GRAPE LEAF DISEASES DETECTION USING DEEP LEARNING

BY

ALIF JAWAT
ID : 191-15-2649
AND

MD. FORHAD HOSSAIN RAHI
ID : 191-15-2402

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

Supervised By

Ms. Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Md. Sabab Zulfiker
Sr. Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
FEBRUARY 2023

APPROVAL

This Project/internship titled **"Grape Leaf Disease Detection Using Deep Learning"**, submitted by Alif Jawat, ID No: 191-15-2649 and Md. Forhad Hossain Rahi, ID No: 191-15-2402 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 *05 February, 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

Raja Tariqul Hasan Tusher
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Mushfiqur Rahman
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Ahmed Wasif Reza
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Ms. Naznin Sultana, Associate Professor, Department of CSE** Daffodil InternationalUniversity. 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:

Ms. Naznin Sultana

Associate Professor, Department of CSE
Daffodil International University

Co-Supervised by:

05.02.2023

Mr. Md Sabab Zulfiker

Sr. Lecturer, Department of CSE
Daffodil International University

Submitted by:

05.02.2023

Alif Jawat

ID: -191-15-2649, Department of CSE
Daffodil International University

05.02.2023

Md. Forhad Hossain Rahi

ID: -191-15-2402, 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 **Ms. Naznin Sultana, Associate Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "***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 draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Professor Dr. Touhid Bhuiyan, Professor, and Head,** Department of CSE, for his kind help to finish our project and alsoto other faculty member and the staff of CSE department of Daffodil InternationalUniversity.

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

Even though a lot of research has been done in this area, classifying grape leaf diseases remains a very difficult task in the agricultural industry. This task is made more challenging by the spread of leaf diseases and the variations in their structural makeup. The scale-invariant feature transform (SIFT) and the histogram of directed gradients are two examples of manually created features that are used as the basis for practically all classifications nowadays (HOG). Convolutional Neutral Network (CNN), a feature learning technique, will be used in this instance. Here, the major objective is to train a model that has already been trained to find and extract the features that are essential for the given task. It finds several grape leaf diseases such as Black Rot, Esca, Leaf Blight, etc. Our goal is to create a deep learning- and image-based system for detecting grape leaf disease. The Convolutional Neural Network (CNN) method and Transfer Learning Implementation Technique are used in this approach to learning from a large number of unlabeled picture patches and a large number of healthy and infected leaf images. To detect the presence of infection in the grape leaf during the study, we will use some previously photographed diseased leaves and analyze the data using the deep learning algorithm Convolutional Neutral Network (CNN). A deep learning method will be applied to a sample dataset of the infected leaf in this project (CNN). For model building, we'll employ a few TensorFlow preprocessed models. Because this system will be able to readily identify various diseases, we hope that we will be able to create a system that will offer us more accurate results for detecting grape leaf diseases. Our approach aims to decrease grape leaf illnesses before they become contagious infected.

Table of Content for Research Project

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Introduction

We know that small things can have a big influence because of a disease in a grape field. There are many different types of diseases in grape leaves today. Some diseases are known, while others are unknown, but the majority pose a serious threat to grape crops and can often completely destroy them. One of the elements that affects grape productivity is disease and infection, which is maybe the most crucial one. Farmers also use insecticides that are not allowed to be used on certain diseases, which leads in unintended loss and unhealthy grapes, which has a significant negative impact on the GDP and growth of the nation.

Due of the significant economic yield loss, it is a significant problem for both farmers and the country's economy. According to a research by the Food and Agricultural Organization (FAO), diseases cause a loss of 5 to 80 percent of agricultural production worldwide each year, with an estimated cost of $105 billion [1]. Therefore, there is still a great need for research into disease detection in the current environment.

The fact that there are millions of pests means that each one is unique in terms of size, color, and shape, making identification and recognition difficult. Many techniques have been tried in the past to prevent disease invasion in grape fields, but they typically don't perform well. Despite some of those methods' usefulness, they are not as useful in actual areas. Therefore, in order to readily identify those infections and avoid unnecessary pesticide use, grape fields today urgently require some advanced technology. Because pesticides are equally harmful to the environment and grape plants.

In order to achieve precise control of infection in the grape fields, we wish to introduce some advanced agricultural technology that is employed in this field in the form of artificial intelligence, information, and wireless communication technology. By heavily utilizing machine learning and artificial neural networks, image processing systems have been able to tackle these kinds of issues. One of the greatest methods for spotting diseases or infections and determining which sickness it is from the image is the convolutional neural network (CNN).

1.2 Motivation

Only a few illnesses are responsible for a significant portion of grapes being damaged. That not only has an impact on the farmer but also on national development. This loss is unacceptably costly in a nation like ours where agriculture is the most important sector of the national economy. Thus, one of the factors that piqued our interest in working on this project was this. Additionally, we chose this issue for research because doing so will help us grow the desired crop and increase our agriculture sector's earnings. One of the most overlooked groups in society, farmers undoubtedly provide the most to the nation but receive little assistance or attention from the general public. In this situation, it would be almost intolerable for them to lose value due to the poor economy. Therefore, it is crucial that you work on this issue.

1.3 Rational of the Study

Most of the research done in this field of agriculture have hand-crafted features like the histogram of directed gradients and the scale-invariant feature transform (SIFT) (HOG). The drawback of this strategy is that the features are chosen manually by the person. The approach has a flaw in that we cannot be certain that the quantity of inclinations is an accurate criterion for categorizing pest photos.

However, in this case we'll build a deep learning model to find and extract the crucial features for categorization. Convolutional Neural Networks (CNN) are mostly used to find the most effective feature to categorize photos. They are particularly effective in this method since learnt features are automatically drawn out to carry out a specific task. This method performs better than the manual method.

1.4 Research Question

In our general research, there is an aggregate inquiry that is followed by some more questions that are divided into some more targeted queries.

1.3.1 What research has been done so far on detecting grape disease in Bangladesh?

a) What methodology was used to categorize the diseases?

I. What approach was used?

II. What infections were fought off?

b) How much accuracy was improved?

I. Which model was trained using the dataset?

II. Which model was most accurate?

c) How they used on the field?

I. Which information—local, national, or international—was used?

II. which organization was impacted?

1.5 Expected result

The major objective of this initiative is to find grape diseases. A model has been put out that can recognize input photographs with a high degree of accuracy after being trained on a sizable number of healthy and infected images.

1.6 Project Finance and Management

No financial support was given for this research project.

1.7 Report Format

The remainder of the paper's discussion is described in the way that follows. Chapter 2 discusses related work. In chapter 3, the methodology is displayed. Chapter 4 presents the experimental findings, a discussion, and specifics regarding our application. Chapter 5 shows the impact on society, the environment, and sustainability. Chapter 6 also offers a summary and a conclusion.

CHAPTER 2
Background

2.1 Preliminaries/ Terminologies

Infection identify, Diseases detection, Deep Learning, Processing image, CNN, Classification.

2.2 Related Works

Researchers, agricultural scientists, and agricultural development institutes around the world have used some approaches to detect grape diseases and infections with a view to agricultural development because it is thought that grape diseases have a massive effect on a large portion of agricultural development. Recent years have seen a rise in interest among researchers in studies that focus on deep learning.

For the purpose of identifying the presence of infection in an input image, Jerry Gao [2] developed a method for grape disease detection and classification. They gathered about 3670 pictures. They achieved 85.12% accuracy in this study.

In their study, Moh. Arie Hasan used a dataset of 1000 grape leaf photos from Kaggle, 800 of which were used for training and 200 for testing. A 0.0001 learning rate during the CNN training phase resulted to findings with an accuracy rate of 91,37% [3].

Prof. Girija J [4] used public data sets, 3,646 images were collected, and a total data set of 107,366 grape leaf images is generated through enhancement techniques applied to images.

The study on cotton leaf disease was proposed by Meunkaewjinda, A., et al. In his research, the researcher proposed a modified self-organizing feature map that employs support vector machines for classification and evolutionary algorithms for segmentation optimization. The segmented image is filtered with a Gabor wavelet, which enables the system to identify and classify cotton diseases more effectively by detecting and analyzing leaf disease color features [5].

For the Cotton Disease Control system, Hui Li et al. proposed the work based on the Web-Based Intelligent Diagnosis System. A system test is part of a research plan that used 80 distinct samples, including 10 of each type and samples from 10 different major disease species. Their project had an accuracy rate of 89.5% [6].

A method of Fungus/Disease Analysis in Tomato Crops Using Image Processing Technique was put out by Shruti and Nidhi Seth. In this study, a camera is used to capture an image of the crop leaves, which is then processed to create a segmented, grayscale image based on the type and size of the fungus. Using the level of fungal growth as a reference, acceptable and rejectable crop quality are set [7].

The proposal for Detection of Diseases on Cotton Leaves and Its Potential Diagnosis was made by Ajay A. Gurjar and Viraj A. Gulhane. They came to the conclusion that, depending on the image quality, the system can accurately detect diseases in 85 to 91% of cases [8].

2.3 Comparative Analysis and Summary

The three study papers [1–3] deal specifically with grape diseases. However, they used a different plant, like cotton, and a different type of crop leaf in the papers [4] to [8]. However, the basic goal remained the same. The approach and technique of execution are likewise comparable. The accuracy of the second publication [2]'s research, which utilized more than 3,000 photos, was 85.12%. They used the CNN algorithm in their third paper. 800 photos were used for training, while the remaining 100 were used for testing. Their project's accuracy is 91.37%. In the fourth publication, they employ 107,366 picture dataset for enhancing approaches and 3,646 public image photos. In the papers [5] to [8], they worked with a different plant, but their method of disease identification for grapes is comparable to theirs. They worked with a cotton plant on their project for the paper [8], and their project's accuracy ranges from 85 to 91%.

2.4 Scope of the Issue

Agriculture is the main industry in our nation. In an agricultural nation, there are several potential applications for plant disease detection. Every year, a paddy pest causes a large amount of grain to be lost in the grape field. In addition, there are various other fields, such as cornfields, cotton plants, wheat fields, and other fruit fields, such as other fields fruits, etc.

2.5 Challenges

The collecting of photos of various types of infections and in big quantities is one of the most difficult issues with that research. The caliber of the photographs is another area of concerns. The photos' poor resolution makes it difficult to easily identify disorders. As a result, detection accuracy will decline. Another difficulty is the variety of the photos within a single illness.

CHAPTER 3

Research Methodology

We divide our information into two groups in this research to make it easier to discover diseases. We used the Kaggle dataset. The training dataset is one of them, while the test dataset is the other. We started by enhancing both of our datasets. It is crucial to take all of the photographs with the same attributes in order to train all of our images and fit our dataset into a model. The CNN model's training is hence of utmost importance. Using our augmented data, we work on it. Finally, by fitting our test dataset to our training model, we calculated our accuracy.

3.1 Research Subject and Instrumentation

Our research's primary goal is to categorize diseases. It is challenging to classify diseases. because leaf infections are so widespread in the world. It takes a lot of picture data to categorize infection from such a huge group. because this paper's image processing research. We have categorized the disorders into three groups. They are Leaf Blight, Black Rot, and Esca. We need to manually take a lot of images for this. We visited several recharging facilities and took pictures in their settings.

3.2 Data Collection Procedure/Dataset Utilized

The collection of data is the most crucial aspect for our project. The majority of the information in our dataset was gathered from Kaggle.com. Data has been gathered from a number of websites. Additionally, we personally took images to acquire the remaining data, making the dataset significantly larger. The accuracy increases with the size of the dataset.

We started by gathering several photographs of various diseases that were infected. Here, we made an effort to compile information on every grape disease known to cause agricultural harm, skipping over certain species that are not at fault. To make calling this data for training simple, we put them in a separate folder. Our data was obtained via Kaggle. Figure 1 shows the data set for diseases used in training.

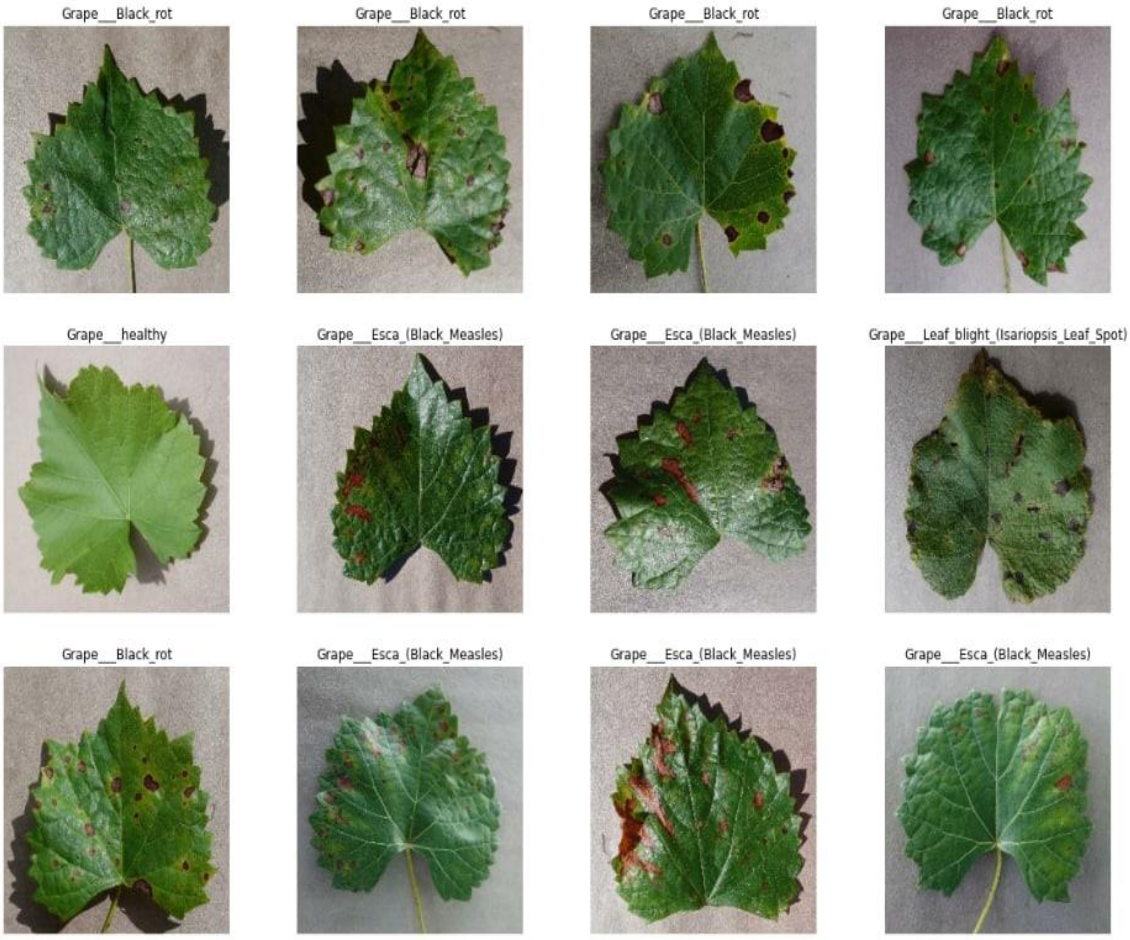

Figure 1: Images of several kinds of diseased leaves.

3.3 Statistical Analysis

In our study, we used images of three different diseases: Leaf Blight, Black Rot, and Esca. We divided the dataset into train and test categories for the model's evaluation. In the training dataset, we used about 4000 images in each category, while in the test dataset, we used 10% of the images for testing and 10% of the images for validation. Our dataset now contains 4000 photos, which is sufficient to provide greater accuracy. To boost accuracy, the dataset has been increased by augmenting it.

3.4 Proposed Methodology / Applied Mechanism

Deep learning is an extremely effective technology. It is capable of managing large data sets. The term "Convulsive Neural Network" is also known as CNN. Most often, it is used in image processing. In the 1980s, studies on the "Convulsive Neural Network" were conducted. Ideas regarding how neurons in our visual cortex function are included in the evolutionary network. This network's architecture differs from that of typical neural networks because of this. The term "convolution" is derived from the mathematics concept of combining two functions to produce a third function. The process of combining two kinds of input results in a convulsion.

Enhancement refers to growth. Images can be changed via a process called data augmentation. The model's training is deep learning's most crucial task. which is accomplished using the dataset. The more data we can use to train a model, the more accurate results we will acquire. To further improve and enrich our dataset, we have added our data.

After completing the above processes, the critical task for us is to re-train our pre-train model with our augmented data, which we have fitted to the expanded data of the model.

After finishing the model's pretraining, we evaluated the test dataset against the model to determine its correctness.

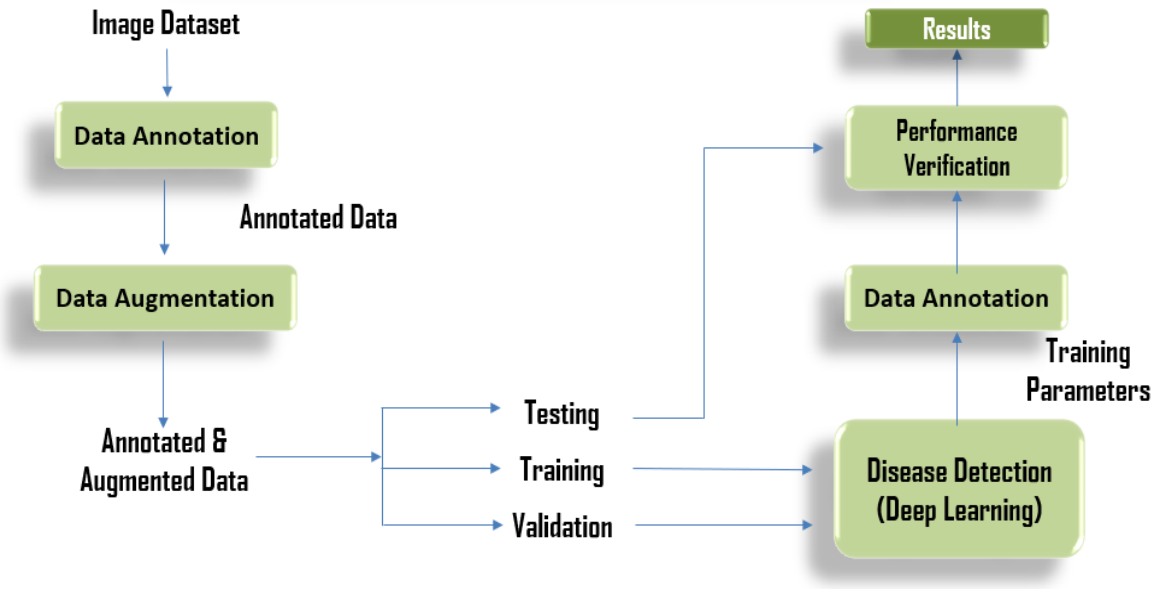

Figure 2: A flowchart of augmenting and training.

3.5 Augmentation diagram summary

This diagram shows the flow diagram for training and augmentation. We can see from this diagram that we first train the dataset, then we train preprocess, then we train augmentation, then it tries to fit in pretrained model, then it learns from the model, then it evaluates the dataset from the test dataset, and last it gives the outcome.

CHAPTER 4

Experimental Results and Discussion

4.1 Experimental Setup

We look forward to describing our model's entire approach in this part. It was created to address a fairly common issue in the actual world. For this situation, we suggested a CNN-based design. We have utilized the 53-layer, pre-trained MobileNet-v2 model from ImageNet. Typically, an inverted residual structure underlies our model. We have two blocks in our model, the first of which is the residual block. Eight layers of the two blocks make up this model. the capacity of a linear classifier on the output domain's non-zero volume region. Here, one of the outputs is the linear classifier's capacity for non-zero volume. In this case, the expansion factor t is the same across all of the major experiments. The internal output has n*t=nt channels if the input has n channels.

4.2 Experimental Result and Analysis

Table 1 shows our accuracy and loss values for our model's training and validation. where the accuracy and loss values for the 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10th epochs are displayed for training and validation. The values of epochs show us that anytime the accuracy grows, the quantity of loss lowers since our model is now more efficient. We were able to get roughly 95% of the accuracy from our models as a result. Below, with 6 epoch, are the test accuracy, training loss, training accuracy, validation loss, and validation accuracy for our system.

Epoch	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
01	1.3260	0.5068	0.7034	0.6953
02	0.5925	0.7622	0.4679	0.8177
03	0.3640	0.8557	0.2461	0.9089
04	0.2698	0.8969	0.1938	0.9219
05	0.1962	0.9257	0.2820	0.8984
06	0.1658	0.9424	0.3792	0.8568

Table 1: Losses and model accuracy on training and validation data.

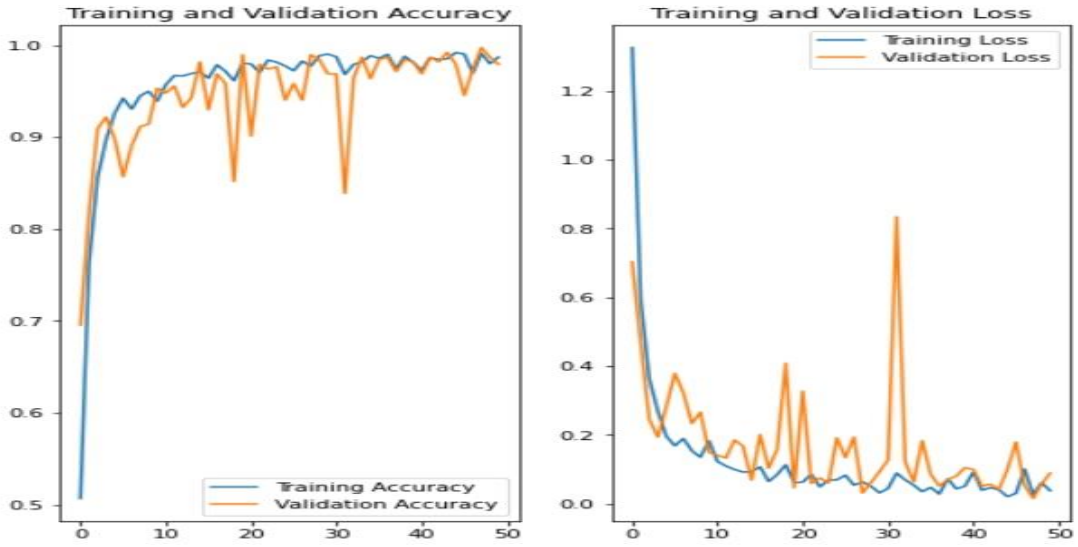

Figure 3: Accuracy and loss differences between training and validation.

Here, we can observe that the accuracy throughout training and validation is roughly comparable. The validation accuracy rose along with the training accuracy. However, the training loss also occurs at the same time as the validation loss.

4.3 Application Features

In this project, we develop a website to disclose our project. Using our application we can call an image and can insert it into our project. Then our application shows the output using our project in the background.

In our application, we have used some features like ;

i. Using the back-end, call the primary project and insert an image.
ii. In the application, display the output result.
iii. Display the accuracy and additional information in the outcome.

Figure 4: View of our application for adding images and results.

4.4 Discussion

In our evaluation, we used the sequential model. This model has been pretreated. With our dataset, we retrained this model independently. Table 1 shows the data we determined from its history. The results show that as the number of epochs has increased, the validation accuracy has also increased. Up to 10 epochs of our model's deployment resulted in 90% accuracy. We will receive better results from our model if we increase the number of epochs. We employed additional models, but the Sequential Model provided the highest level of accuracy. Sequential is the best model for our dataset.

Because of the poor image quality and the little number of photographs in our dataset, our algorithm was unable to quickly identify infections. The photograph is frequently overexposed and difficult to read because of the sunshine, and vice versa when there is no sunlight. We frequently found that the background was so similar to the infection itself that our model was unable to identify the infections. a few images where our model had trouble spotting infections. This precision was decreased to 40–50%.

CHAPTER 5

Impact on society, Environment and Sustainability

Grape is one of the most popular fruit among all the fruits in Bangladesh. It is a source of many types of food nutrient . It has Potassium which can help bring high blood pressure and lower risk of heart disease . So we can say that, the social demand of grape is more than we think . It is grown commercially in our country .

5.1 Impact on Society

We know, Grapes are cultivated commercially for economical benefits . As much we produce grapes, it will be more beneficial for the society as well as farmers . But there are various types of grape diseases which reduce the production ratio and most of the farmers have not proper knowledge about this kind of diseases . That's why we are trying to develop a system, which will help the farmers to detect the diseases easily . As a result, farmers can prevent these diseases and can ensure a good production . By above discussion we are hopeful with our project that it will create a great impact on our society .

5.2 Impact on Environment

Our detection system can create a good impact on environment as well . For example, there are three common diseases of grapes which are Black Rot, Esca (Black Measles), Healthy, Leaf Blight . By using our system, we can easily detect specific diseases . Which help farmers to find out the proper cure of the specific disease . That can reduce the number of infected plants . If we can reduce the number of infected plants, then it will also help to reduce the infection possibility with the other plants .

So, we can say that our project will create a great impact on our Environment .

5.3 Ethical Aspects

First of all, we will not work with any individuals data . That's why, there's have not any confidentiality issues . Rather than our initial project have not any authentication system. As a result anyone can use and take benefits from our projects without sharing their personal information . In future, we will add authentication system like "Login and Logout" and also create user profile and that's time we also keep the user information hidden from everyone else.

5.4 Sustainability Plan

There are few steps of sustainability plan . Some are :

i. Identify what resources are required : We need some technologies for Frontend, Backend and deep learning based model for train . Again, we don't need as much money for the project in the current situation.

ii. Create case statements : We need to discuss with the people like farmers and need to introduce them with this project . Finally we have to take feedback from them and make our project more sustainable according to their feedback .

iii. Create an action plan : We already read five to six paper which based on Grape Leaf Disease Detection . And also, Our model training process is running and we are working on the field stage for collect user reviews to make a user friendly system .

CHAPTER 6

Summary and Conclusion for Research

6.1 Summary of the Study

In order to conduct our research, we first organized the data, pre-trained a model using the data, and then re-trained it. We validated our test data and its level of accuracy after retraining these data.

6.2 Conclusion

Despite extensive research in this area, millions of metric tons of grapes are wasted due to a few infections and illnesses, which create enormous losses in the agricultural sector. The farmers in rural areas are the biggest victims of the loss. From this perspective, in this study, we suggest a CNN (sequential Model) based deep learning strategy to identify illnesses and infections in grape leaves. Three diseases are correctly recognized by our model from a taken image. Our suggested model achieved 90–55% accuracy. We are hoping that our suggested model may have some effect on this circumstance and lessen crop loss from insects or pests.

6.3 Future Research

In this research we working on grape leaf disease detection for this work we collected data from Kaggle in this time but we have a plan to use real data in future and in this time, we work only on leaf disease field but we have a plan to work or more field like cotton and bean tree pest detection using drone, flower disease detection etc. In this work we don't get enough accuracy for this reason we want to improve our model and also want to apply more algorithms to improve accuracy. For the instant solution farmer need mobile application. For this reason we have a plan to design a mobile based application with an easier user interface so that farmer can easily find out anything. Further we also have a plan to include an automated crop heath detection system it will be very easier for the farmer to identify the crops health then they will get an idea that what to do next.

REFERENCES

[1] Grapevine Trunk Disease: The Destructive FMC; https://ag.fmc.com/us/sites/us/files/2020- 07/19-FMC 2337_RhymeGTD_FactSheet_011520_VF_LoRes_Ref%20%281%29.pdf

[2] Jerry Gao,Department of Computer Engineering,San Jose State University,San Jose, United States jerry.gao@sjsu.edu (corresponding author)

[3] Identification of Grape Leaf Diseases Using,Convolutional Neural Network,Moh. Arie Hasan1∗, Dwiza Riana2∗, Sigit Swasono3, Ade Priyatna4,Eni Pudjiarti5, Lusa Indah Prahartiwi6,Program Studi Ilmu Komputer STMIK Nusa Mandiri1−6

[4] A Review on Grape Disease Detection and Classification using ImageProcessing, Niharika S1, N V Sahana2, Prof. Girija J,3,1,2Dept. of Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru, India, 3Professor, Dept. of Computer Science and Engineering, Bangalore Institute of Technology, Karnataka, India

[5] Meunkaewjinda. A, P.Kumsawat, K.Attakitmongcol and A.Sirikaew.2008 Grape leaf diseasefrom color imaginary using Hybrid intelligent system", Proceedings of ECTICON.

[6] Hui Li, Ronghua Ji, Jianhua Zhang, Xue Yuan, Kaiqun Hu and Lijun Qi, "WEB-Based Intelligent Diagnosis System for Cotton Diseases Control" IFIP Advances in Information and Communication Technology, 346:483-490. 2011.

[7] Shruti and NidhiSeth,"Fungus/Disease Analysis in Tomato Crop using Image Processing Techniques."International Journal of Computer Trends and Tecn- ology (IJCTT),volume 13 number Jul 2014.

[8] Mr. Viraj A. Gulhane, Dr. Ajay A. Gurjar, "Detection of Diseases on Cotton Leaves and Its Possible Diagnosis ", International Journal of Image Processing (IJIP), Volume (5) : Issue (5) 2011.

Proj_Rep

ORIGINALITY REPORT

14%	%	%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	11%
2	Submitted to Jacksonville University Student Paper	2%
3	Submitted to Savitribai Phule Pune University Student Paper	<1%
4	Submitted to University of Greenwich Student Paper	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		