’t know what the current situation of plant is and what have to do. Smartphone is now available for all people. So I think I will make a solution for them which will be mobile based application and real time identification so that they can get escape from huge economical loss.

1.4 Research Question

Q1. Is there any technique in Potato Late-blight Identification?

Q2. How can we identify Potato Late-blight disease?

1.5 Expected Outcome

I have created a mobile app based solution which can identify potato late-blight disease from potato. This application will let them know about crops disease condition and give them step by step solution for disease. Its open research so any developer can use this technique to develop their own system. Now this application is currently implementing identification of one disease but in future we will incorporate more disease into our application so that we can help various kinds of farmer.

1.6 Report Layout

In chapter 1, tried to cover basic concepts of “Disease Detection of Potato Using Deep Learning” and also discuss motivation, rationale of the study, research question and expected outcome of my theses. In chapter 2, focus on related works, brief overview on summary, scope of the problem and the challenges.
In chapter 3, discuss about research methodology
In chapter 4, describes the details of experimental results.
The final chapter 5 I have concluded about my evaluation result and also about some other features that can be included in future works for the betterment of my research work.

CHAPTER 02

BACKGROUND

2.1 Introduction

Potato cultivation is one of the most remunerative farming enterprises in Bangladesh [17]. This vegetable is used for many things like, chips, vegetable etc. So its production is huge. But every year it's losses also huge for late-blight disease. Most of the times farmer cannot find out what disease is attacked in potato, so they need to take go long distance to contact with expert. This process actually needs time to get proper solution and implementation. That's why I thought about real time disease detection so that they can implement proper solution immediately and get escape from losses.

2.2 Related Works

In this section, present related works on Potato Late-blight Disease Detection. There no major work has been done in Bangladesh. But in abroad there have some works based on traditional algorithm and some works on feature or texture based but there having still accuracy and performance issue. Some latest works on CNN[7] and deep learning[6]. So it is difficult for me to do it using deep CNN and make proper train able data. As a beginning of my work I tried to do explore deep CNN [7] for specific image identification. There I find out some approach in Deep learning and some image pre-processing technique which can help me to fulfill my desire. But for more information I get some works based on Neural Network (Grape Leaf) [1] and feature based Neural Network (texture feature based classification)[1] works to find out plant disease. There have some different technique but effective. The problem which I feel the performance issues, I need real time solution for farmer so that they can easily cope up with their problem and get instant solution. So I decide I will simplify this performance problem and try to give them more accurate result. That's why my Deep CNN training face is more important and takes more power to make capable to understand and differentiate exact disease.

2.3 Research Summary

I am trying to represent a new approach in vegetable like potato late-blight disease detection based on Deep Artificial Neural Network (Inception v3 model) [6] and give them real time exact solution. The summary of my whole research is given bellow-

First stage

- Data collection and preprocessing
- Feature Reduction

Second stage

- Feed to Deep CNN
- Use cross-validation for divided data into train and test set.

Third stage

- Train the Deep CNN classifier.

Fourth stage

- Identify late-blight from new potato image.
- Give exact solution for detecting disease.

Last and final stage

- Find the accuracy for each detection.

2.4 Scope of the problem

Plant Disease (like potato vegetable) Detection in Bangladesh perspective is the good research field for me. Because there are no major works have been done for Bangladeshi agricultural sector but having various researches work in abroad based on plant disease. When I started the research work I was faced many problems. The first one was data collection, I had to know about potato disease and how it looks like then had to collect data with effected disease. After data collection I also faced problem to learn about Deep learning and Convolutional Neural Network, how to make its model, its behavior and how to shuffle data to train CNN. First time I cannot understand how actually Deep CNN (inception V3)[13] works, what is the responsibility of weights and biases to find actual result, how hidden layer works to extract image features. So I get confused that's why I cannot get desired results. After then I tried to learn tensorflow [9] ML framework to figure out how CNN works, what is the responsibility of weights and biases and how to make ready my data to cope up with Deep CNN. At last after hard working of couple of months I got my desired solution.

2.5 Challenges

I have faced many problems in thesis work. I always try my best to overcome from those problems. For the purpose of this thesis I have to learn how Deep CNN works. Because classification of exact disease is one of the biggest challenges in image processing problem. On the other hand my main work is to detect actual disease and give them proper solution. In the most related work they used ordinary algorithm like feature extraction or texture feature. Which need more processing power and also having accuracy problem most of the time. So I used Deep CNN and train it. After training validate it using test data set. These tasks was challenging for me.

CHAPTER 03

RESEARCH METHODOLOGY

3.1 Introduction

In this session various steps are considered, Firstly data pre-processing, secondly feature reduction then feed Deep CNN and training it, then potato disease (late-blight) detection. In the first phase, feature reductions are reduce extra things from data using opencv [15] custom function, background subtraction and only potato getting functions. In the second phase train our Deep CNN with shuffled data and validate result.

3.2 Research Subject and Instrumentation

My research subject is “Disease Detection of Potato Using Deep Learning”. It is very interesting and challenging subject for research field. Because accuracy and performance issue is the main fact in this area. Country like Bangladesh its main concern to give them solution with is cheap mean. Though smart phone and internet is now more available then previous, So I decided I will make mobile based application which can give real time solution and give them instant solution. For this project there both mobile app and server side works done. Mobile apps done using JAVA (android) and server side actually did processing and decision task. So server side is in python, there use popular ML [7] library tensorflow [9] and most accurate Deep CNN [7] model Google Inception V3 [16] and for pre-processing task used OpenCV [15].

3.3 Data Collection Procedure

Data collection is the main problem in this image processing task. I was collected the dataset from Vegetable Bazar and website by choosing exact disease image, and collected more than 100 effected images. These files are stored in standard image format (‘png’,’jpg’,’bmp’) in labeled folder. For fresh potato or not potato image, I have made other 2 dataset for better classification of disease image. All of the dataset stored in labeled folder which will be the classification label. Table 3.1 is describing about data collection details.

Table 3.1: Detailed statistics of image collection

Label	No. of image.	Image Format	Image Size
Late-Blight	100	png, jpg	Various
Fresh Potato	100	jpg	Various
Others	1000	png, jpg	Various

As we can see in Table 3.1 we have collected 3 different data. where late-blight 100 different images, same for fresh potato and for other we are collected 1000. why actually we need fresh potato data and others? Answer is simple, we will train our data in Deep CNN model (Inception V3) [13]. So we are teaching algorithms how it can differentiate what is actually this, is it affected potato like late-blight, or its not.

There are several steps in my experiment. I choose 80% and 20% data respectively of training data set and test data set for disease identification.

Table 3.2: Divided the whole dataset in training and testing set

Label	No. of Images	Training set(80%)	Testing set (20%)
Late-Blight	100	80	20
Fresh Potato	100	80	20
Others	1000	800	200

3.4 Statistical Analysis

Statistical analysis is a component of data analytics. Statistical analysis involves collecting and scrutinizing every data sample in a set of items from which samples can be drawn.

3.4.1 Data Pre-processing

In this step, data is sent to pre-processing method, for outsider removal, feature reduction and image data reshape.

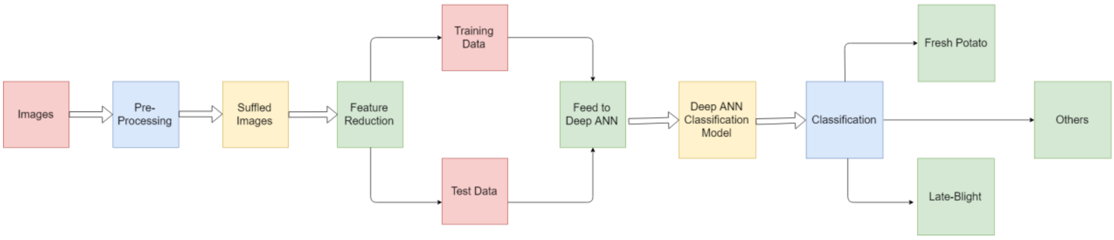


Figure 3.1: System architecture

In terms of garbage removal, I filter dataset to eliminate extra things from image data. Like

Figure 3.2: Outsider removal

3.4.2 Image Data Richness

Image data richness is measurement of diversity of varieties images. Here we have collected 3 types of data. First of all we have collected Potato Late-Blight images with many sides, many faces and many types. These varieties help our classifier to understand the exact face of the late-blight disease. Secondly collected fresh potato images, which is actually fresh. Last one is non potato images so that our classifier can understand which not a potato.

3.4.3 Feature Reduction

Feature reduction is an important state for making good dataset and accurate disease detection. Potato late-blight disease is looking to much odd. So it has similarity with some other view perspective. That's why feature reduction is important to reduce common feature and make disease looking unique view.

3.4.3.1 Reshape Image

Data are collected in different format and contains different shape. To reduce large size image we will reshape all in standard size (220x220). As we know image are converted to a matrix format so if we does not resize this image then it will be a computational issue. Think we have a image size 1920x1080 so if we convert it to matrix then it will be 1920 row and 1080 column which will computationally expensive.

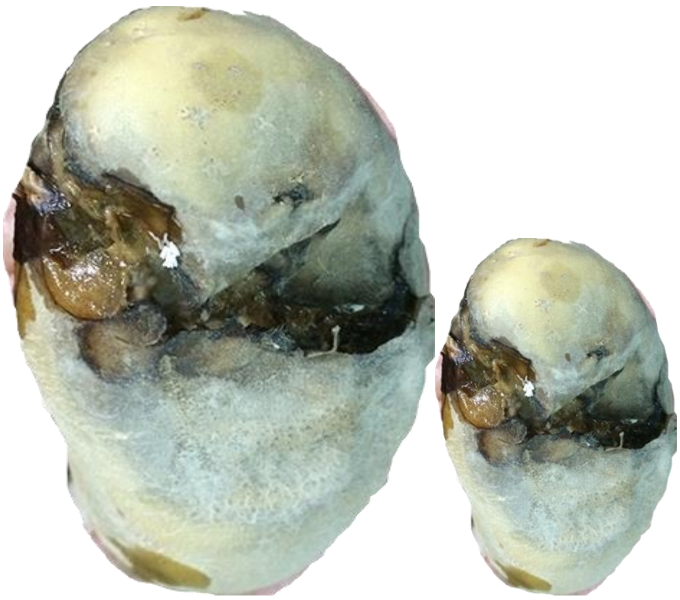


Figure 3.3: Image Reshape

3.4.3.2 Background Subtraction

Classification process is sometimes confusing for computer. Whenever we are trying to feed our classifier with some extra component in data then our classifier will consider these also as a classification area. That's why we will remove background and outsider from our image data.

Figure 3.4: Background Subtraction

3.4.3.3 Noise Reduction

Noise reduction is vital issue for making object sharp and clear identification. So we applied Non-Local Means algorithm to remove a group of pixels surrounding a target noise pixel to smooth the image.

$$u(p) = \frac{1}{C(p)} \int_{\Omega} v(q) f(p,q) dq.$$

Where **u(p)** is the filtered value of the image, **v(q)** is the unfiltered value of the image, **f(p,q)** is the weighting function.

C(p) is a normalizing factor, given by:

$$C(p) = \int_{\Omega} f(p,q) dq.$$

Figure 3.5: Noise Reduction

3.4.3.4 Color Quantization

Color quantization is a process that reduces the number of distinct colors used in an image, usually with the intention that the new image should be as visually similar as possible to the original image. We should go for this approach because large number of variation in RGB value will need more computational power. So we will try to keep unique feature values and reduce common values from our image data.

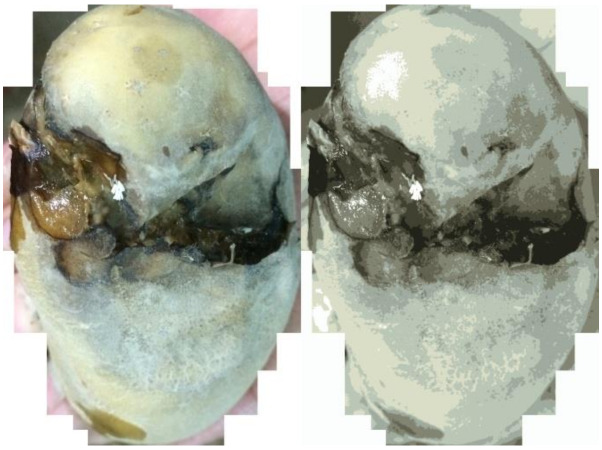


Figure 3.6: Color Quantization

3.4.4 Deep CNN

CNN is now most interesting algorithms for both supervise and unsupervised learning. It's neural network makes relation with classification problem and we don't need to code like our procedural way. We just build our model, set models parameters like activation function, weights, biases and suitable hidden layer.

In this potato late-blight detection problem, we are going to use Deep CNN which is Inception v3 (GoogleNet)[13]. It is so much clever, it can solve identification issues more accurately if we can train these models in better way.

Before feeding to Inception v3, we will talk about how these Deep CNN (Inception v3) works.

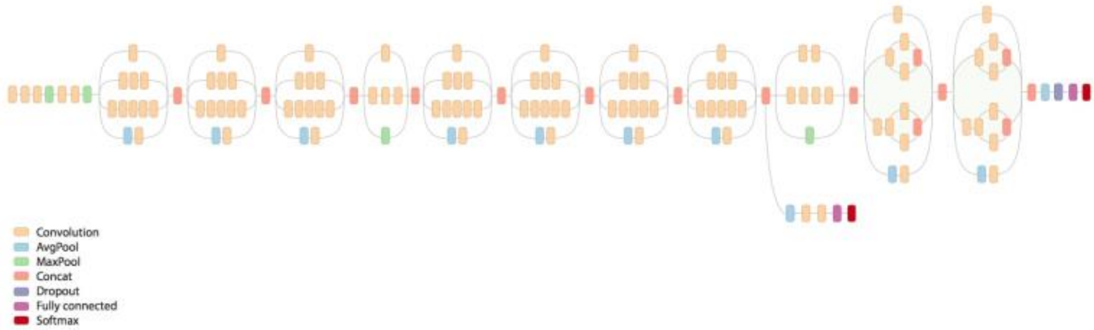


Figure 3.7: Inception v3 architecture [13]

3.4.4.1 Model Architecture

What is the magic behind this Deep CNN inception module [11]. How this actually works. These are the common questions when you heard about this model first time. So now let see it's architecture.

Inception is layered architecture [13]; each layer contains different types of convolution

3×3? Or a 5×5 etc. Each layer each convolution in parallel and concatenating the resulting feature maps before going to the next layer.

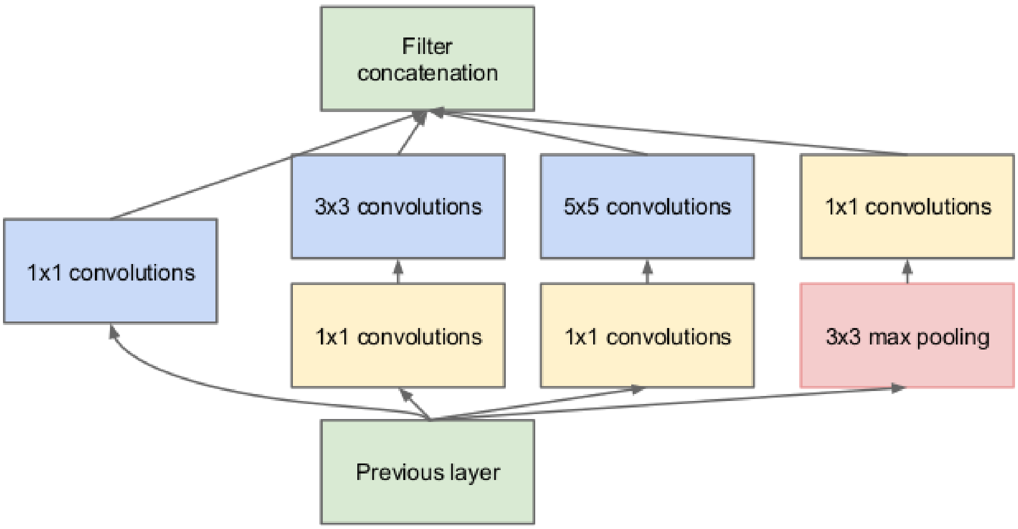


Figure 3.8: Inception module [13]

Notice that here the variety of convolutions. We can use 1×1, 3×3, and 5×5 convolutions along with a 3×3 max pooling. If you're wondering what the max pooling is doing there with all the other convolutions, we've got an answer: pooling is added to the Inception module for no other reason than, historically, good networks having pooling. The larger convolutions are more computationally expensive, so the paper suggests first doing a 1×1 convolution reducing the dimensionality of its feature map, passing the resulting feature map through a relu, and then doing the larger convolution (in this case, 5×5 or 3×3). The 1×1 convolution is key because it will be used to reduce the dimensionality of its feature map.

3.4.4.2 Dimensionality reduction

This was the coolest part of the paper. The authors say that you can use 1×1 convolutions to reduce the dimensionality of your input to large convolutions, thus keeping your computations reasonable. To understand what they are talking about, let's first see why we are in some computational trouble without the reductions.

Let's say we use, we the authors call, the naive implementation of an Inception module.

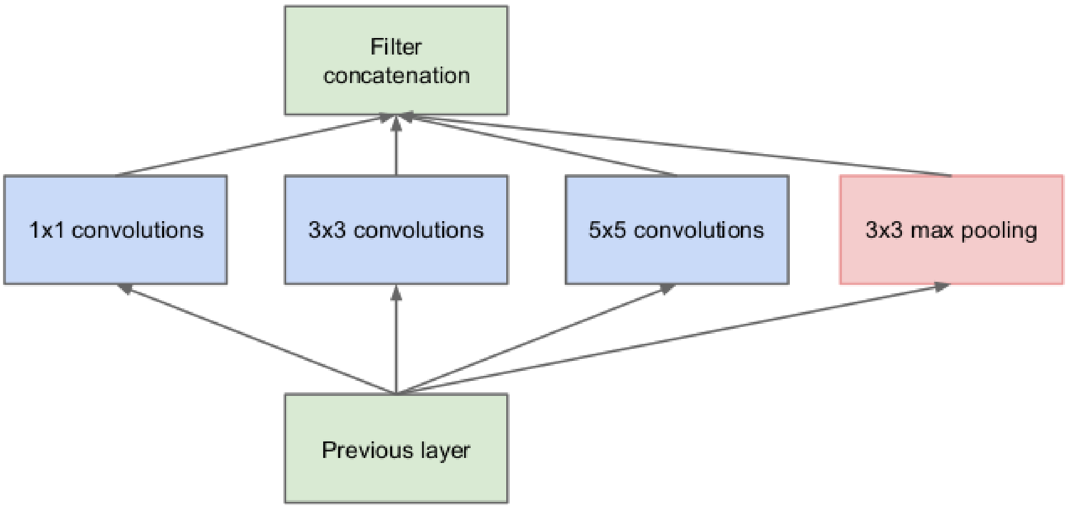


Figure 3.9: Naive Inception module

Figure 3.9 shows an Inception module that's similar to the one in Figure 3, but it doesn't have the additional 1×1 convolutional layers before the large convolutions (3×3 and 5×5 convolutions are considered large).

Let's examine the number of computations required of the first Inception module of GoogLeNet. The architecture for this model is tabulated in Figure 5.

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

Figure 3.10: GoogLeNet architecture with the previous layer dimensions and the Inception module of interest highlighted in green and red, respectively

We can tell that the net uses *same* padding for the convolutions inside the module, because the input and output are both 28×28. Let's just examine what the 5×5 convolution would be computationally if we didn't do the dimensionality reduction. Figure 6. pictorially shows these operations.

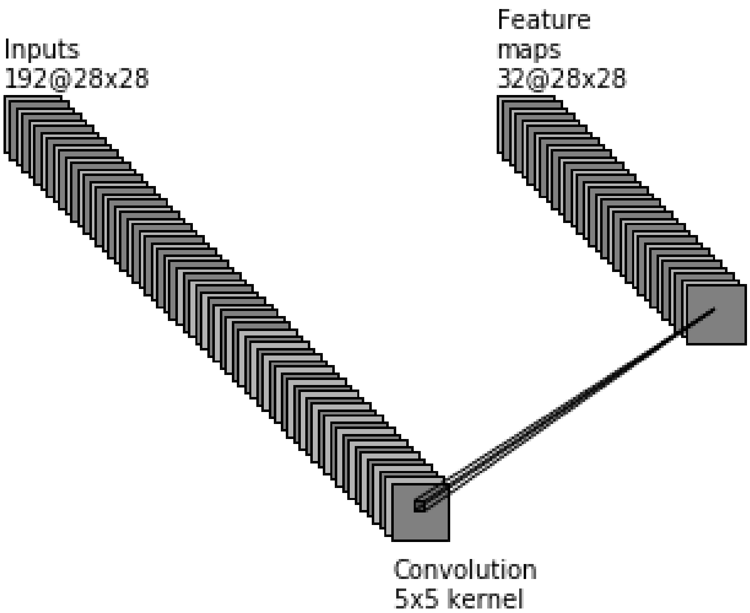


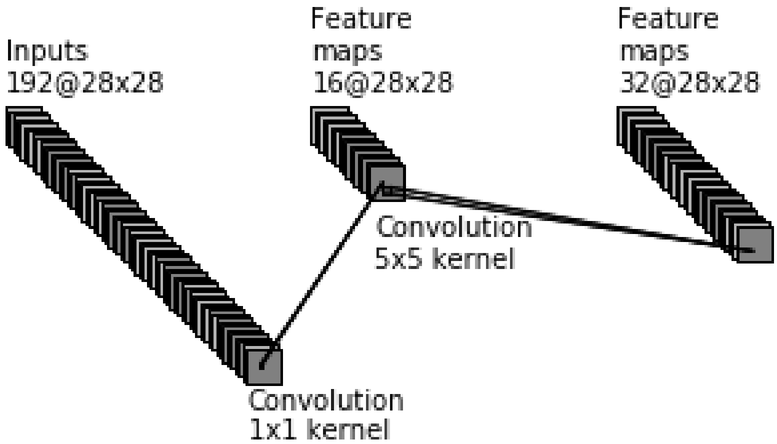
Figure 3.11: 5×5 convolutions inside the Inception module using the naive model

There would be

$5^2(28)^2(192)(32) = 120,422,400$ operations.

Wow. That's a lot of computing! You can see why people might want to do something to bring this number down.

To do this, we will ditch the naive model shown in Figure 4 and use the model from Figure 3. For our 5×5 convolution, we put the previous layer through a 1×1 convolution that outputs a 16 28×28 feature maps (we know there are 16 from the #5×5 **reduce** column in Figure 3.10), then we do the 5×5 convolutions on those feature maps which outputs 32 28×28 feature maps.



In this case, there would be

$[(1^2)(28^2)(192)(16)]+[(5^2)(28^2)(16)(32)] = 2,408,448+10,035,200 = 12,443,648$ operations.

Although this is a still a pretty big number, we shrunk the number of computation from the naive model by a factor of ten.

We won't run through the calculations for the 3×3 convolutions, but they follow the same process as the 5×5 convolutions. Hopefully, this section cleared up why the 1×1 convolutions are necessary before large convolutions!

3.4.5 Feed to Deep CNN

Now it's time to enter into our action. We have to feed our deep CNN. As we discussed above, all things are for preparing our data to feed deep CNN (Inception v3) [12] model. Our model graph is looks like

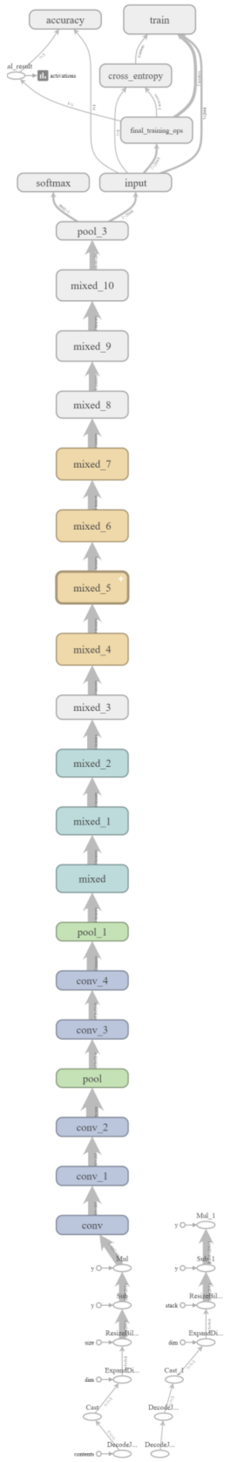


Figure 3.12: Our CNN model graph

Here every step is related with some hidden layer. Each steps process its data and result should be the input data for next steps layers.

To training our model, we have prepared our dataset with large number of verities image.

Table 3.3: Collection of dataset

Label	No. of image.	Image Format	Image Size
Late-Blight	100	png,jpg	Various
Fresh Potato	100	jpg	Various
Others	1000	png, jpg	Various

To improve our results we will alter the details of the learning process. We were selected 8000 training steps to get better accuracy.

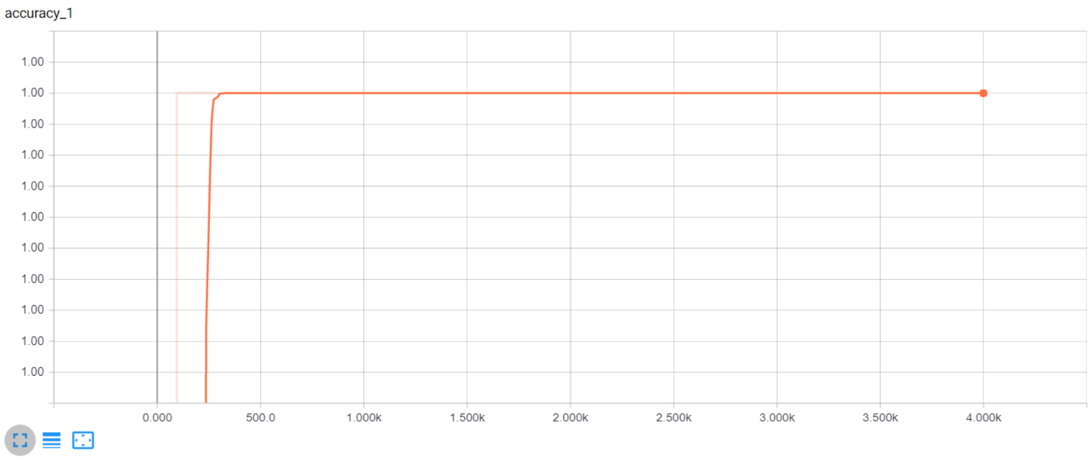


Figure 3.13: Training Accuracy

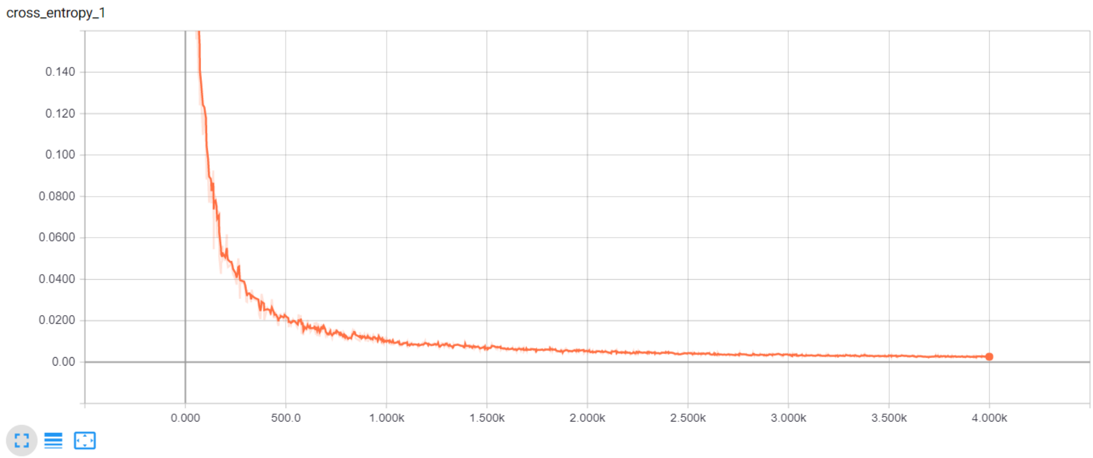


Figure 3.14: Training Cross Entropy

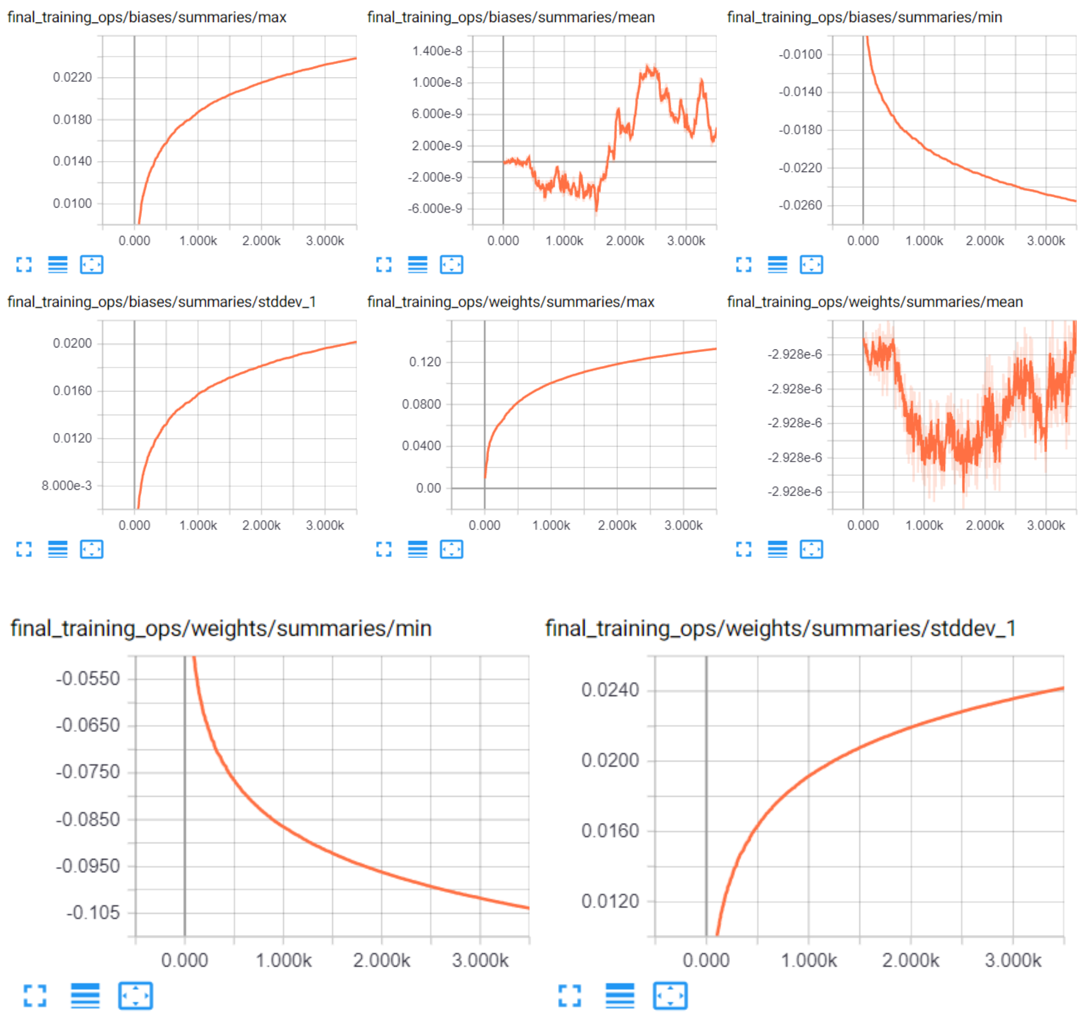


Figure 3.15: Training Parameters

3.4.6 Cross-Validation

After training process has completed, we need to evaluate our classifier. This evaluation is completed by validating trained model. According to Inception model, validation unease to improve accuracy level and choose right weight, biases values. Here some validation reports of different types.

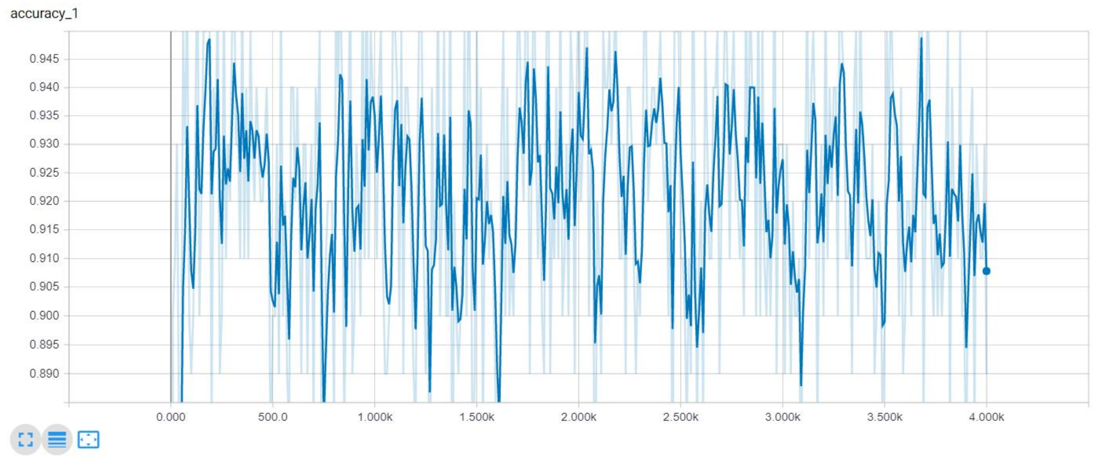


Figure 3.16: Validation Accuracy

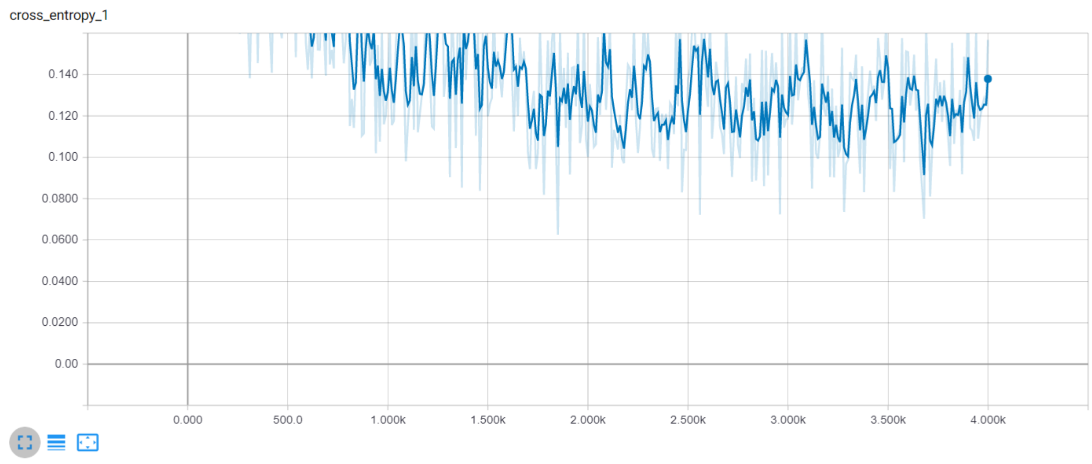
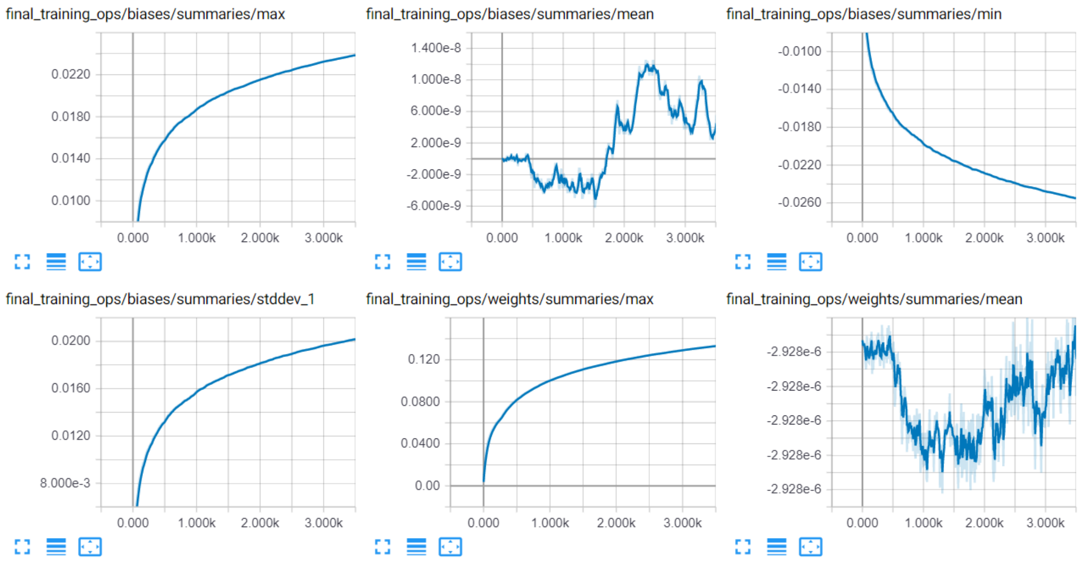


Figure 3.17: Validation Cross Entropy



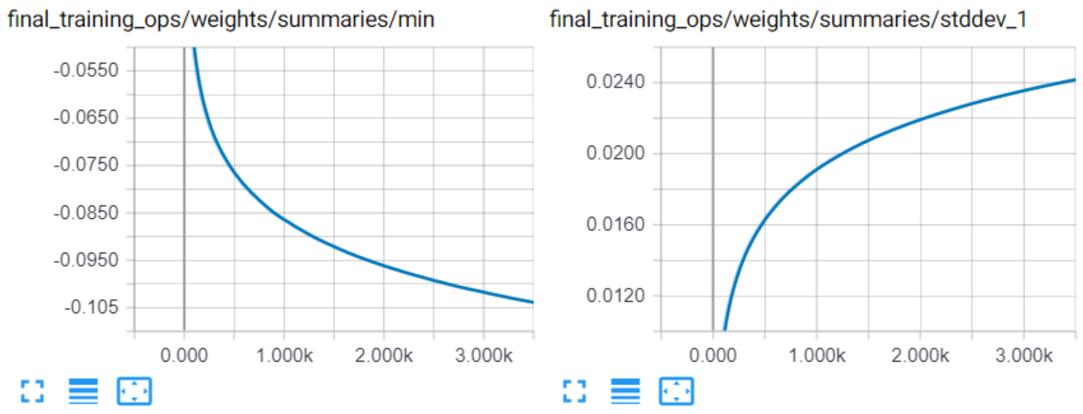


Figure 3.18: Validation Parameters

3.4.7 Deep CNN Classifier

Deep CNN is much smarter then we think. During training our deep CNN model, it started to transfer its learning according to our dataset. Every layer has their own feature extraction technique. Actually main work is done from last layer. It fine-tune all of the operators which can use to classify object more accurately.

3.4.8 Pipelining

A pipeline does exactly what it sounds like: connects a series of steps into one object which you train and then use to make predictions. In short, we can use a pipeline to merge the feature extraction and classification into one operation.

3.5 Implementation Requirements

Hardware Requirements (Mobile version & server)

- Minimum Quad Core processor and hardware.
- 2GB of RAM or higher for mobile and 4GB or higher for server
- 2.0GHZ processor or above

Software Requirements

- Operating system windows/linux for server and android for mobile

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

All of my experiments have been done on the basis of test data. So from the first stage of my experiment our deep CNN [11] model was trained so that it can learn features of our disease from training data set.

After fine-tuning our model parameters, our classifier will able to classify disease more accurately. After training of 8000th iteration our model get overall 97.8% accuracy. These improvement of accuracy based on the number of iteration we are executing to train our classifier. Here the accuracy of our training set.

All the experiments have been done on the test data .In the first stage of my experiment unigram word and vocabulary richness is taken as feature set. In the second stage several features from character and lexical feature are used for my experiment.

4.2 Experimental Results

Most of the time same disease have different face. So feature identification sometimes confusing for classifier. From digitalization of the computer there have many techniques were invented to handle detection issues more accurately.

Artificial Neural Network is best algorithm to handle detection related problem more easily. So in our deep CNN [7], we trained our model to classify 3 classes. First class is, Potato Late-Blight, second class is Fresh Potato and last is others. Our deep CNN there have several faces to extract feature from image dataset and runs over many times to get better result. In the mean time of training, fine-tuning is start.

After fine-tuning our model parameters, our classifier will able to classify disease more accurately. After training of 4000th iteration our model get overall 87.8% accuracy. These improvement of accuracy based on the number of iteration we are executing to train our classifier. Here the accuracy of our training set.

Table 41: Accuracy percentage with different classes

Class	Accuracy (%)
Late-Blight	0.87
Fresh Potato	0.86
Others	0.87

Details histogram about training accuracy achievement is shown.

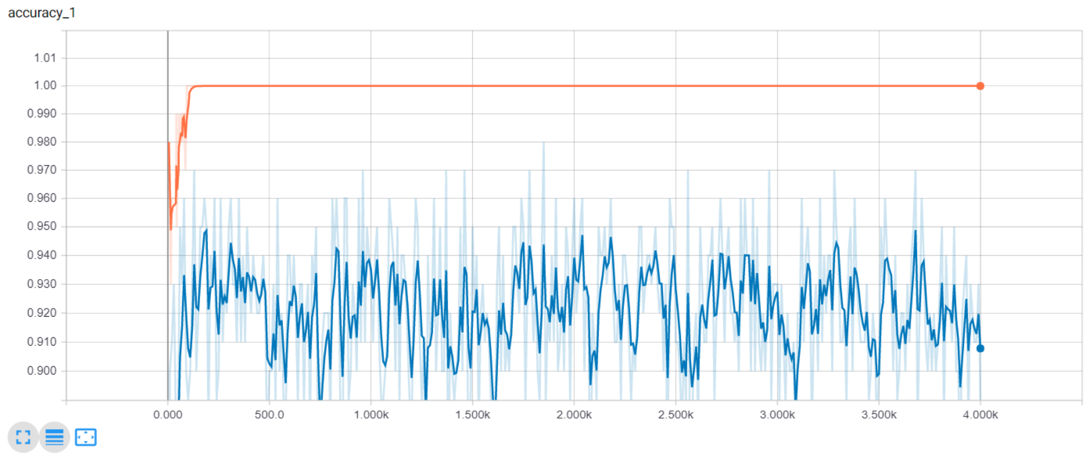


Figure 4.1: Accuracy of training and validation.

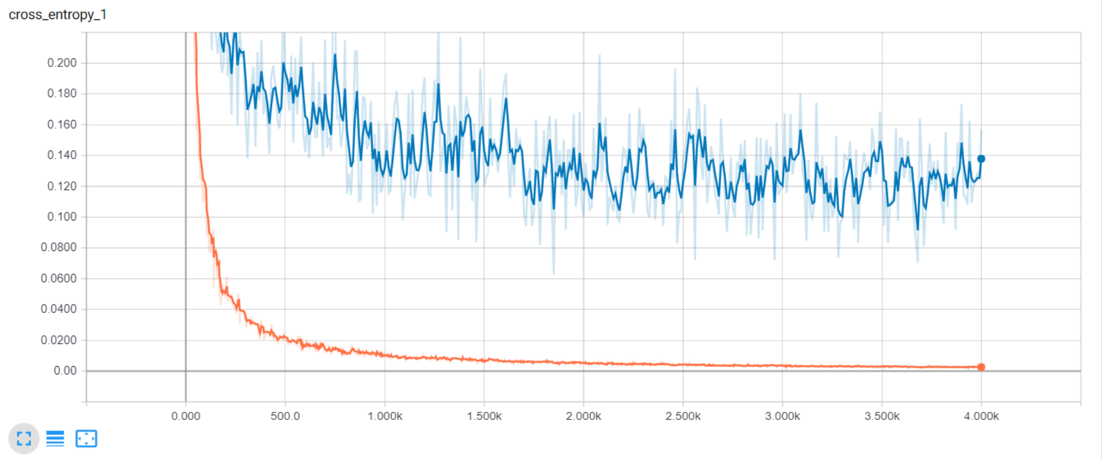


Figure 4.2: Cross Entropy of training and validation.

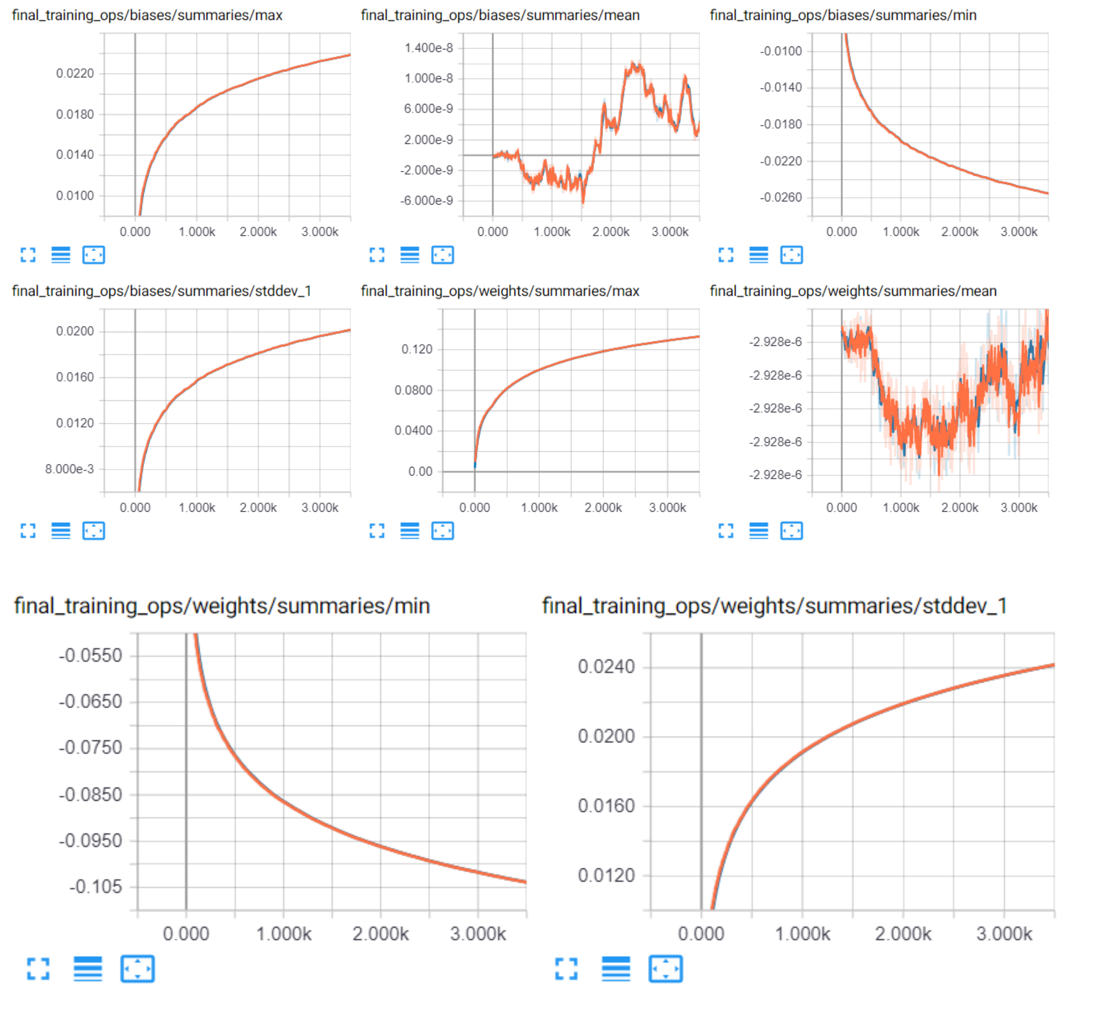


Figure 4.3: Model Operators Training and Validation.

4.3 Descriptive Analysis

As you know that our experiments have been done based on the test data. So first stage of the experiment there we used 80% data for training set and 20% data is used for test set. You already heard about our deep CNN classifier [11], it can now classify 3 classes almost accurately.

In table 4.1 there you can see the percentage of the correct classification accuracy results. Now we already know about how our model classify these disease. As you can see figure 4.1 there the graph shows the training and validation accuracy level according to iteration. You can see that training accuracy bar (orange color) is become almost stable after 250-300 training iteration but in validation process, the frequency of accuracy value is changing. But why this is not getting stability ? Answer is pretty simple, as we already knows about the working process of deep CNN and its layers. In validation process, our model parameter selection works that's why the value is frequently changing over iteration.

In figure 4.2 shows the cross entropy loss function which is used to reduce the loss of the parameter. Cross entropy simply calculate the loss from input dataset and output label. In this figure we can see that the training entropy is going down after 400-500 iteration and becomes lower after 4000th iteration. But in validation side this line is to much distorted as like as accuracy graph.

If we take a look in the figure 4.3, we can see 8 graph beside there. Each graph represent different operators value. First 1-4 graph is for biases, first one is showing the max values, second one is mean values, third one is showing min values, and fourth one is showing standard deviation. Last 4 one is for weight operators. There we can look same max, mean, min and standard deviation. Whatever, as we can see we cannot differentiate validation line and training line from these graph because both of the line almost same.

4.4 Summary

In our experiment we are try to solve plant disease detection problem more accurately using deep convolutional neural network (GoogleNet inception v3 model) [13]. After fine tuning and transferring the learning we achieved almost 85% - 87% accuracy over 3 classes. There late-blight detection accuracy is near about 87%, fresh potato detection accuracy 86% and our last class if there have no potato image then it's identification accuracy is 87%.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

I am trying to represent a deep learning approach in plant (potato late-blight) disease detection problem using convolutional neural network model [7]. The short summary of my whole research is given bellow-

I. First stage
   - Data collection and pre-processing
   - Feature reduction

II. Second stage
   - Transfer learning.
   - Use 10-fold cross-validation for divided data into train and test set.

III. Third stage
   - Classify the late-blight disease using deep CNN.

IV. Last and final stage
   - Find the accuracy and entropy to get better results.

Fine tuning the model operators to get better accuracy and solution.

5.2 Conclusions

In this work, disease detection has been experimented on the potato late-blight disease but we are trained our model with 3 classes. So our model can identify 3 types, whether first label is Late-Blight, second label is Fresh Potato and last label is others.

The classifier which is implemented in our experiment is Deep Learning which is actually a Convolutional Neural Network. It has different hidden layer. Each hidden layer extracts different information from image data and improves biases and weights to make classification more accurate. In my experiment Deep Learning based classifier shows the following remarkable points:

- ✓ The mid stage of late-blight effected image can be detected more accurately than earlier stage. Its accuracy almost 74%.
- ✓ Fully late blight effected image accuracy is more higher than mid level. It’s best accuracy recorded about 93%.
- ✓ It can easily detect any fresh potato with higher confidence.
- ✓ For other things it can say easily it’s not a potato.

5.3 Recommendations for Further Study References Appendices

In this experiment I show the Late-blight disease detection from Potato. I was used midsized dataset and transfer learning to our CNN classifier model (Inception v3). Presently I was selected 3 classes to classify, but in future I will retrain our model with multiple diseases so that it can classify more disease at the same time.

REFERENCES

[1] Sanjeev S Sannakki, Vijay S Rajpurohit, V B Nargund, Pallavi Kulkarni.: Diagnosis And Classification Of Grape Leaf Diseases Using Neural Network. IEEE - 31661

[2] Jagadeesh. D. Pujari, Rajesh. Yakkundimath and A. S. Byadgi.: Reduce Color and Texture features based Identification and Classification of Affected and Normal fruites image.

[3] Jayamala K. Patil , Raj Kumar.: ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES. ISSN 0976-2604 Vol 2, Issue 2, June-2011, pp 135-141

[4] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. DeepnNeural Networks Based Recognitaion of Plant Disease by Leaf Image Classification. Volume 2016 (2016), ArticalID 3289801, 11 pages http://dx.doi.org/10.1155/2016/3289801

[5] Yu Sun and Jianxin Wang. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. Volume 2017, Article ID 2917536, 8 pages https://doi.org/10.1155/2017/2917536

[6] Federico Martinelli, Riccardo Scalenghe, Salvatore Davino, Stefano Panno, Giuseppe Scuderi, Paolo Ruisi, Paolo Villa, Daniela Stroppiana, Mirco Boschetti, Luiz R. Goulart, Cristina E. Davis, Abhaya M. Dandekar. Advanced Method of Plant Disease Detection. A review. https://doi.org/10.1007/s13593-014-0246-1

[7] Deep learning, available at << https://en.wikipedia.org/wiki/Deep_learning/>>, last accessed on 21-03-2018 at 6:00pm.

[8] CNN, available at <<https://en.wikipedia.org/wiki/Convolutional_neural_network />>, last accessed on 22-03-2018 at 10:00PM.

[9] Learn about tensorflow, available at<<https://www.tensorflow.org/>>, last accessed on 23-04-2018 at 10:00AM.

[10] Learn about tensorboard and tensorflow graph data visualization, available at <<https://edwardlib.org/tutorials/tensorboard/>>, last accessed on 25-04-2018 at 10:30AM

[11] Building a Convolutional Neural Network using Tensorflow Library, available at <https://www.tensorflow.org/tutorials/layers/>>, last accessed on 10-03-2018 at 11:00PM

[12] Object Recognition with Google's Convolutional Neural Networks, available at <<https://medium.com/@williamkoehrsen/object-recognition-with-googles-convolutional-neural-networks-2fe65657ff90/>>, last accessed on 25-03-2018 at 4.00 PM.

[13] Google's Inception modules: explained, architecture and implemented, available at <<https://hacktilldawn.com /2016/09/25/inception-modules-explained-and-implemented// >> last accessed on 20-03-2018 at 4.00 PM.

[14] About cross-validation, available at blog << http://www.zacstewart.com/>>,last accessed on 20-03-2018 at 4.00 PM.

[15] OpenCV, available at <<https://opencv.org/>>, last accessed 26-03-2018 at 10 AM

[16] Learn more things from Wikipedia site, available at <<http://www.wikipedia.org/>>, last accessed on 27-03-2018 at 12:00PM

[17] Potato production statistics, available at << https://www.potatopro.com/world/potato-statistics/>>, last accessed on 29-03-2018 at 11:00AM

[18] Estimated worldwide loss due to potato late blight disease, available at << https://research.cip.cgiar.org/confluence/display/GILBWEB/Social+Impact+and+Economic+Importance+of+Late+Blight/>>, last accessed on 29-03-2018 at 12:00PM

APPENDICES

Appendix A: Research Reflection

During my research I was faced many problems. I had to learn new ML and deep learning techniques and Convolution Neural Networks. After then I had to learn how to train or retrain model. After completing research I have gain depth knowledge and skill about Neural Network, how its hidden layer works, server side implementation of these processing and android application. This knowledge will bring me a better future.

Appendix B: Related Issues

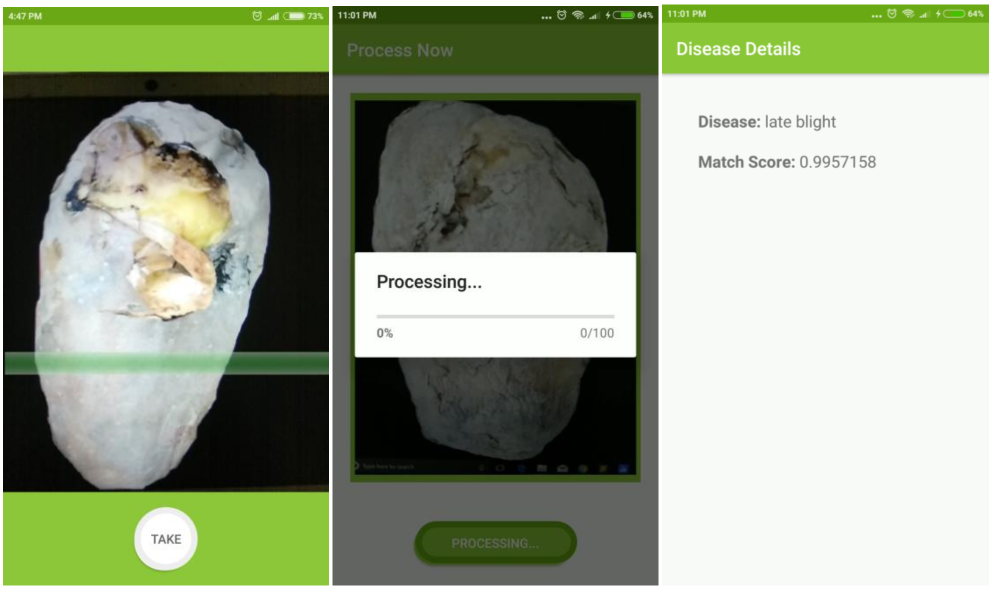


Figure 5.1: Late-Blight Detection test

Figure 5.1 shows the Late-Blight detection test. I use a test effected late-blight potato image and after scanning our classifier successfully detect it. Here accuracy gain 99.5%

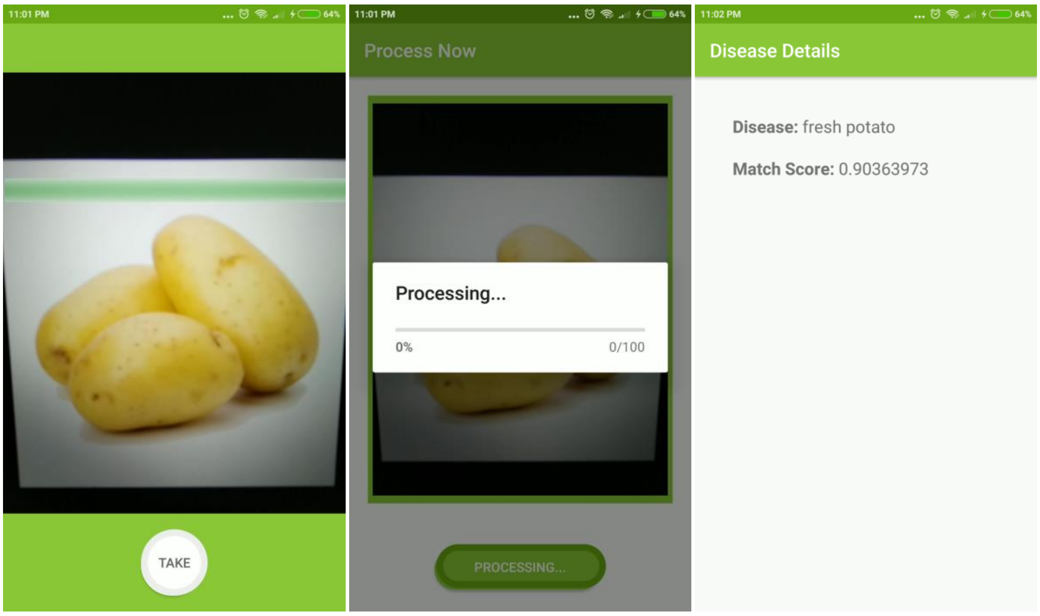


Figure 5.2: Fresh Potato detection test

Figure 5.2 shows the Fresh Potato classification test. Here I used fresh potato image and our classifier successfully classify it as a fresh potato. Accuracy is 90%

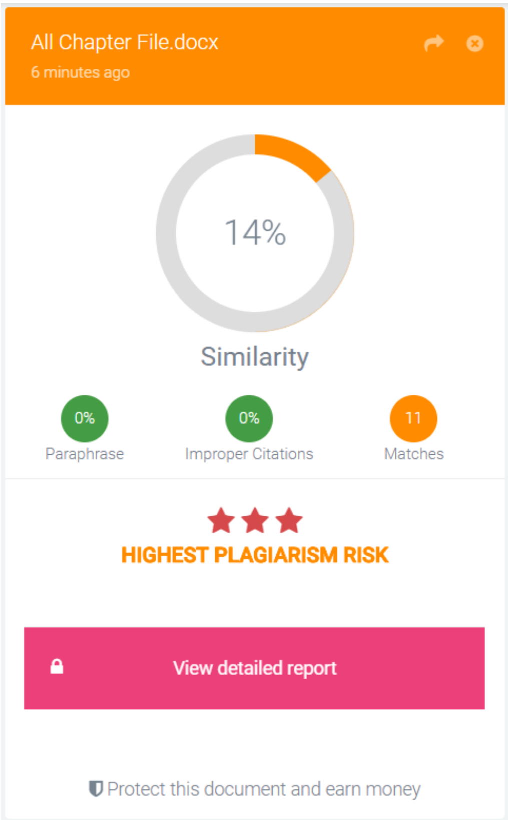


Figure 5.3: Plagiarism check