

TOMATO'S DISEASE DETECTION BY DEEP LEARNING USING CNN

BY

SAJIB HOWLADER

ID: 191-15-12953

AND

MAHMODUL HASAN

ID: 191-15-12933

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

Supervised By

Md. Firoz Hasan

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Fariha Jahan

Lecturer

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 28, 2023

APPROVAL

This Project/internship titled “**Tomato’s Disease Detection by Deep Learning using CNN**”, submitted by Sajib Howlader and Mahmodul Hasan, ID No: 191-15-12953 and 191-15-12933 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 B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 28/01/2023.

BOARD OF EXAMINERS

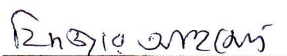
Chairman


Dr. Touhid Bhuiyan

Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Dr. Fizar Ahmed

Associate Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Taslima Ferdous Shuva

Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner



Dr. Md Sazzadur Rahman


Associate Professor

Institute of Information Technology
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Md. Firoz Hasan, 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 or diploma.

Supervised by:

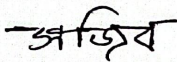
 28.01.2023

Md. Firoz Hasan
Lecturer
Department of CSE
Daffodil International University

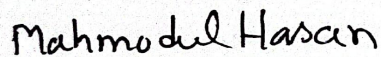
Co-Supervised by:

Fariha Jahan
Lecturer
Department of CSE
Daffodil International University

Submitted by:



Sajib Howlader
ID: - 191-15-12953
Department of CSE
Daffodil International University



Mahmodul Hasan
ID: - 191-15-12933
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Md. Firoz Hasan**, Lecturer, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Machine Learning and Deep Learning*” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Prof. Dr. Touhid Bhuiyan, our 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

Leaf disease have been one of the riskiest impedences to food society, because it cut down the crops yield and compromise the quality. Tomato is one of the most demandable vegetables in Bangladesh. Due to the insect and disease of leaf the yield falls down and cultivating system going through an unclear future. Most of the farmers and agricultural farm are going to off their agricultural profession. It hearts us most. We decided to build a model that can detect the disease earlier using the affected leaf image by CNN with 87.5% accuracy which is cleat how accurate CNN work for image classification. We took 1359 data and 80% use for training and 20% data for testing. The model has not needed any human supervisor for using its any features. The CNN model takes its decision with Convolutional layer, Maxpooling layer, Flatten layer, Dense layer. Difference layer has different work and by a fully supervision all the layer the decision has been taken. This detection includes some phase like image acquirements, pre-processing of the image, data segmentation, classification and feature education.

TABLE OF CONTENTS

CONTENT	PAGE
Board of Examiners	i
Declaration	ii
Acknowledgement	iii
Abstract	iv
List of Figure	vii
List of Tables	vii
CHAPTER 1: Introduction	1
1.1 Introduction	1
1.2 Motivation	3
1.3 Rational of the Study	4
1.4 Research Questions	5
1.5 Expected Output	7
1.6 Project Management and Finance	8
1.7 Report Layout	10
Chapter 2: Background	13
2.1 Preliminaries	12
2.2 Related Works	14
2.3 Comparative Analysis and Summary	15
2.4 Scope of the Problem	16
2.5 Challenges	17
Chapter 3: Research Methodology	19
3.1 Research Subject and Instrumentation	19
3.2 Data Collection Procedure/Dataset Utilized	20
3.3 Statistical Analysis	22
3.3 Proposed Methodology/Applied Mechanism	24
3.4 Implementation Requirements	29

Chapter 4: Experimental Results and Discussion	30
4.1 Experimental Setup	30
4.2 Experimental Results & Analysis	31
4.3 Discussion	38
Chapter 5: Impact on Society, Environment and Sustainability	41
5.1 Impact on Society	41
5.2 Impact on Environment	42
5.3 Ethical Aspects	42
5.4 Sustainability Plan	44
Chapter 6: Summary, Conclusion, Recommendation and Implication for Future Research	46
6.1 Summary of the Study	46
6.2 Conclusions	46
6.3 Implication for Further Study	47
References	49

LIST OF FIGURES

FIGURE	PAGE
Figure 1: The accuracy of the model.	7
Figure 3.1: Disease images.	21
Figure 3.2: Accuracy and Error graph.	23
Figure 3.3: Model Architecture.	24
Figure 3.4: Plot Model.	28
Figure 4.1: ANN visualization.	31
Figure 4.2: Accuracy and Loss graph.	33
Figure 4.3: Confusion Matrix of CNN.	35
Figure 4.4: Logistic Regression.	36
Figure 4.5: Support Vector Machine.	36
Figure 4.6: K-Nearest Neighbor.	37
Figure 4.7: Decision Tree.	37
Figure 4.8: Naïve Bayes.	38

LIST OF THE TABLE

Tables	PAGE
Table 3.1: Nine disease leaf and a healthy leaf.	30
Table 4.1: Model's Information	31
Table 4.2: Different model accuracy.	33

CHAPTER 1

INTRODUCTION

1.1 Introduction

Plant disease is one of the major factors for agricultural sector. In every year Bangladesh losing huge percentage of crops just for disease. According to the FAO (April 3, 2019), 20% - 40% of global product damage due to pests. For this reason, the whole world going to an unclear economy. One of the main reasons for famine is crops disease and damage. It can be reduced by proper treatment. And it would be better to detect the disease before affect. There are several technics are implemented to solve the problem. Image classification is one of them. Bangladesh has six season and it has seasonal vegetables. Tomato is one of them. Tomato's diseases detection is one of the most challenging work. The most nine disease are identified, for that the tomato's yield are going down. Prevention of one of the most effective and useful technics. Here, we implement a model that can detect the tomato's disease earlier and can prevent the damage. Machine Learning plays a vital role to solve such kinds of problems. In additionally, deep learning is very essential for this specific area. With the support of machine learning computer vision have acquire a tremendous success of its field. In this achievement leads to implement new model which can be consider as Deep Learning (DL) technics. Deep Learning approach display the agricultural domain and it has a hug acceptance for its characteristics. Researcher have newly invented Convolutional Neural Network to get a solution for image recognition with the help of image. Advance Artificial Network make it possible to choose automatic detect and take new feature along with disease.

Convolutional Neural Network (CNN) is of the most useful and effective algorithm for image classification. It hast some techniques and different layer what can sense the disease if a highest number of accuracies. CNN is one type of Deep neural

Network (DNN). Traditional compute vision requires for selection manually of its feature. In contrast, Convolutional Neural Network learn automatically its feature of multiple processing.

To build a model, training, testing, validation if a great factor in this segment. There are some limitations in this area that are: -

1. Lack of farmer's training.
2. Digital Agricultural system.
3. Bandwidth.
4. Farmer's education.
5. Proper Guideline from local agricultural institute.

To create a digital agricultural system and to get a good yield the limitations need to solve. This is not a huge problem and its solvable. The deep learning techniques is so popular for its performance and its accuracy. There are many approach the has applied such as Kent, AlexNet, VGGNet, GoogLeNet. But the numerous challenges are still hidden for the correct classification and the exact salutation. To get a highest accuracy and an easy model what can use by farmers our model can be a salutation. Convolutional Neural Network vastly uses for the image classification of its easier technics. The input layer, hidden layer and output layer are purified the image in different technics and provide the solutions of the model display. The first layer convolutes the data. During the operation it require input data, filtering and feature map. The rest of the layer will work for classification, pre-processing and assuming with the 3D matrix along with different features. During the import feature CNN need not any human direction. The hidden layer is used to increase the accuracy of model. The number of neurons is equal to subset of the data set.

1.2 Motivation

Behind this research, we have a motivation history. To see the huge amount of tomato's yield, fall down and farmers are facing lose. We get motivation from this segment. As we know most of the farmers are poor, that's why we build a model that would be useable by the farmers and agricultural farm. Bangladesh produce it require vegetables itself. If our model use over the country then Bangladesh will be benefited and Bangladesh will be able to export its local crops and can earn foreign currency. The economy and GDP will increase along with international business field will be expand. To think overall of the point, we have started this project and research. It can be seen that many farmers take loan for articulate product from local bank, epistyle, agency and others source. They also take usury with the highest number of percentages. They think that they by selling their agricultural product they will repay their loan or interest. But when their production is not expectable and crops are destroyed due to insect, they become proletarians unable to bear the burden of deb. They are forces to leave their homes. Then they lead a non-human life. It hearts us most and one of the top reasons to build this model. We also focus no to be complex and the prices also. We try to make as reasonable that farmers would be able to buy it. During the pre-checking and pre-plan, we have found may model on online and have found many papers what gave the salutations. But we notice all the possible prices and the model. We think there is not a single model what is exactly perfect and farmers friendly what can solve the specific problem. So, we take to solve out this problem. On the other side, government are losing its GDP, if it got the government then the economic conditions would be strong. We got motivated when we fixed the Convolutional Neural Network (CNN). As, we know CNN vastly use over the world to such types of problem. And Convolutional Neural network will give us the highest accuracy. It easy to conduct and self-selected feature. During the operation the model faces many technical and non-

technical problems but it would be overcome by the management. The more problems we solved we our self-getting confidence and we our self-motivated. One of the biggest things is our team were committed to gift the nation a such accurate and easier model. Makes the model thinking that we can put a smile on the farmers face. At last we have done it with a pleased and satisfied accuracy. Now this time our model is ready to solve the problem and help farmers, people, government and the world.

1.3 Rationale of the study

It was so needful study about Tomato's disease detection. Prevention is far better than treatment. We study about image classification as we can prevent the disease. If it can be prevented according to the earlier symptom then the planting cost would low down and farmers would be benefited. The study was not just for cut the cultivating system cost also decrease the cachectic product. As human can harm by the cachectic product. The study and plan were based on proposed and unique mode, highest accuracy, a useable cheap product with the best quality and easy to use as farmers can use easy. We highlight the farmers as most of farmers of our country are uneducated. They are not familiar with the new technology, computer, internet and digitalization. We know Bangladesh has an internet services limitation. In everywhere internet is not available. And cost of the internet was our study contain. Training for farmers on a session is another thing what is included on our session. We have to conduct a session for the uneducated people or who didn't know before about Convolutional Neural Network, Machine learning, Deep Learning. Tomato's has various diseases. This paper highlights the top nine disease and classified all the specific criteria. It checks all the possible outcomes and the future too. This leaf disease detection by machine learning with the help of Convolutional Neural Network (CNN). At first, we research the process and what the studies was previous happened. We found may works and their different

criteria. Many models can detect varieties number of plants diseases. But tough we need to solve a specific problem and we choice Tomato's. Cause in our country tomato is vastly cultivating and it is a popular vegetable. We analysis different plant and its insect. One of the most reason that previously there are few researches had done for tomato's insect detection and prevention method. In Bangladesh is seldom work. So, we decided to solve specific the Tomato's leaf disease detection and earn a good yield. The studies about all the connected network and all of its possible algorithm. The are many technics are applied for this problem and different salutations have been come out. But all the approach is not accurate and applied able, specifically in our country name Bangladesh as we mentioned the limitations in our society. To consider the all limitations and all side of the project out mode is suitable for our farmers, country and agricultural firm. During the plan it also re-check abuts the risk scenario, risk management, possible real-life problem solving and the expected outcomes. Checking all the requirement we study about is and did the model to solve out this specific thing. The studies took place in single point of view and also a group point of view. The studies topics was to specific and clear to discuss. A successful session has been happened from the first and last time of the research. When one critical thing found the team, group discuss and fined the salutation of the problem. All the things go through a pure plan and right direction.

1.4 Research Questions

During the research we have faces may question and have some self-questions too. All the questions we completed by our self and find out all the answer separately and specifically. The more question we solve we solve the clearer we are. Before starting the project, we have something that is unclear, thought that was not a major point. But when we starting our project to doing the all critical and unclear going to easy to the team. At first the question was “which model or technics is best to image classification”? we found the salutation that is Convolutional Neural

Network (CNN) because of its feature and the highest accuracy. That's why the word is vastly using this method. Another one was "How we build up the model"? is was really a focusing section. But the aim was clear and according to the pre-plan the model builds up and gave us a pleased result, what is enough for such types of classification. In the other side the farmer is uneducated and they are not connected with the internet. So, here raising the question "How it possible to train and make familiar with this technology of farmers"? It just possible to a serial session with the farmer and the team teaches the farmer how to conduct the model. Though, the model is user friendly and easy process so it wasn't hassling for us. Awareness make a positive vibe to us. In additionally, secondary queries were raising hand during the work like: -

SQ1 – What type of technics create use of recent architectures or models?

SQ2 – What area unit the characteristics of the info sets preponderantly used?

SQ3 – What styles of crops area unit most investigated with approaches victimization CNNs?

SQ4 – What styles of approaches and frameworks area unit ordinarily used?

SQ5 – Which CNN formula is current in current approaches?

SQ6 – What styles of plant diseases area unit most investigated with approaches victimization CNNs?

As we told that, we took all the awareness during our whole process and all things we under control. We didn't think for a single time that we can't. We can and we have a plan and we have successfully done is crossing many more questions ahead of us and also, we were able to remove all the question that was raise on the mind of users. All the possible outcomes solve the all question and the critical think what

would can raise on the mind. We have a very general formation to solve the provable and the formation have been cleared after taking the project. It's funny things the WH question circulate all the time. It's been pleasure to solve problem and questions at a time.

1.5 Expected Output

we have a pleased outcome of our research. The model provides a great accuracy what is enough for building a model as useable. There are many models what work for the same problems but different outcomes come out. Though, all the researcher has the same intention to acquire 100% detection system. But is not possible to detect 100% this time. Cause, there are some imitations. Outcome depend on mostly Coding language, data, and model. Data is a big factor to get a good accuracy. With our model we got 87.5% accuracy.

Figure shows the outcome of the model,

```
Epoch 42/50
39/39 [=====] - 56s 1s/step - loss: 8.0567e-04 - accuracy: 1.0000 - val_loss: 1.4355 - val_accuracy: 0.8750
Epoch 43/50
39/39 [=====] - 58s 1s/step - loss: 6.5834e-04 - accuracy: 1.0000 - val_loss: 1.4656 - val_accuracy: 0.8750
Epoch 44/50
39/39 [=====] - 58s 1s/step - loss: 5.6728e-04 - accuracy: 1.0000 - val_loss: 1.4848 - val_accuracy: 0.8750
Epoch 45/50
39/39 [=====] - 56s 1s/step - loss: 4.7831e-04 - accuracy: 1.0000 - val_loss: 1.5125 - val_accuracy: 0.8750
Epoch 46/50
39/39 [=====] - 55s 1s/step - loss: 4.3324e-04 - accuracy: 1.0000 - val_loss: 1.5360 - val_accuracy: 0.8750
Epoch 47/50
39/39 [=====] - 56s 1s/step - loss: 3.7213e-04 - accuracy: 1.0000 - val_loss: 1.5611 - val_accuracy: 0.8750
Epoch 48/50
39/39 [=====] - 58s 1s/step - loss: 3.3279e-04 - accuracy: 1.0000 - val_loss: 1.5769 - val_accuracy: 0.8750
Epoch 49/50
39/39 [=====] - 58s 1s/step - loss: 2.9776e-04 - accuracy: 1.0000 - val_loss: 1.5945 - val_accuracy: 0.8750
Epoch 50/50
39/39 [=====] - 57s 1s/step - loss: 2.7474e-04 - accuracy: 1.0000 - val_loss: 1.6139 - val_accuracy: 0.8750
```

Figure 1: The accuracy of the model.

The accuracy 87.5% is a pleased accuracy as it is convolutional Neural Network (CNN). With this accuracy it can build a model. Mostly accuracy depend of the that. How accurate data is used for training and testing along with validation. Most of our dataset is from Kaggle. One of the most data are found in this online platform.

We also took some raw data what added with the raining and testing dataset. 90%-100% accurate prediction model is rate to find out. All the researcher finding hard to get more accuracy. But available 80%-90% accuracy is come out.

On the other side, it expected to reduce the crops insect as the farmers get a goof yield. Bu using this model is can be solving out and the user will be benefited. In addition, from here is not just benefited by farmers or agricultural farm, also will be benefited the local people who loves vegetables and special Tomato. Many researchers are satisfied with their model where their expected outcome was 90-100%. From their point of view, we have successfully over the operation and our outcome pleased. Though it tries more to acquire more accuracy but due to data pre-processing, raw data, source data, collecting type and model limitations it not possible to more. But if new thing found to us with what we can increase the accuracy of our model. We definitely developed our model with the new technology. But it can seem that the model outcome and expected outcome is not far from the destination. We tried to help the people and help the nation by our model that is most focus system in our destination. It is such inspiration also for us. The expected outcomes and our acquire things are almost similar. It really such hard things to seems crops disease and the loss of the farmer along with nation. At last, as the model working smoothly with great accuracy so out are passed on out journey and it would be happy to the farmers are benefited by using our model. Then the whole process makes the creator happy seems.

1.6 Project Management and Finance

Management and finance are a requirement field for research. In every sector management is a big role. If the management segment is being well decorated the whole project will be a decent form. And there is connection between management and finance. Cause, management depends on the finance. According to the fund,

the team will be mapping the management. Suppose, for a project 50,000-taka budget. Here management don't know about the budget. Then management do their budget and the cost for the project 80000-take. So, the project is not possible to complete as the budget 30000 is short. So, before taking any plan the finance should be published. And depending on the budget the management should arrange their plan. In this research project, we have collected a financial fund. We collect money from our supervisor and Co-supervisor as they have sent us money from their professor fund. And we (two members) contribute this project. After collecting this amount from all the members of our team, it looks a healthy amount and we(management) team are think it enough to build this project. We calculate all the necessary step how much money we will spend and how this amount will income from our model. In this management system, the members find out some income possible field from which site our income can come. All the possible field calculate and it finds that this research project will be successful and will be profited. When the users will star using this model and when you will be getting the benefits from this model, they will suggest other farmers and agricultural farm. Then the income will come a healthy amount. In additionally, there is a possible chance to buy it by the government. Besides, if any agricultural firm or any company wants to buy it then we can sell a copy it with discussion. It depends on the situation and board team. Many times, can seems that many authorities donate and support financially. It is a great opportunity. The strong site is that before doing this research project the team were so concern about its funding. After collecting all the money, it takes a deep breath. We weren't higher any management team from out site. We two team members manage all the criteria and the needful site. We note down all the field where we spend how much money and keep money receipt as to make this solid and clear finance report. All the finance things need to go through the explanation. Case, as we are two members are including with this finance sector and we collect fund from our supervisor and co-supervisor. So, we need to prover our self-fresh in

front of them. Cause they trust us and we try to keep their believe on us. By calculating all possible field and all the needful things we are started our research field and during our work field we didn't found any extra source where we have to spend money. It proves that, the management was so good. The management always tries to expend less amount as the less we expand the more we benefited. In this criterion we fulfill through finance and management process.

As we have some possibilities source from where we can earn by our project so, it not a risky thing to invest in our research project. Getting fund from third party is not easy before doing model. So, it must need to invest first. And as this is our own project so we our self-invested as I mentioned earlier. This is a cheap project, that's why we were able to do with this four person's investment. Overall, we weren't facing such types of financial crisis and the financial and management team is nicely decorated all the needful site. Getting extra support from others was impossible to our project. We hardly try not to expend any extra money and were not buy any extra material for the project. All the needful material was bought what exactly need to this project. The team were so careful about this.

1.7 Report Layout

A research layout contains different types of information. In a paper, there are different section like introduction, background, research methodology, result, discussion, impact, summary, conclusion and recommendation. This verities information also classified and part into another criterial. In this paper we can classified this report into three sectors, like: -

- i. The preliminary section
- ii. Main Text
- iii. The reference sections

For a preliminary analysis paper, or for the preliminary a part of an exploration paper (esp. dissertation), the content can usually include: -

- a. Introduction and Background
- b. Problem Statement
- c. Purpose
- d. Research queries
- e. Conceptual Framework
- f. Nature of the Study
- g. Brief Review of the Literature
- h. Conclusion

Together, these usually apply towards treatise topic approval, or as CNN introduction to the rest of the study, followed by the literature review, methodology, then forth. The importance of those preliminary elements of the analysis paper can't be underestimated to make sure topic viability and significance.

The second main text contains the main part of the research what we did on our paper. In the theme and conceptual area are so diplomates. Cause it also has some criteria. The part of the main text what added in this section. Amon the research layout we have concerted about the instruction like (Text size, typing style, figure contents, table contains, background, Bold up mark, line spacing size etc.). The whole report contains all of the instructions. In the paper we highly try to detect the Tomato's disease detection. Earlier detection is a very good way to detect the insect. If it can early detect then it would be able to prevent the disease that was the main focus of this research. Main goal is a biggest issue for any project. The yield down, losing economy, farmers los are concerted issues for us.

In the endo of the paper the reference section has been added. We have downloaded 17 research paper from google as to take reference and having a look which way or

technics they used. All the papers have cited and added in the reference section. To get a new mode it should to view others model what actually they did and what the main lacking their have. When we implement our model, we were so careful about their lacking and we carefully hand that specific sectors. By doing this works we got a valid model. In comparatively our model is best than other model and it can checkable with the reference paper.

CHAPTER 2

BACKGROUND

2.1 Preliminaries

The preliminaries about the research was so clear to us, to solve tomato's disease detection and get a good yield. We work for that and to identify the tomato's leaf insect earlier this model uses Convolutional Neural Network (CNN). CNN use for detecting the insect as is can take if feature automatically without any human supervision. In this model it can use any situation through the internet. As our country on the way o digitalization and model of the village have 4G internet connection and Broadband services too. So, the user can easily use. Before building this model, we noticed that the most of the user will be remote corner of the country as most of the last is cultivating villages. We work for that and as it is earlier mentioned there are some lacking and their lacking will be cover through training and providing the necessary equipment. In it needed to supervise the model through online or offline, we are committed to do that. We had a plan to connect the model as it is cloud based. Satellite base monitoring and services also include this plan. But this time it not possible as some limitations. Local satellite and cloud performance won't be a good decision. It would be costly and difficult process. As most of the user will be not up to the mark for education, we decided not to go any hard process, it tries to make as user possible as it. In this model there are some features will be added, like, suggestions about the disease, more information, kinds of insect and others. Not just user detect their disease also will get extra information from the model. At first this model will define different categories so it helps to detect the work. The model is simplicity rather than previously the model has been published. The coding site also so easily done.

2.2 Related Works

There are many works have been done previously. As this insect problem have been facing from many years so many works have been done by several researcher over the world. But it needs to remember that all the country's weather is not same. So, the characteristics of the disease, harmful type of the insect will not be same. But it might be same way not exactly same. All the researchers do not follow the same pattern. Different model will give different values and accuracy. With advancements in pc vision, progress has been achieved within the identification and designation of plant diseases. various designation and identification techniques square measure planned by the subsequent image segmentation procedures, feature extraction, and pattern recognition. Before the evolution of deep learning, the favored classification approaches that were used for unwellness detection in plants embrace random forest, artificial neural network (ANN), k-nearest neighbor (KNN), and support vector machine (SVM). Recognition ways exploitation the procedures mentioned before improved plant disorder designation. However, these approaches depend on the extraction and choice of visible unwellness options. Recently, several works on machine-driven disease designation and identification are developed using deep learning techniques.

Kawasaki et al. planned CNN architectures to acknowledge cucumber plant disease and obtained ninety-four.9% accuracy. CNN is that the most helpful classifier for image recognition in each little and large-scale dataset. it's shown glorious performance in image processing and classification. Mohanty et al. trained a deep learning model for recognizing fourteen crop species and twenty-six crop diseases with ninety-nine.35% accuracy exploitation GoogleNet and AlexNet design. CNN will perform each feature extraction and image classification. Srdjan et al. planned a disease recognition approach to classify healthy leaves and thirteen completely different diseases supported CNNs. The results demonstrate that sturdy

computing infrastructure makes CNN an appropriate candidate for disease recognition.

2.3 Comparative Analysis and Summary

In competitively the model for what we work for is the best model. And it has some reason too. The main reason is to simplicity use and architecture of it. Convolutional neural Network (CNN) as usual use most of the paper and many papers used traditionally Artificial Neural Network (ANN). That is time consuming. This model will have a new feature what will suggest the next step and will also analyzing with the fertilizer, medicine, weather conditions and PH of the soil. That will be the future step. But now we maid user friendly model and gives a satisfactory accuracy with what farmers or agricultural farm can predict their Tomato's disease. This model is based the Deep Learning (DL) technics and in especially Neural Network.

On the other hand, other technics is not so easy to conduct as their complexity. Artificial Neural Network (ANN), may be a cluster of multiple perceptron's or neurons at every layer. Ann is additionally referred to as a Feed-Forward Neural network as a result of inputs square measure processed solely within the forward direction. This type of neural networks square measure one among the best variants of neural networks. They pass info in one direction, through varied input nodes, till it makes it to the output node. The network could or might not have hidden node layers, creating their functioning additional explainable.

At last it can be considering that, in general compering our model will be the best solution for tomato's disease detection. As our training and testing data was pure along with the accuracy was 87.5% as it is mentioned earlier. Comparing our input data was some raw data. We both collect data from online and offline. Thought most of the data is from Kaggle but we took manually picture and clicking time we

were serious about any filter. Since, the data is raw, the model would be able to detect the disease.

2.4 Scope of the Problem

Research is the reverse things in our society. All the research based on any problem on our real-life or the problems what facing society. To make an easy life research are doing and also to know it happening, what exactly doing about any elements or things, then there makes a problem. And to solve that problem research are doing by the researchers. Several researchers are doing to know deep about the research topics and several researches are doing how exactly going to solve the problems.

However, analysis in tomato's, indeed within the social sciences usually, has not forever shared this halo. There are a unit in all probability as several analyses of this state of affairs in concert might realize analysts to gift them, however it appears that most points of read would include- a minimum of these aspects. Researchers found the problems then it tries to solve out the problems. From different aspect of the scope are come in front of the researchers. But all the probabilities to come problems are handle by the researchers. We faced different type of the problems and we also work with the problems. The scope is unlimited but it is listable and according to the real-life problem and research problem, if the team work hard to solve then it would be able to solve out the problems. We are work for detection of tomato's disease using the life image but there are different problem salutations research work what have done the same way. All the detect a problem from the environment and find out a salutation in an easy way with experiment. In additionally, most of the problems are created by the human or human need easy process. To help and make easy of our life researcher solving the problems.

2.5 Challenges

In every researcher have faces different challenges. It can be environmental, research topics-based challenges or internal challenges. During the operating we have facing some challenges like: -

- I. Data collecting
- II. Selecting the model
- III. Method and materials arranging
- IV. Getting best accuracy
- V. Management.

We successfully have passed all the challenges. Dur any hard problems we discuss team-wise and finding the salutations. Building a model with a grate accuracy is not easy. Tough none challenges have suffering more as we were confidence and clear about our goal. If any researcher stop himself/herself by facing any challenges then he/she can't find out the salutations. Team working is the best way to faces any kinds of challenges.

- (i) As it mentioned the challenges, one of data collection. As we all know accuracy and the performance of the model mostly depend how accurate the data is. So, collecting the raw and accurate data was a challenging matter to us. After collecting the data selecting was another concerted matter.
- (ii) With which model we will implement. Cause there were verities types of model. Which model would be better to us and would give us the best accuracy? We research about all model to select or model mode.
- (iii) All the material for the whole operation and the method is another part. As it is our first time work such field so we are finding hard to arrange the method and material.

- (iv) After the training and testing the accuracy is come out. But training data and testing data selectin is a part of challenge.
- (v) The last process and the first process is management. Managing the total process from up to bottom is a hard factor and the most challenging part.

But we passed all the challenges and have built the model. Though all the challenges were real life so we have been able to solve out it. From this challenge the team learn many things and for any further project it will be the planning list.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

Our research topic is tomato's leaf disease detection by Deep Learning using CNN. We did it using image classification method by using CNN. We used Jupyter notebook as platform and python as python as programming language. We used several python libraries such as NumPy, cv2, os, random and matplotlib. We used TensorFlow as machine learning library and keras as neural network library. "Sequential" is used to build CNN model.

Jupyter Notebook: Jupyter notebook is web-based application which is using for data driven projects. It visualizes the whole process of data processing, model building and final result. It also uses to collaborate across project and tools, data organization and cleaning and developing data science skill.

NumPy: NumPy is a python library. It's works with array. It also works for others domain like linear algebra, Fourier transform and matrixes. NumPy gives a large range and efficient ways to create arrays and manipulating numerical data. NumPy array consume less memory than python list. Because NumPy array store at a continuous place in memory. NumPy allows code optimization even further.

CV2: Open-CV is a python library which is used for image shows by cv2.imshow() method.

OS: 'os' module helps to do several tasks automatically such as creating or removing directory, changing, fetching and changing directory.

random: It works for creating floating number within range 0.1 to 1.0. Every time it gives different floating number.

matplotlib: ‘matplotlib’ is a visualization function which use to visualize statical graph or chat and animated image. ‘pyplot’ is mostly used as sub-module.

TensorFlow: “TensorFlow” is an open-source library which used for machine learning and artificial intelligence. It works for data automation, model training, model retraining and performance monitoring. It accepts data in form of multi-dimensional array. Because it makes more easier to handling large amount of data. TensorFlow works based on data flow graphs which have nodes and edges because it helps execute in a distribute manner across the clusters. TensorFlow build by parts. They are data-preprocessing, building the model, training and estimating model.

Keras: “keras” is a deep learning API (Application Programming Interface) for implementing neural network. It makes neural network implementation easier. It also easy learn and work. Several frameworks are supported by keras. Mostly TensorFlow is used as framework. Keras is embedded, it helps to perform deep learning fast. Keras has simple APIs which makes it easier and reduce the code size also explain user error clearly. That is how it can reduce implementation time. Keras runs easily on both CPU and GPU and it support almost all neural network models.

3.2 Data Collection Procedure/Dataset Utilized

Data in one of key component in research. To do any research first we data. Data sources can mainly divide into two sources. One is primary source and another is secondary source. Primary source is containing interview, survey, observation and experiment. Secondary source is containing internal source, external source and

other sources. Internal sources data means market record, sales record, customer data. External sources mean data from various publications We collect data from secondary source. We collect a tomato leaf image dataset Kaggle with more than ten thousand images. Our dataset has ten subsets. We selected more than thirteen hundred images for our project. Images are in various sizes. We convert all images into 256X256 size. We used 80% data for training and 20% for testing and validation. From dataset we know that tome leaf can affected from nine types of diseases. They are Bacterial spot, Early blight, Late Blight, Mold, Septoria Spot, Spider Mite, Target Spot, Mosaic Virus and Yellow Curl Virus.

This image shows the verities disease of tomato,

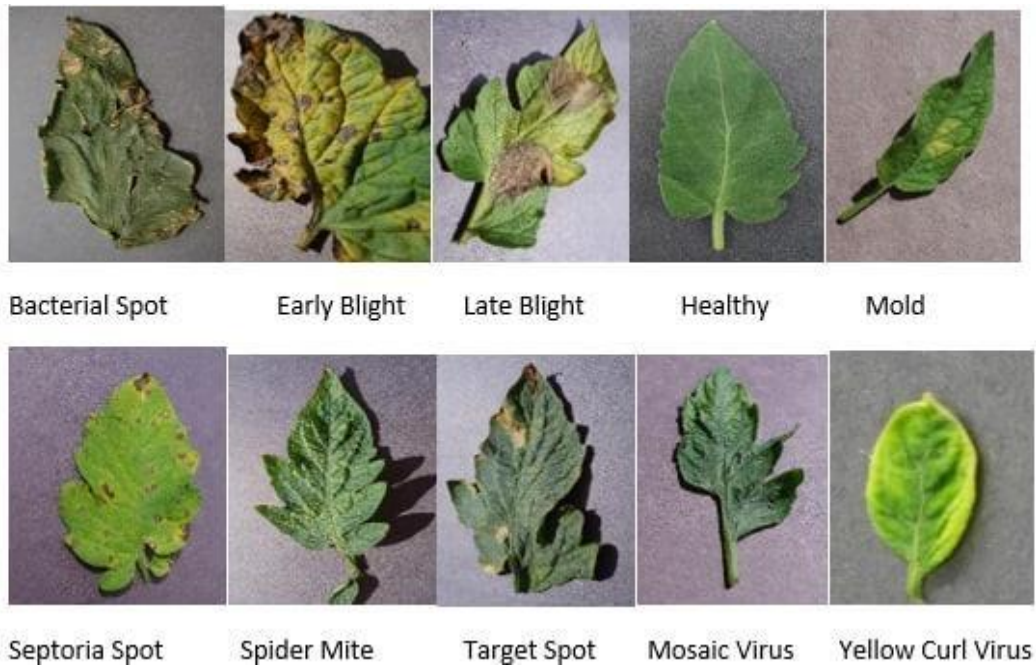


Figure 3.1: Nine leaf disease leaf a healthy leaf.

This table shows the all of the disease for tomatoes,

Tomato's disease name

Number	Disease Name
1	Bacterial Spot
2	Early Blight
3	Late blight
4	Mold
5	Septoria Spot
6	Spider Mite
7	Target Spot
8	Mosaic Virus
9	Yellow Curl Virus

Table 3.1: Tomato's disease name.

3.3 Statistical Analysis

The graph at below shows the result of our model. First shows accuracy where blue line shows training accuracy and yellow line shows validation accuracy. Second graph shows error of model. Where blue line shows training error and yellow line shows validation error. In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5% and last seven epochs give same result which is 87.5%. So, our model accuracy is 87.5%. This model is an overfit model. Because in this model, training accuracy is higher than validation

accuracy. At the same time training loss is much less than validation loss. That is why it is an overfit model.

This figure shows the model accuracy and error,

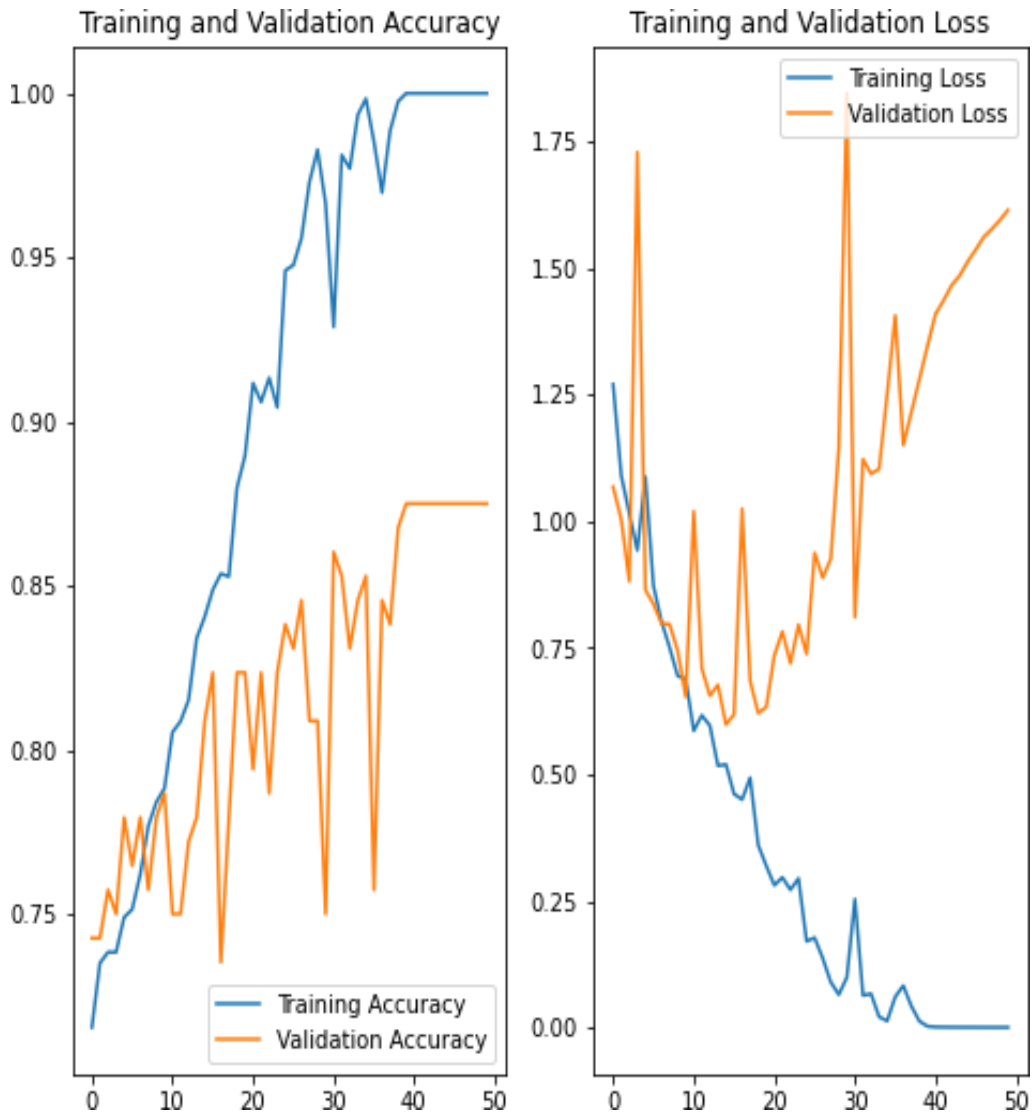


Figure 3.2: Accuracy and Error graph.

3.4 Proposed Methodology/Applied Mechanism

We used Sequential model to build CNN (Convolutional Neural Network) model. As we know CNN model has basically four layers Convolutional layer, Maxpooling layer, Flatten layer and Dense layer. We also used some hidden layer also. We used “relu” as an activation function for convolutional layer and hidden layer. “SoftMax” is used as activation function in Dense layer. We used “sparse_categorical_crossentropy” as loss function and “Adam” as an optimizer. We use 32 filter for our convolutional layer and filter size is 3X3. And pool size is 2X2. Dataset has ten subset and model has ten types output. So, the model has ten neurons.

This figure shows the model architecture,

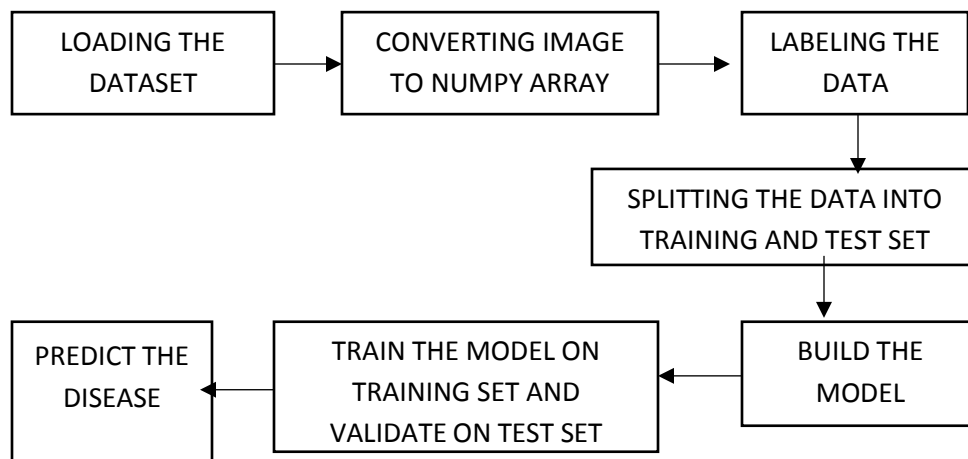


Figure 3.3: Model Architecture.

Convolutional Neural Network: Convolutional neural network is a part of machine learning. It works like human. Human has billions of neurons. Just like human It builds by multiple artificial neurons in a specific way. It specially used for image identifications. It works like human brain. Firstly, model needs to train by data. Then it can recognize similar things like human. In convolutional neural

network, the input image goes through a set of filters. Each filter contains a part of image and pass it's out to filter in next layer. Each layer train to identify to identify different features of input. This process done in several time. This is how the whole model get train to identify certain object.t has mainly four layers. They are Convolutional Layer, MaxPooling Layer, Flatten Layer, Dense Layer.

Convolutional Layer: Convolutional layer is the main layer of CNN model. It builds with some parameter and filters. Filter size must less than image size. Filter has width and depth. It divided input parameters according filter size. It can extract image features such as size, texture and edges. Low level features make it difficult to identify. In our model, 32 filter is used and it size it 3X3. Several activation function can be used ReLU, Leaky ReLU, Linear and Sigmoid. We use ReLU as activation function.

MaxPooling Layer: Maxpooling layer calculate maximum value of features. It uses after convolutional layer. It helps model to become over-fitting by providing an abstracted form of representation. It reduces the parameter size to reduce computational cost.

Flatten Layer: Flatten layer is a function which convert the pooled featured map into a single column. Flatten layer is used when multidimensional output needs to pass in dense layer in linear form.

Dense Layer: Neural network has several neurons. Dense layer works to make connection among layers. Dense layer get result from each neuron and it perform matrix vector multiplication. Dense layer also needs activation function and in our model, we use “Softmax”.

Model summary

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 64)	0
conv2d_2 (Conv2D)	(None, 60, 60, 64)	36928
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 64)	0
conv2d_3 (Conv2D)	(None, 28, 28, 64)	36928
max_pooling2d_3 (MaxPooling2D)	(None, 14, 14, 64)	0
conv2d_4 (Conv2D)	(None, 12, 12, 64)	36928
max_pooling2d_4 (MaxPooling2D)	(None, 6, 6, 64)	0

conv2d_5 (Conv2D)	(None, 4, 4, 64)	36928
max_pooling2d_5 (MaxPooling2D)	(None, 2, 2, 64)	0
flatten (Flatten)	(None, 256)	0
dense (Dense)	(None, 10)	2570

```

=====
Total params: 169,674
Trainable params: 169,674
Non-trainable params: 0
=====

```

In the model summary we have shows all the model and the layer too. It can describe that which model use and what the layer in here. Convolutional layer, pooling layer, hidden layer, max_pooling layer, flatten layer and all of the layers are shows in this position. It can show that total params 169,174, trainable params 169,674 and non-trainable params in here are 0. Convolutional layer is the main layer of CNN model. It builds with some parameter and filters. Filter size must less than image size. Filter has width and depth. It divided input parameters according filter size. It can extract image features such as size, texture and edges. It specially used for image identifications. It works like human brain. Firstly, model needs to train by data. Then it can recognize similar things like human.

The below figure shows the plot model,

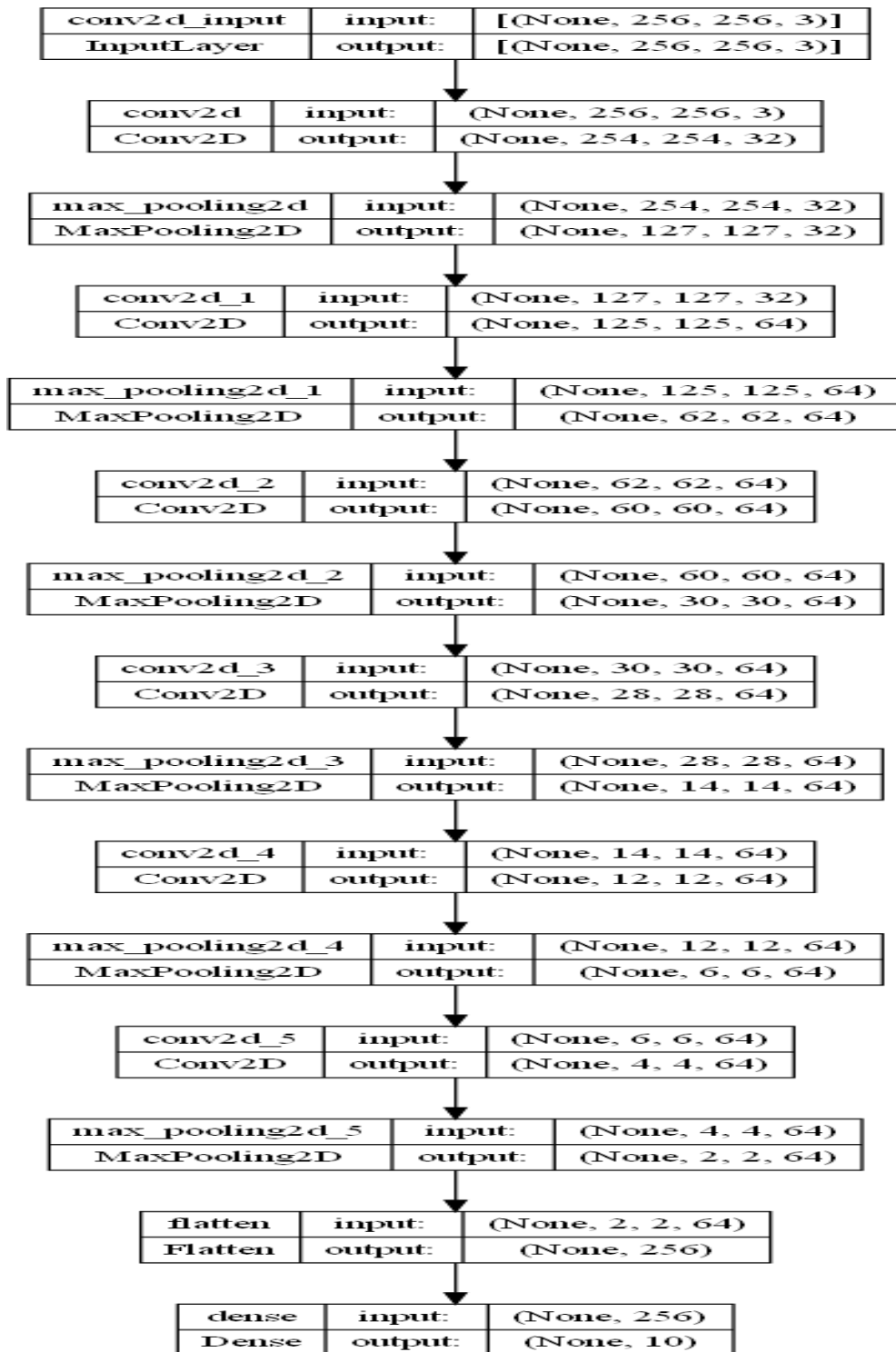


Figure 3.4: Plot Model

3.5 Implementation Requirements

For kind of experiment, first we need some data which is relevant with the experiment. Our experiment is based on image classification. So, we need image data. We collect our data from online published data from “Kaggle”. We collect multiple data sets. Then we judge them which will be more suitable for our model. We judge image quality, resolution. Which is a secondary source. Images are in different sizes. First convert the images to the same size by scaling.

Implementation requirements mean necessary components and skills to complete a research or project. Before starting research, researchers have to learn necessary skills and collect those components. It also includes data and team. Me and my team mate made a group. Our supervisor supervised us to complete this project.

As we said earlier our project is a machine learning project. To complete this project, it needs skill of programming, python language and some of python libraries such as NumPy, cv2, os, random, matplotlib, TensorFlow and keras. A programming platform is needed to do this project. We use Jupyter notebook as programming platform. We make a project to identify “Tomato Leaf Disease”. We build this project by using CNN. “Sequential” is used as model to build CNN.

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental Setup

To do machine learning project we setup anaconda in our PC. We done our project on Jupyter notebook. Before that we install some of important libraries in Jupyter notebook such as NumPy, os, cv2. We also install TensorFlow and keras in Jupyter notebook. We use “Sequential” to build our model.

Data is the key things for any kinds of research. When any wants to do research, first thing comes on his or her mind that is data. Any research begins by data collection. So, we also collect data firstly. Data collects from two main sources. One is primary source and another is secondary source. Primary source is containing interview, survey, observation and experiment. Secondary source is containing internal source, external source and other sources. Internal sources data means market record, sales record, customer data. External sources mean data from various publications. We have collected our data from secondary source. We collect our data from Kaggle’s. Dataset has more than ten thousand data. We have selected more than thirteen hundred images for our research. Our research topic is tomato leaf disease detection by CNN. From data we know that tomato leaf has nine types of diseases. So that our data set has ten subsets. One is healthy and others are unhealthy.

As early we said, we used CNN (Convolutional Neural Network) for our model. We used “Sequential” model to build CNN structure. Our model has ten neurons. Because model has ten different outputs. Because our data set has ten different subsets. As we know CNN model has four main layers. They are convolutional layer,

Maxpooling layer, Flatten layer and Dense layer. Our model has some hidden layer. Hidden layers are built by both Convolutional layer and Maxpooling layer. This table shows the model's information,

Model's Information

Total Data	Training Data	Testing Data	Epoch	Loss of Accuracy	Model Accuracy
1359	80%	20%	50	12.5%	87.5%

Table 4.1: Model's Information.

The below figure shows the ANN visualization,

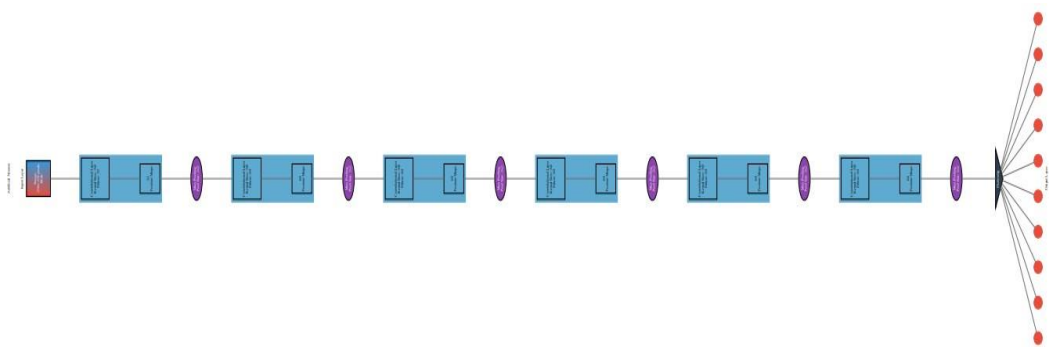


Figure 4.1: ANN visualization.

4.2 Experimental Results & Analysis

As we know CNN is the best fit for image classification, so we experiment our model by using CNN. We got a classic result 87.5%. With this accuracy building a model is effective. There are few models what has the more accuracy than us but that model is complex. Cause, they have use varieties plants detection system. But this model is till now just for Tomato's disease detection. A single problem will solve a single model. That's why it is usable model to detect tomato's disease. As

it just solves the tomato's disease there are no chance to mix-up the input image with another plant image. It can be for others model. So, this model will give accurate value, it has the possibilities to give more than 87.5%, when the input image will be perfect for detection and classification.

Result Shows our model is overfit model. Because our model more accuracy in training than validation and our validation loss is more than training loss. In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. And last seven epochs give same result which is 87.5%. So, our model accuracy is 87.5%. Our data set has ten sub-sets. So, this CNN model has ten neurons. Every contains some marks out of hundred. And which neuron will get highest marks that will granted as output result. So, in 50 epochs any of ten neurons will get heights marks out hundred. For our model it will be between 75% to 87.5%. During the experiment, the model took the input. The ispre-process by the input layer, the hidden layer do the actual experiment like classification, matching, and detection based on the training data. This figure showsthe model experimental result,

Logistic Regression: Logistic regression follows supervised learning system. It used to predict the probability of binary like (Yes/No) or 1/0 of an event. It is a statical model to use the logistic function or logit function which is sigmoid function where 'y' considers as a sigmoid function of 'x'. Logistic regression can also analyze multiple independent variables.

$$\text{Sigmoid function, } S(x) = \frac{1}{1 + (e)^{-x}}$$

Support Vector Machine: Support Vector machine learning is a supervised machine learning algorithm is used classification specially. It crates the best line to sperate n-dimensional space into class to put new data point in correct category and best line is called hyperplane. SVM choose the extreme points to create the hyperplane or best decision line. There are two types of SVM. They are Linear

SVM and Non-linear SVM. In linear data can be separate by straight line. But it is not possible in non-linear.

K-Nearest Neighbor: K-Nearest Neighbor is supervised learning algorithm which is used for regression and also for classification on nonlinear data. known as KNN. KNN algorithm creates some data point and closest data point are considered as a one cluster and others closest data points create another cluster. In KNN ‘K’ has value and algorithm choose the data point which is closed to value of ‘K’.

Decision Tree: Decision Tree is a supervised machine learning algorithm use for regression and classification model. It creates a tree which has root, branches and nodes. It follows hierarchical method. It considers best attribute as tree root and split dataset into subset. Each subset takes same amount of data. Subsets are repeat the same procedure till get the leaf node.

Naïve Bayes: Naïve Bayes is a classification algorithm use for both binary and multiclass classification. It follows the bayes theorem to measure conditional probability. Bayes theorem is given bellow.

$$P(H | E) = \frac{P(E|H)*P(H)}{P(E)}$$

P(H) means the probability of hypothesis.

P(E) means the probability of evidence.

P(E|H) means the probability of truthiness of hypothesis by given evidences.

P(H|E) means the probability of hypothesis given that evidence is true. Tables 4.2 shows the different Deep Learning model accuracy,

Different model accuracy

Model Name	Accuracy
CNN	87.50%
Logistic Regression	69.12%
SVM	72.05%

KNN	50.73%
Decision Tree	58.82%
Naïve Bayes	71.32%

Table 4.2: Different model accuracy.

In the model have been applied six different types of Deep Learning model and every model have given us a specific accuracy but Convolutional Neural Network have sent us a great accuracy. In this part we have implement a model based on CNN as it gives us the highest accuracy. The most vastly usable model it is as it mentions previous. For automatic feature map and classification power it is popular to use.

In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. But last seven epochs give same result which is 86.76%. So, our model accuracy is 86.76%. Our data set has ten sub-sets. So, this CNN model has ten neurons. Every contains some marks out of hundred. And which neuron will get highest marks that will granted as output result. So, in 50 epochs any of ten neurons will get heights marks out hundred. For our model it will be between 75% to 87.5%. But last some of epochs give same result that 86.76%.

Our experiment is to identify “Tomato Leaf Diseases” by image classification. To classified images, we build CNN (Convolutional Neural Network) model. “Sequential” is use as model to build CNN. To build CNN, four needed layers are given. They are Convolutional layer, Maxpooling layer, Flatten layer and Dense layer. We also use some hidden layer to increase model accuracy. We use 32 filters for convolutional layer and 64 filters for hidden convolutional. To activate convolutional layer, we use “ReLu” as activation function and “SoftMax” for denselayer. As a model optimizer “adam” is used.

The training and validation avalidation lose and training and validation accuracy shown in the below graph

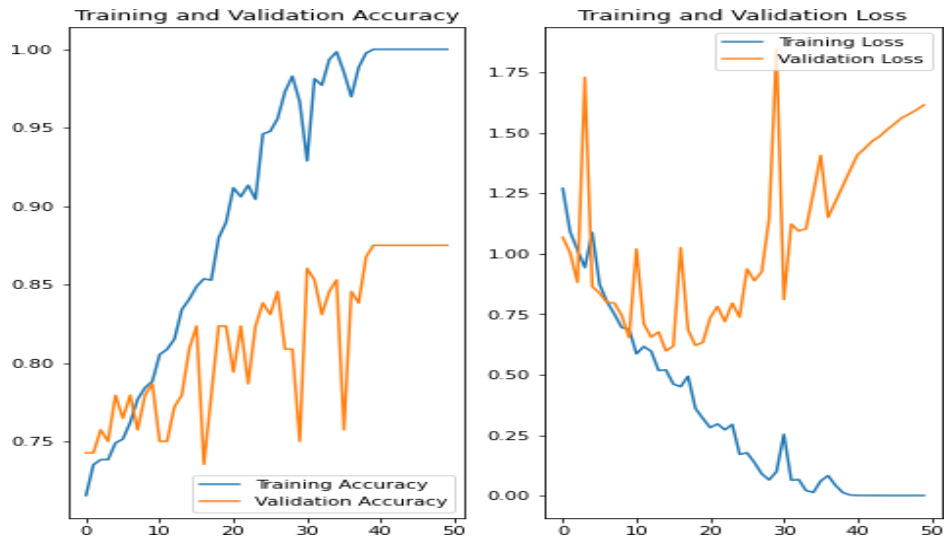


Figure 4.2: Accuracy and Loss graph.

For the experiment this model took 1369 data and use 80% training and 20% for testing. In addition, that, we have downloaded then thousand (10,000) of data for this research but 1369 we have selected according the data quality to get an accurate value. Is can say the experiment section have been done under a subtle observation. As is there is no shortage. In this part we specially, focus and re-check for every part. So, competitively is model will be the best solution for single detection of tomato's disease.

Figure 4.3 shows the Confusion Matrix for CNN and the accuracy is 81.82%

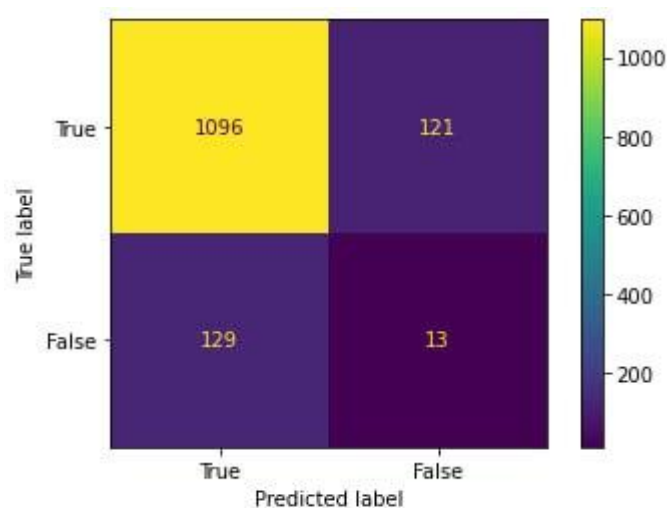


Figure 4.3: Confusion Matrix of CNN.

In the above figure 4.3 shows the confusion matrix on the where True Positive (TP), True Negative (TN), False Positive (FS) and False Negative (FN) have been shown.

Figure 4.4 shows the Confusion matrix for Logistic Regression. And the accuracy score is 69.11%

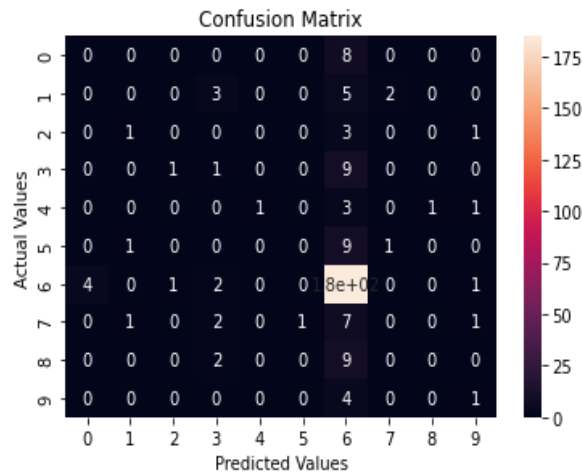


Figure 4.4: Logistic Regression

Figure 4.5 shows the Confusion Matrix for SVM. And the accuracy is 72.05%

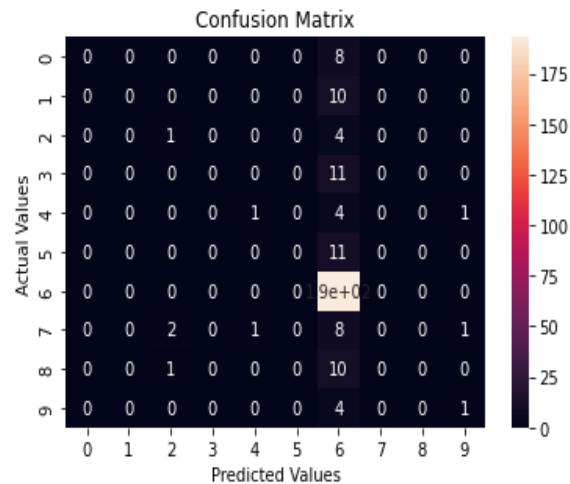


Figure 4.5: Support Vector Machine.

Figure 5.6 shows the Confusion Matrix for K-Nearest Neighbor and the accuracy is 50.73%

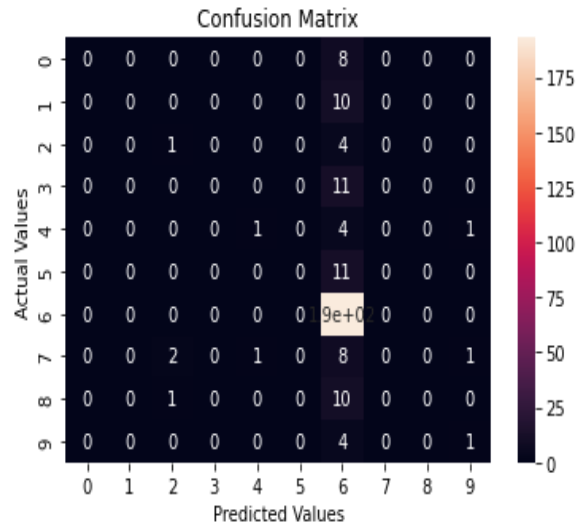


Figure 4.6: K-Nearest Neighbor.

Figure 4.7 shows the Confusion Matrix for Decision Tree and the accuracy is 69.92%

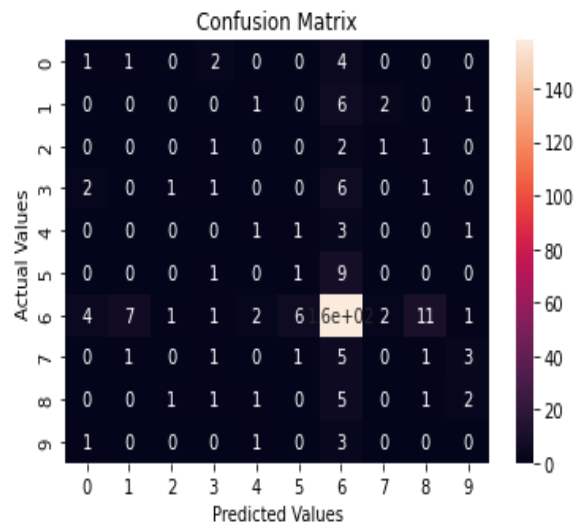


Figure 4.7: Decision Tree.

Figure 4.8 shows the Confusion Matrix for Naïve Bayes and the accuracy is 71.32%

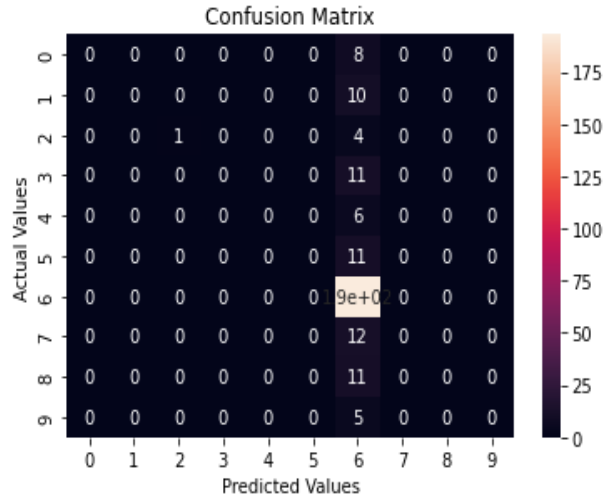


Figure 4.8: Naïve Bayes.

For CNN the matrix got the accuracy 81.88%, Sensitivity 12.68%, Specificity 89.38%. For Logistic Regression the accuracy is 69.11% and SMV the accuracy is 72.05%, for KNN the accuracy is 50.73, Decision Tree 56.92% and for Naïve Bayes the accuracy is 77.32%.

4.3 Discussion

When any one wants do any kinds of research, first thing comes on his or her mind is data. Without data research is impossible. Data can be collected from experiment, interview or survey. This primary data sources. Other source is secondary source, it can be any published data on online. We collect our data from “Kaggle” which is

In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. But last seven epochs give same result which is 86.76%. So, our model accuracy is 86.76%. Our data set has ten sub-sets. So, this CNN model has ten neurons secondary source. Our experiment is to identify “Tomato Leaf Diseases” by image classification. To classified images, we build CNN (Convolutional Neural Network) model. “Sequential” is use as model to build CNN. To build CNN, four needed layers are given. They are Convolutional layer, Maxpooling layer, Flatten layer and Dense layer. We also use some hidden layer to increase model accuracy. We use 32 filters for convolutional layer and 64 filters for hidden convolutional. To activate

convolutional layer, we use “ReLU” as activation function and “SoftMax” for denselayer. As a model optimizer “Adam” is used.

In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. But last seven epochs give same result which is 86.76%. So, our model accuracy is 86.76%. Our data set has ten sub-sets. So, this CNN model has ten neurons. Every contains some marks out of hundred. And which neuron will get highest marks that will granted as output result. Our data set has more than ten thousand data. We used more than thirteen hundred images for our model. Our data set has ten sub-sets. One contains healthy leaf images and others contained unhealthy leaf images. In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. But last seven epochs give same result which is 86.76%. So, our model accuracy is 86.76%. Our data set has ten sub-sets. So, this CNN model has ten neurons.

Our experiment is to identify “Tomato Leaf Diseases” by image classification. To classified images, we build CNN (Convolutional Neural Network) model. “Sequential” is use as model to build CNN. To build CNN, four needed layers are given. They are Convolutional layer, Maxpooling layer, Flatten layer and Dense layer. We also use some hidden layer to increase model accuracy. We use 32 filters for convolutional layer and 64 filters for hidden convolutional. To activate convolutional layer, we use “relu” as activation function and “SoftMax” for dense layer. As a model optimizer “adam” is used. In our model, we run 50 epochs. Several epochs give several results. Among them lowest is 75% and highest is 87.5%. But last seven epochs give same result which is 86.76%. So, our model accuracy is 86.76%. Our data set has ten sub-sets. So, this CNN model has ten neurons. Every contains some marks out of hundred. And which neuron will get highest marks that will granted as output result. So, in 50 epochs any of ten neurons will get heights marks out hundred. For our model it will be between 75% to 87.5%. Our experiment is to identify “Tomato Leaf Diseases” by image classification. To classified images, we build CNN (Convolutional Neural Network) model. “Sequential” is use as model to build CNN. To build CNN, four needed layers are given. They are Convolutional layer, Maxpooling layer, Flatten layer and Dense layer.

Our data set has more than ten thousand data. We used more than thirteen hundred images for our model. Our data set has ten sub-sets. One contains healthy leaf images and others contained unhealthy leaf images.

Our experiment is to identify “Tomato Leaf Diseases” by image classification. To classified images, we build CNN (Convolutional Neural Network) model. “Sequential” is use as model to build CNN. To build CNN, four needed layers are given. They are Convolutional layer, Maxpooling layer, Flatten layer and Dense layer. We also use some hidden layer to increase model accuracy. We use 32 filters for convolutional layer and 64 filters for hidden convolutional. To activate convolutional layer, we use “relu” as activation function and “SoftMax” for dense layer. As a model optimizer “adam” is used.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The model makes a positive impact on the society. As this model made to help the people and solve a problem then it would be a great impact on our society. This research is for tomato's disease detection and helps the farmers or agricultural farm who will cultivating tomato's for commercial purpose. For the leaf insect the farmers don't get a good yield and good profit. When the user uses this model then the user will be able to detect the disease along with, they will get a god yield. If they get a good yield then by selling their product, they will get a huge profit. And the local society's economic conditions will be developed. The farmers family will change their life style. And in future if they want, they can change their profession of they can stay their same profession and can cultivate more tomatoes'.

On the other hand, since the production is growing as the detect tomato's disease earlier so overall production will be huge. So, by fulfill the domestic demand the government can export tomato's in the international market. Bangladesh will be able to earn foreign currency, with the foreign currency the countries reserve will be increase. In the same time, the government and country will be acclaim. All the system depends on the production and production depends use of our model. In additionally, the cultivation cost will be decrease by using it. If cultivating cost decrease then the cultivator has the chance to be more benefited. There are huge possibilities to cut the percentage of the jobless people. All the unemployed people will get motivation from another people when they will see the profit of this

farming. We know the unemployed person is a big problem for the country and society. Here has a chance to solve unemployed people by cultivating tomato with the help of technology “Leaf disease detecting by CNN”. So, it can seem that our model not just solve one problem it also solves the jobless problem too. One problem is connected with the another one. Like this way most of our society’s problems will be solve by starting our model. The model will create a positive impact on the society and helps the society to solve out its problems.

5.2 Impact on Environment

As like the impact of society the research will create a great positive vibe on the impact on the environment. Using insecticides is a great problem for the environment. Most of the farmers and agricultural farm use pesticidal material for reduce the insect. But it creates a bad impact on the environment. But when they implement our model on their field then they will be able to detect the disease earlier and will get good yield. Most of the time farms use this harmful material by the suggestions of another unknown or inexpert person. They suggest to use more insecticides but it is dangerous for the tomato and environment too. It mixed with the rain and come to pond, cannel, river and pollute the river. By using this fish also attracted by disease. As it is connected each other the elements so, if one element be attracted or hampered then another element will be hampered. So, using our model will reduce all the negative impact and make our environment clean. This research is not just a research, during the pre-plan about the research we careful about environment. It is an environment friendly model which will help our environment to build disease free environment.

5.3 Ethical Aspects

All types of work have the ethical aspects. This research is not excepting of that. There are several ethical issues in this research. When we design our research, we

highly try to maintain all the ethical issues. As without ethics any work has zero value. During the research every researcher maintain the ethics. We also maintained the ethics. Few but most important ethics are: -

- a) Voluntary participations
- b) Informed assent
- c) Anonymity
- d) Privacy or Confidentiality
- e) Probable for hard
- f) Result Communication

Voluntary participation means all analysis subjects are liberated to like better to participate with none pressure or coercion. All participants are ready to withdraw from, or leave, the study at Any purpose while not feeling an obligation to continue. Our participants don't ought to offer a reason for going the study.

Informed consent refers to a scenario during which all potential participants receive and perceive all the data they have to make a decision whether or not they need to participate. This includes info regarding the study's edges, risks, funding, and institutional approval.

Anonymity implies that you don't apprehend United Nations agency the participants are and we can't link any person participant to their information.

We can solely guarantee namelessness by not grouping any in person characteristic information for example, names, phone numbers, email addresses, science addresses, physical characteristics, photos, and videos. In several cases, it's going to be not possible to actually anonymize information assortment. for instance, information collected in the flesh or by phone can't be thought-about totally

anonymous as a result of some personal identifiers (demographic info or phone numbers) are not possible to cover.

Confidentiality implies that we recognize WHO the participants are, however we take away all distinctive data from our report. All participants have a right to privacy; thus, we must defend their personal information for as long as we store or use it. Even after we can't collect information anonymously, we must secure confidentiality whenever we'll.

As an investigator, we've got to think about all doable sources of damage to participants. damage will be available in many alternative forms.

Psychological harm: Sensitive queries or tasks could trigger negative emotions like shame or anxiety.

1. Social harm: Participation will involve social risks, public embarrassment, or stigma.
2. Physical harm: Pain or injury may end up from the study procedures.
3. Legal harm: reportage sensitive knowledge may lead to legal risks or a breach of privacy.

The method we communicate our analysis results will typically involve moral problems. sensible science communication is honest, reliable, and credible. It's best to form your results as clear as attainable. Take steps to actively avoid plagiarism and analysis misconduct where attainable.

5.4 Sustainability Plan

The plan for near future like 2,5- or 10-years plant. In a plan to fulfill the research or work have to continue. It can be considering a goal. During the research this plan can be fix or before starting the research it considering the pre plan. Five contents can be considered as sustainability plan: -

- a. Appreciation.
- b. Progress.
- c. Implementation.
- d. Evaluation.
- e. Reassessment.

In our research we consider all the plan. We asses our research to build a great model. Plan have to implement all the disease to on frame. To build it the work progress will be count by the team. And it just not sustainability, it nee to implement on the field. We already implement on our program on the real life. So, implementation have a work in this area. We evaluate our plan and work. We will count how it can make possible and also will do for the plan. It just not to make a plan, we work for fulfil the plan. After make all the method and requirement things it needs to reassessment. We have a clear sustainable plan and we will try to touch our plan. The research plan has different types, many researchers have tries to collaborate with the future plan and sustainability plan. But we define it two types and make different work for those plans. Based on the answers to the queries higher than, develop the subsequent components to be inserted into the property set up. for every of those elements:

- ✓ Fixed measurable, quantitative goals.
- ✓ Discuss the means that for achieving these goals.
- ✓ Fixed timelines for achieving your goals.
- ✓ Fixed new goals once previous ones are achieved.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION AND FUTURE RESEARCH

6.1 Summary of the Study

This research has done by using Deep Learning with the Convolutional Neural Network (CNN). As the model build with this technic. Convolutional Neural Network has some advantages such types of work like automatic feature selection and a great accuracy. The framework is very effectively workable without human supervision. In the whole process have been done through a sequence of study. All the field like model, data and researchable things is connected with each part to another part. We have learnt many new things and gather a new experience during the studies. The team members help each other during studies and their own part. By doing the study before our any work helped us to the work. Tomato's disease detection using Deep Learning (DL) with the convolutional neural network is a field work research. In this research have separated into two categories. That are field work and compute based. Data collection are based open the field work, another data part (Kaggle) and building model including training, testing, validation is based on the computer part. We need not any external supervision except our team and we didn't take any support or such kinds of help from other. Finally, I seem that we have done our research through a great study.

6.2 Conclusions

Convolutional Neural is a significant branch of deep Learning (DL). Because it has strong feature capabilities power and it mostly introduce itself for plant disease. There are several methods for "plant disease detection", but specifically tomato's

disease detection using Convolutional Neural Network (CNN) are rare with a great accuracy. What's additional, there square measure withal no business preparations accessible, apart once people managing plant species acknowledgment addicted to the leaf's photos. The full strategy was delineated, separately, from collecting the images utilized for getting ready and approval to data preprocessing and growth in conclusion the procedure of preparing the deep CNN and fine-tuning. As CNN takes its feature automatically the is use that the feature requires for the time. This method will individually detect the tomato's disease earlier.

6.3 Implication for Further Study

All the work has some futures plan or sustainability. It can call by vision. Our research also has some these visions. Right now, this research has the ability to detect the tomato's disease. But in near future we will work to develop it as all types of leaf will be include in this model and all kinds of disease can detect through our model. Then the farmers/agricultural/user need not to find other solutions to detect their plant disease through another model. We all merge it in our further implication and we have another plan to give the instant salutations. Suppose, the detection has done and the user wants the next step to prevent the disease. That time our model will suggest a treatment based on the disease. If this segment is done then it would be a unique model whatever have implement in our country or the world. Then Our research and mode will be a best solution for leaf/plant disease detection by Convolutional Neural Network (CNN). Through our work this model will be a world class model. Right now, our research accuracy is 87.5%, we will study more to improve our accuracy and give a more accurate model. If it can be done realistically then, this research will get more acceptance to the user. As this report earlier mentioned that aim was to build this model to help the people. Implementation this future study the aim will be successful. When all kinds leaf disease can be detecting from this research then the leaf will get a new disease-free

life and crops will give more production. Crop's yield will be increase. At last it can say that, our future study will add a new vibe on this model and the whole world will be benefited from our research.

Reference:

- [1] Mishra, S., Sachan, R., & Rajpal, D. (2020). Deep convolutional neural network-based detection system for real-time corn plant disease recognition. *Procedia Computer Science*, 167, 2003-2010.
- [2] Francis, M., & Deisy, C. (2019, March). Disease detection and classification in agricultural plants using convolutional neural networks—a visual understanding. In *2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 1063-1068). IEEE.
- [3] Yadhav, S. Y., Senthilkumar, T., Jayanthi, S., & Kovilpillai, J. J. A. (2020, July). Plant disease detection and classification using cnn model with optimized activation function. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 564-569). IEEE.
- [4] Abade, A., Ferreira, P. A., & de Barros Vidal, F. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Computers and Electronics in Agriculture*, 185, 106125.
- [5] Mohsin Kabir, M., Quwsar Ohi, A., & Mridha, M. F. (2020). A Multi-Plant Disease Diagnosis Method using Convolutional Neural Network. *arXiv e-prints*, arXiv-2011.
- [6] Kumar, S. (2021). Plant disease detection using CNN. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(12), 2106-2112.
- [7] Walleign, S., Polceanu, M., & Buche, C. (2018, May). Soybean plant disease identification using convolutional neural network. In *FLAIRS conference* (pp. 146-151).
- [8] Toda, Y., & Okura, F. (2019). How convolutional neural networks diagnose plant disease. *Plant Phenomics*, 2019.

- [9] Hassan, S. M., Jasinski, M., Leonowicz, Z., Jasinska, E., & Maji, A. K. (2021). Plant disease identification using shallow convolutional neural network. *Agronomy*, 11(12), 2388.
- [10] Bedi, P., & Gole, P. (2021). Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. *Artificial Intelligence in Agriculture*, 5, 90-101.
- [11] Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2015, December). Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. In *International symposium on visual computing* (pp. 638-645). Springer, Cham.
- [12] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, 145, 311-318.
- [13] Prashanthi, V., & Srinivas, K. (2020). Plant disease detection using Convolutional neural networks. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3).
- [14] Marzougui, F., Elleuch, M., & Kherallah, M. (2020, November). A deep CNN approach for plant disease detection. In *2020 21st International Arab Conference on Information Technology (ACIT)* (pp. 1-6). IEEE.
- [15] Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, 8(11), 468.
- [16] Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In *2018 3rd international conference on computer science and engineering (UBMK)* (pp. 382-385). IEEE.
- [17] Shrestha, G., Das, M., & Dey, N. (2020, October). Plant disease detection using CNN. In *2020 IEEE Applied Signal Processing Conference (ASPCON)* (pp. 109-113). IEEE.

Demo

ORIGINALITY REPORT

17%
SIMILARITY INDEX

12%
INTERNET SOURCES

12%
PUBLICATIONS

7%
STUDENT PAPERS

PRIMARY SOURCES

- 1** Melike Sardogan, Adem Tuncer, Yunus Ozen. "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018
Publication 1%
- 2** www.mdpi.com
Internet Source 1%
- 3** "Ubiquitous Intelligent Systems", Springer Science and Business Media LLC, 2022
Publication 1%
- 4** Submitted to University of California, Los Angeles
Student Paper 1%
- 5** ebin.pub
Internet Source 1%
- 6** Jihan Alam. "An Efficient Approach for Skin Disease Detection using Deep Learning", 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 2021 1%