SUNNET: A DEEP LEARNING APPROACH TO DETECT SUNFLOWER DISEASE

BY

TASLIMA AKTER SATHI
ID: 191-15-2669

AND

MD. ABID HASAN
ID: 191-15-2603

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

Supervised By

Mohammad Jahangir Alam
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md. Sabab Zulfiker
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

04 FEBRUARY 2023

APPROVAL

This Project titled **"SunNet: A Deep Learning Approach to Detect Sunflower Disease"**, submitted by Taslima Akter Sathi, ID No: 191-15-2669 and Md. Abid Hasan, ID No: 191-15-2603 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 in *04 February 2023*

<u>BOARD OF EXAMINERS</u>

7/2/23

Dr. Touhid Bhuiyan **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

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

Internal Examiner

Dewan Mamun Raza
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

4/2/2023 **External Examiner**

Dr. Shamim H Ripon
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 **Mohammad Jahangir Alam, Senior Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Jahangir 06.02.23

Mohammad Jahangir Alam
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

06.02.2023

Md. Sabab Zulfiker
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Sathi

Taslima Akter Sathi
ID: 191-15-2669
Department of CSE
Daffodil International University

Abid

Md. Abid Hasan
ID: 191-15-2603
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 **Mohammad Jahangir Alam**, **Senior Lecturer**, 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 drafts and correcting them at all stage have made it possible to complete this project.

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

Helianthus annuus, often known as sunflower, is a crop that is only mildly affected by drought. The agricultural sector of the economy benefits greatly from this. However, various illnesses have imposed a halt on sunflower cultivation over the world. However, many severe diseases will have affected plants if corrective measures are not taken sooner. Therefore, it will have a negative impact on sunflower yield, quantity, and quality. Diagnosing a disease by hand can be a time-consuming and difficult process. Object recognition methods that use deep learning are becoming increasingly commonplace today. In this study, we put out a strategy for identifying diseases in sunflowers. A total of 1428 photos were utilized to complete this task. Images have also been processed using methods like resizing, adjusting contrast, and boosting color. We have segmented the area of the photos afflicted by the disease using k-means clustering, and then retrieved characteristics from those regions. Five deep learning classifiers were used to complete the classification. For the purpose of comparing classifier quality, we computed four performance evaluation measures. The best performing classifier overall was a ResNet50 classifier, which had an average accuracy of 97.88%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Sunflower (Helianthus annuus) is an Asteraceae plant with hundreds or thousands of small florets. Sunflowers have economic and decorative importance [1]. Both the leaves and the oil extracted from the seeds can be used in cooking, with the former serving as a substitute for wheat and the latter for olive and almond oils. Feeding utilized poultry might also include sunflower oil cake. Also used in cosmetics, cleaning products, paints, and ointments, the oil has a wide variety of applications. In spite of Russia and Ukraine being the world's leading sunflower producers, sunflowers are grown in much of the world [2]. A total of 11 and 12 million tons of are produced in Ukraine and Russia, respectively.

Sunflower production is falling at an alarming rate due to the prevalence of illnesses. Gray mold, downy mildew, leaf scars, leaf rust, phoma blight, septoria leaf spot, etc., are just some of the most common diseases that attack sunflower plants. Most farmers are not well-versed in the modern equipment that assist them to generate massive yields with fewer damages because they are illiterate and live in rural areas. A farmer may miss the first sign In this research, we present an automated technique for identifying sunflower diseases. Using this method, we were able to categorize sunflower diseases such as Leaf Scars, Gray Mold, Fresh Leaf, and Downy Mildew. We used a myriad of preprocessing methods to resize and improve the image's quality during this stage. For this reason, we use a k-means clustering-based image segmentation technique to extract the desired areas from the overall image.

1.2 Motivation

Bangladesh's economy is mostly based on agriculture, but it is still unable to produce enough food to meet the needs of its large population. When a pathogen infects a leaf, most farmers have no idea what to do. That is exceedingly harmful to one's health. The production of crops that are free of disease is vital. The population is the driving force behind any nation's economy. When people are healthy, they are better able to work and contribute to the expansion of the economy. As a consequence of this, the farmer will gain a better understanding of how to rapidly diagnose plant illnesses as a result of this research.

The goal of this project is to establish a forum that would facilitate communication between farmers and agricultural experts at various levels. It will also make it simpler for farming specialists to travel to far-flung corners of the country in search of new diseases. This will help farmers grow more crops, making the economy stronger in the long run.

1.3 Applied Classifier

In this study, we'll explore various approaches to modeling the universality of plant tissue disease. The primary goal of this model is to minimize losses in damaged crops so that output can be increased. With the help of the CNN algorithm, six models have been used in this project to find plant diseases:

- ResNet50
- VGG16
- MobileNet V2
- Inception V3
- Xception

1.4 Research Questions

The questions that this thesis primarily addresses are as follows:

- Can you tell me about the most widespread diseases that attack sunflowers?
- Where can I have these disorders treated?
- How can image processing models be used to help detect plant diseases in advance?
- Where can I read up on the most effective strategy for training with visuals?
- Where will these advantages be felt, and by whom specifically?
- What is the model's accuracy potential?

1.5 Expected Outcome

The participants in this research were tasked with developing a cutting-edge method that would assist users in recognizing afflictions that affect plants and administering any necessary treatments. The primary objective of this piece of software is to identify plant diseases by analyzing user-submitted pictures of sick leaf tissue. Users have access to a wide selection of extra options that the program offers and can select any of these at their discretion.

The single most significant realization that we've made is:

- Using our method for early disease detection in sunflowers could make it much less likely that the crop would lose a lot of money.
- Only ResNet50 achieved acceptable results among all the models. greater than average sensitivity and identification precision compared to competing models.

1.6 Report Layout

In Chapter 2, we shall talk about the history of this study and the other relevant research. We'll go into more depth on research techniques, data collection, and other nitty-gritty in Chapter 3. In Chapter 4, you'll look at the information you've collected and think about the best model for your project. In Chapter 5, we will talk about the final results and what comes next.

CHAPTER 2
BACKGROUND STUDY

2.1 Introduction

The importance of the visual in 21st-century science, art, medicine, business, education, and daily life cannot be overstated. useful in almost any circumstance. We rely heavily on visuals to convey information. Thousands of lines come together to form an image. Each passing day brings new technological marvels and, with them, increasingly intelligent machines. Computers have no trouble understanding visuals. Through the use of image processing, a machine may do tasks independently. Diverse forms of technology are used for such tasks. One option for classifying pictures is the use of convolutional neural networks. Observing the plant's leaves is the most reliable method for determining whether or not it has been infected. Some other symptoms of a diseased plant include the following. Identifying these, however, is a challenge unless a farmer conducts thorough testing on the plants. It's possible to learn a lot about disease just by looking at the leaves. Farmers can't tell the differences between the symptoms of different diseases. Even though there are simple treatments for most diseases, even a minor mistake could lead to a significant loss of crops.

Recent years have seen a dearth of studies in Bangladesh concerning smart agriculture systems and the identification of plant diseases. There are, however, numerous initiatives afoot to usher in a paradigm shift in the country's agricultural sector using agro-based methods.

2.2 Related Work

An approach for the "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" was proposed by S. S. Sannakki and V. S. Rajpurohit[3]. This approach relies primarily on segmenting the defective area, with color and texture serving as characteristics. The categorization, in this case, was performed by a neural network classifier. Categorization is determined to be accurate 97.30% of the time, and the key benefit is that it converts to L*a*b to extract chromaticity layers of the image. The fact that it can only be utilized for a small number of crops is probably its biggest drawback.

P. R. Rothe and R. V. Kshirsagar [1] first presented cotton leaf disease identification using pattern recognition techniques. This method uses snake segmentation, and in this case, Hu's moments are used as a way to tell them apart. Active contour models like BPNN classifiers are used to limit the liveliness within the infected area by reducing the amount of space available for it to spread. The overall categorization was found to be 85.52% of the total.

Machine learning techniques were used by A. Majumder et al. [2] to create a system for identifying diseases in carrots. Using a k-means clustering algorithm, scientists could divide the area hit by the illness into several sections. They next used a support vector machine (SVM) classifier to sort the data. They combined 11 feature sets with a total of 202 photos. The success rate of this method was 96%. It has been pointed out, though, that bad image quality and changes in the color of the background can be distractions that make it hard to get accurate results.

In [3], S. Sasirekha et al. described image processing for the diagnosis and classification of carrot rot. They started by shifting the images from RGB to L*a*b. The k-means clustering method was then applied for the purpose of segmentation. With the help of texturing and classification methods, we were able to extract a total of 13 characteristics. Carrot illnesses were subsequently identified by a classification process using multiclass SVM.

It was proposed by H. Zhu et al. [4] to use deep learning to determine the quality of a carrot's exterior. AlexNet has been used to analyze carrot photos for characteristics and grade the quality of the carrots. In this study, we found that a recognition rate of 98.70% was possible for binary classes.

Using deep learning, Rupali Saha et al. [5] devised a method for identifying orange fruit with a disease. CNN was utilized for classification, and eight different feature sets were applied to a 68-item dataset. They said their method was 93.21% accurate.

The papaya disease detection system used by M. T. Habib et al. [6] was developed using a machine learning approach. They used k-means clustering to divide the area hit by the disease into manageable chunks, and then SVM classification was applied for further analysis. The dataset size was 126, and 10 feature sets were employed. The method had a 90.15% success rate.

Identifying imperfect areas of produce was the focus of a method given by L. J. Rozario et al. [7]. Four different kinds of produce were tried out. For color-based image segmentation, they turned to a combination of modified k-means clustering and the Otsu approach. There were 63 photos utilized in total for this study. It should be noted that this effort did not involve any sort of categorization.

Rajbongshi et al.[8] proposed an approach for Sunflower Diseases Recognition Using Computer Vision-Based Approach. They used 650 image data and separated them into training data 80% and testing data 20%. They used the seven performance evaluation measures to assess each classifier's performance. The Random Forest classifier that outperformed others had the highest average accuracy of 90.68%.

U.Fulari et al.[9] suggested a strategy called Three-channel convolutional neural networks for detecting vegetable leaf diseases. The Grey Level Cooccurrence Matrix (GLCM) is employed for feature extraction. The categorizing process uses a machine learning technique known as Support Vector Machine (SVM). The recognition accuracy was improved when the Convolutional Neural Network (CNN) method was compared to the SVM method. 12,949 images from the dataset are utilized to train the system. Using the 80/20 splitting ratio, the database is divided into two datasets. The image dataset includes healthy and ill crop leaves; CNN has provided 97.71% accuracy, which is higher than the accuracy attained using hard coding techniques.

P. Bedi and P. Gole[10] conducted a method to detect Plant disease using a hybrid model based on a convolutional autoencoder and convolutional neural network. The collection has 4457 peach plant leaf photos, evenly split into two categories: healthy and diseased (Bacterial Spot). With just 9,914 training parameters, the model's training accuracy was 99.35% and its testing accuracy was 98.38%.

N. Ganatra and A. Patel [11] introduced a method for Fruit and Vegetable Identification Using Machine Learning for Retail Applications. The experimental phase and the implementation phase are the two parts of the project. From ImageNet, a dataset with 400 photos per class has been taken. The camera used for this project has also captured 30 photographs per lesson. When using the top 5 ranks, MobileNet achieves a 60% accuracy rate out of 10 photos. MobileNet's overall top-three accuracy is 97%.

A.Bantan et al. [12] proposed a method for discriminating between sunflower seeds utilizing multispectral and texture datasets in conjunction with region selection and supervised classification approaches. The six classes Syngenta CG, HS360, S278, HS30, Armani, and High Sun 33 were used to classify a collection of sunflower seed cultivars. 53 multi-features are collected from each of the regions that are chosen for analysis in their novel region-oriented seed-based segmentation. Finally, accuracy rates of 98.2%, 97.5%, 96.6%, and 94.8% were reached with an ROI size of (180 180).

Supriya and A. Shukla [13] proposed a method to identify the four types of paddy leaf disease using MATLAB R2015a image processing tools.The proposed model, which pre-processes an unseen image and extracts its features for CSV input to the classifier to categorize them into healthy and unhealthy categories, achieved an accuracy of 97% using SVM in this step. By accounting for the train-test set to be 80% - 20%, the Machine Learning algorithms used in this work enable us to detect a diversity of samples with high accuracy.

S. Ramesh and D. Vydeki [17] introduced a technique for using a machine learning algorithm to detect disease early in the agricultural process. The simulation results show that the KNN-based classification technique has an accuracy of 86% for conventional leaf shots and 85% for blast-affected leaf photos. The ANN-based classification techniques now have accuracy scores of 99% and 100%, respectively.

A. Malik et al. [18] propose an Approach for the Design and Evaluation of a Hybrid Technique by Using Deep Learning to Detect Sunflower Leaf Disease. In this study, they utilized a data set that consists of 329 sunflower images classified into five categories and was created by the author with the help of Google Images. The stacking ensemble learning approach is used to combine VGG-16 and MobileNet, two transfer learning models that are employed for classification applications. After calculating the accuracy of these models, including AlexNet, CNN, DenseNet-121, Inception V3, ResNet-50, ResNet50V2, ResNet-101, and VGG-16, it can be said that the proposed technique provides better accuracy than the others. The final step is to compare the proposed technique to the existing approaches. The proposed method outperforms the competition with an accuracy of 89.2% using the same set of data.

M. Zabeeulla A N and Dr. C. Shastry[19]developed a technique to automate a framework for predicting leaf disease in several crop species using machine learning and deep learning. The suggested study work uses three different frameworks to identify plant leaf disease with the highest degree of accuracy, the lowest possible processing time, and a true positive rate. Three distinct sections comprise the developed research study.

R. Dawod and C.Dobre[20] suggested a technique for the Automatic Segmentation and Classification System for the Classification of Foliar Diseases in Sunflowers. The residual neural networks ResNet50 and ResNet152 were used to categorize diseases based on lesions. At least one impacted area has been split in more than 90% of the photos. In the event of diseases like powdery mildew, where the entire leaf turns pale, segmentation is more challenging to execute. In 30% of the photos, diseased regions could not be segregated.

S.Sladojevic[21] proposed an approach for classifying leaves in order to identify plant diseases using deep neural networks. With the ability to differentiate between plant leaves and their surroundings, the developed model can identify 13 different types of plant illnesses from healthy leaves. For various class tests, the experimental findings had a precision of between 91% and 98%. The trained model's ultimate total accuracy was 96.3%.

M.Brahimi et al. [22] suggest a technique for visualizing tomato diseases using classification and symptoms using deep learning. This study utilized a massive dataset in comparison to the state-of-the-art. 14,828 images of tomato leaves with nine distinct ailments are included in this collection. They have introduced the Convolutional Neural Network (CNN) learning approach to training their classifier. The findings are encouraging and can be used by farmers as a helpful tool to protect tomatoes from illness. They outperform significantly shallow models with 99.18% accuracy.

A.MEUNKAEWJINDA et al. [23] developed a method for detecting grape leaf disease using a convolutional neural network. Using a modified self-organizing feature map, support vector machines are used to classify the data and genetic algorithms are used for optimization. Finally, the segmented image is filtered using a gabor wavelet, allowing the system to analyze the color aspects of leaf illness. The ability to remove confusing color pixels from an image's backdrop is subject to several restrictions.

P. TM et al. [24] constructed a technique for detecting tomato leaf disease using convolutional neural networks. In this work, 54,306 photos of 14 crops plagued by 26 diseases are used. Around 18160 photos of tomato leaf diseases are included in the subset. The average accuracy of this proposed system is between 94 and 95 percent, demonstrating the viability of the neural network approach even in unfavorable circumstances.

S. Zhang et al. [25]created a method for Vegetable leaf disease detection using three-channel convolutional neural networks. 100 photograph from each of the five types of cucumber diseases Scab Angular, Powdery Mildew, Downy Mildew, Anthracnose, and Scabare chosen for the experiment's 500 diseased cucumber leaf images. To verify the efficacy of the suggested method, every original image of a diseased leaf is enhanced 14 times, then split into a training set and a test set with varying percentages of the photos. The lesion segmentation and manually created feature process are not used in the suggested strategy. The experimental outcomes show that the multi-channel CNN is efficient and practical.

In this way, the study can serve as a reliable starting point for additional research. In this part of the literature review, the authors talk about different ways that people have come up with to spot harmful organisms. The main purpose of this study is to diagnose illnesses and provide better data for farmers, and the literature review paper is related to that objective in some ways.

2.3 Comparative Analysis

After looking at several research papers and projects, it was decided that CNN would be the best option for this endeavor.

- It works well for classifying forests and other natural environments, recognizing and categorizing eco-friendly products, and detecting environmental sights.
- It compares images and is useful in identifying plant diseases.
- It is user-friendly and comes with several resources to aid in its continued development.
- It could help contrast different illnesses and leaf varieties.

2.4 Scope of the Problem

Sunflower (Helianthus annuus L.) is an important oilseed crop that is widely grown in different regions of the world. However, the crop is vulnerable to various diseases caused by pathogens such as fungi, bacteria, and viruses. These diseases can cause significant yield losses and affect the quality of the seeds. Therefore, the identification and control of sunflower diseases is crucial for sustainable crop production.

The scope of this research problem encompasses the following key areas:

Identification and classification of sunflower diseases: This includes determining the causative agents of the diseases, their symptoms and signs, and the conditions that favor their development.

Detection methods: Developing accurate and efficient methods for the detection of sunflower diseases in the field, such as visual inspections, laboratory analysis, and the use of remote sensing and machine learning techniques.

Disease management: Investigating the most effective ways to control and prevent the spread of sunflower diseases, including the use of chemical and non-chemical methods, and the development of resistant cultivars.

Economic and environmental impacts: Evaluating the economic and environmental consequences of sunflower diseases and identifying ways to minimize these impacts.

New and emerging threats: Keeping track of any new or emerging diseases that may pose a threat to sunflower production and developing strategies to mitigate their impact.

Overall, this research aims to provide a comprehensive understanding of the challenges facing sunflower disease detection and management, and to develop practical and sustainable solutions for the benefit of farmers, industry, and society.

2.5 Challenges

To develop such a system, some of the features will initially need to be trained, and then further data will have to be gathered in order to evaluate the model. A number of different hospitals' records were used to compile the data.

A substantial obstacle is presented by the strategy's lack of completeness, inconsistent data, and redundant information. Imbalanced data is difficult to process. It is impossible to arrive at an accurate calculation using data that contains redundant information. The pre-processing of the data, which must be clarified first, is more crucial to the assessment of the more accurate data. The process of sorting through the data is also quite important.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

The process of diagnosing illnesses in sunflowers is broken down into numerous stages. Here, we'll go down each stage in great depth. Identifying sunflower diseases follows the process shown in Fig-1.

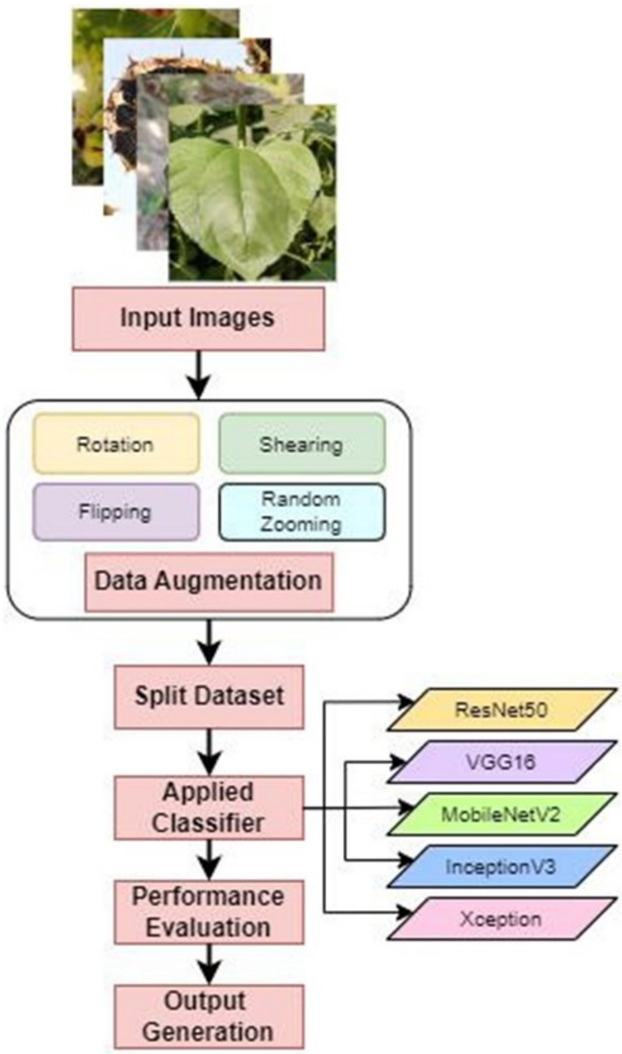

Fig 3.1.1: Procedures for identifying illnesses in sunflowers.

3.2 Image Collection

The process of acquiring images is the first step in the operation of any computer vision system. The act of extracting an image from its source is what is meant by this term. If both the overall image quality and the total number of photos are of a sufficient size and level of excellence, respectively, it will be possible to attain the desired result. Within the scope of this study, we gathered several types of picture data for sunflowers at the field level. There are four distinct illnesses that can affect each of the four varieties. These diseases are referred to as leaf rust, leaf scars, gray mold, and downy mildew. In addition, there is a class for disease-free sunflowers. Table-I contained the information about class wise data.

TABLE-3.2.1: DATASET OVERVIEW

Class	Original Data	Augmented Data	Total Data
Leaf Scars	120	212	332
Gray Mold	72	386	458
Fresh Leaf	134	186	320
Downy Mildew	120	198	318

3.3 Pre-processing

It is vital to preprocess the acquired images to effectively carry out a certain activity that is related to eyesight. A wide variety of processes and procedures can be utilized to carry out image preprocessing. The steps of resizing the image, increasing the contrast, and changing the color are included in the preprocessing stage for images. The scaling of the obtained images is an important step in the preprocessing phase of computer vision. In general, the training of deep learning models can be completed more quickly on images that are smaller. The augmentation of contrast is a necessary step that must be taken in order to produce a higher-quality image. The acquired image needs to have its colors converted in order to be used effectively in subsequent steps. Because the photographs that have been collected come in a variety of sizes, it is essential to resize them to the appropriate proportions. Firstly, the photos were scaled down to a set length of 224*224 by utilizing bicubic interpolation. Assume I and f progressively, and the derivatives are going to be fx, fy, and fxy. These are going to represent the four corners of a unit square, which are going to be (1, 1), (1, 0), (0, 1), and (0, 1), respectively, and mij is going to stand for the coefficients. The stability of the interpolation surface, designated by the number [18], is characterized by the following:

$$f(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} m_{ij}x^{i}y^{j} \qquad (1)$$

3.4 Image Segmentation

The approach of separating an image into a variety of segments is known as the "image segmentation technique." It is considered an extremely important and critical step in the field of computer vision. The goal of image segmentation is to take a representation of an image and convert it into something that is both more meaningful and easier to understand. It is routinely conducted for the sake of discovering objects and establishing boundaries. Working with a full image is not a particularly brilliant notion due to the fact that different components of an image might not include any information that is of any use. As a result, through the process of image segmentation, we are able to do processing on only the relevant areas. Using the k-means clustering method, we carried out image segmentation on the sunflower picture data that was provided to us for this work.

k-means is an unsupervised method that may be used to find separate clusters in the data that is provided on the basis of the degree to which the data is related. This particular clustering approach is one of the most widely used, where k signifies the total number of groups to be clustered. In this study, we have set the k parameter for segmenting a picture to equal 3, which means that it will differentiate between three different clusters inside the image. The K-Means clustering algorithm works quite effectively with relatively small datasets.

3.5 Data Set Splitting

To develop CNN architectures, we splitted our data into 75:15:10 ratio. We used 75% to train the model 15% for validation and finally 10% data for testing.

3.6 Applied Classifier

For the purpose of diagnosing diseases affecting sunflower, researchers utilized a total of five distinct deep learning classifiers, namely ResNet50, vgg16, MobileNet V2, Inception V3, and Xception. Sunflowers are susceptible to a variety of diseases, and these sorters make it possible to diagnose them.

Foundational Base: CNNs are constructed using many layers of artificial neurons [10]. The most common use for CNN is the analysis of visual data with DL algorithms [21]. Convolutional layers, input layers, pooling layers, fully connected layers, hidden layers, and activation functions were used to build the core architecture of CNN models. A CNN structure will be created as a result of stacking these levels one on top of another. The primary components of CNN's architecture are the selection of features and the categorization of those features.

ResNet50: An extremely sophisticated residual network, ResNet50, is a deep neural network. The classification process consists of two sets: the training set and the test set. Each data instance in the training set has a number of features and a single goal value. In all, there are roughly fifty levels of preprocessing [22]. The layout of the models is shown in Fig. 3.

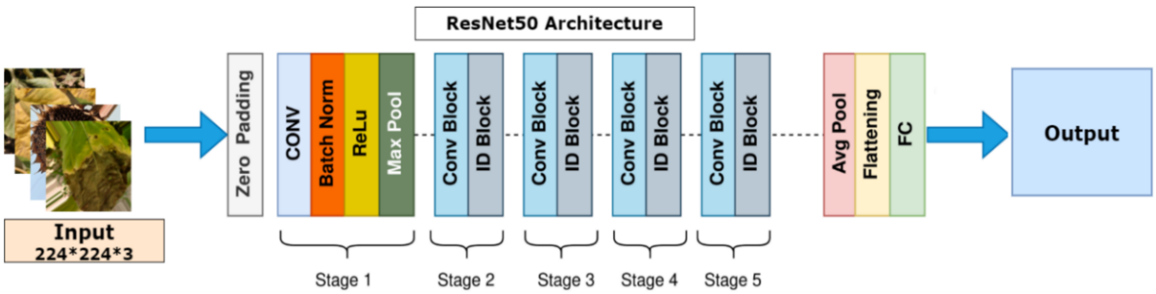

Figure-3.6.1: Architecture of ResNet50 Model.

VGG16: VGG16 is a VGG (Visual Geometry Group) variation that adds 16 layers and 16 billion FLOPs [23] to the original model. Each VGG16 block consists of a Max Pooling layer and a 2D Convolution layer [24]. VGG16 is presented in Fig. 4.

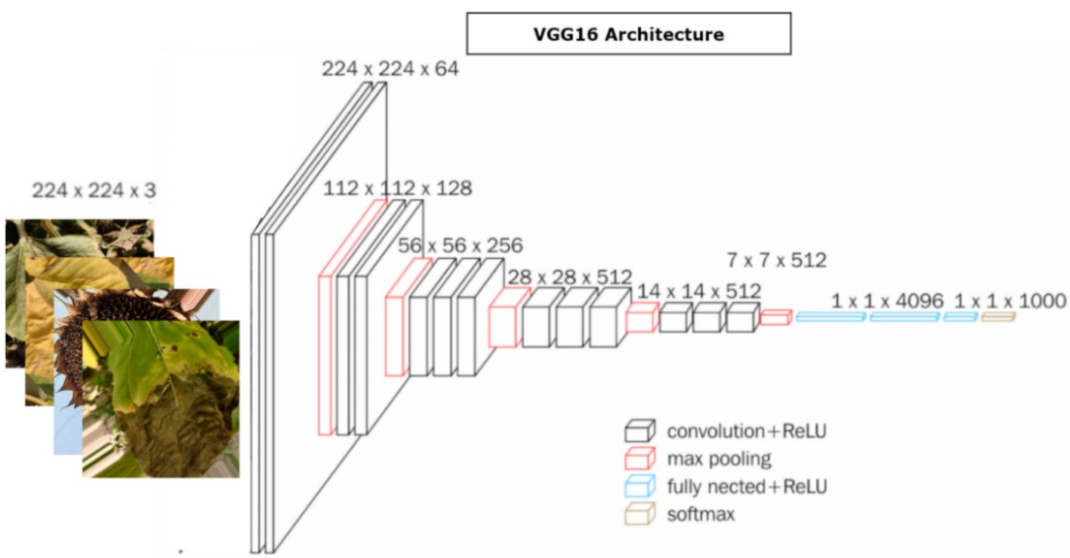

Figure-3.6.2: Visualization of VGG16 Architecture.

MobileNet V2: A CNN model has 53 layers. The details of MobileNetV2's residual inversion architecture can be seen in Fig. 5. It relies on a backward-looking residual structure that makes use of residual connections between bottleneck stages. As a source of non-linearity, lightweight depthwise convolutions are used to filter features in the intermediate expansion layer. Therefore, the model's structure consists of 32 completely convolutional layers followed by 19 bottleneck layers [25]. A network trained on over a million photos from the ImageNet database [26] might be loaded using this method.

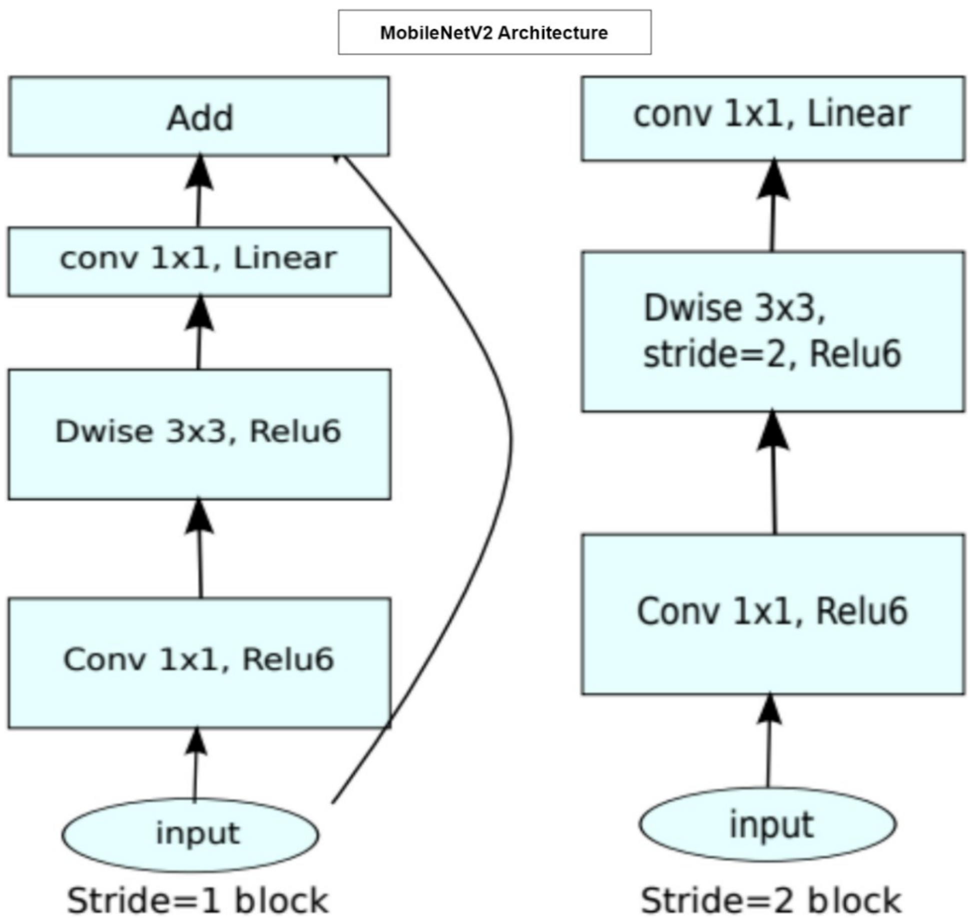

Figure-3.6.3: Architecture of MobileNetV2 Model.

Inception V3: The image size is automatically changed to 224 * 224 * 3 when using the Inception-v3 algorithm, which is part of the convolutional neural network classification model. With this network model, the number of network model parameters is drastically reduced in contrast to the AlexNet by using all connections from the AlexNet layer instead of averaging for pooling, and the use of asymmetric convolution kernels increases the diversity.

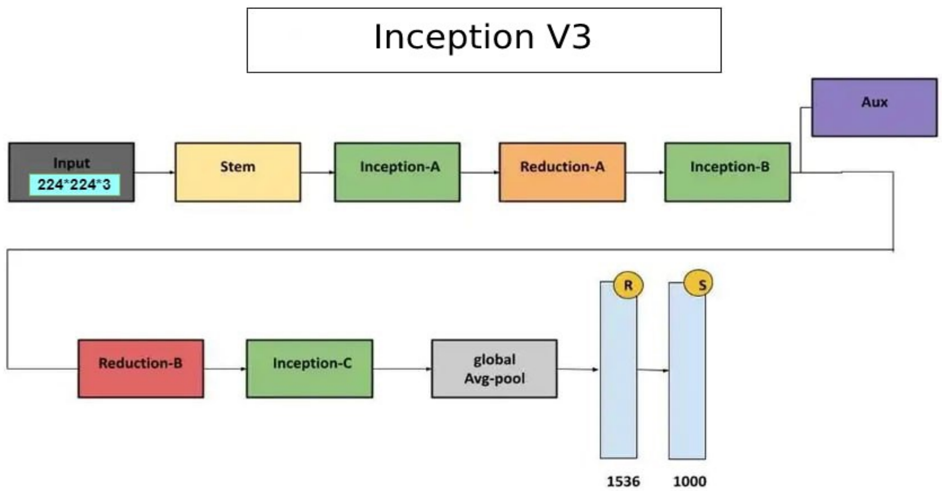

Figure-3.6.4: Architecture of InceptionV3.

Xception: The Xception network employed here can be seen as a translation of the Inception protocols. Extreme inception is another inspiration for the name Xception. To further comprehend the Xception architecture, aquick discussion of Inception is necessary.

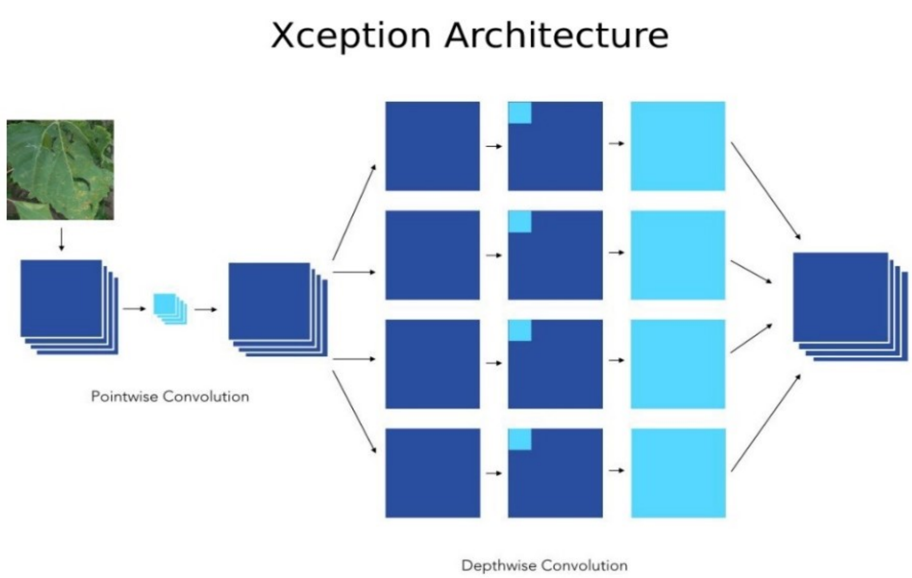

Figure-3.6.5: Architecture of Xception.

3.7 Measurement classifier performance with performance measurement matrix

Multiple metrics for the used classifier's efficacy (FNR) has been investigated. Accuracy, error rate, specificity, sensitivity, precision, false-positive rate, and false-negative speed are some of these [27–29]. The true positive (TP), the true negative (TN), the false positive (FP), and the false negative (FN) apply when P is the positive category and N is the negative category, respectively. The classifier's effectiveness is evaluated with the help of the following formula:

$$Accuracy = \left(\frac{TP+TN}{TP+FP+FN+TN} \times 100\right)\% \qquad (2)$$

$$recall = \left(\frac{TP}{TP+FN} \times 100\right)\% \qquad (3)$$

$$Precision = \left(\frac{TP}{TP+FP} \times 100\right)\% \qquad (4)$$

$$F1 = \left(\frac{2*precision*recall}{Precision+recall} \times 100\right)\% \qquad (5)$$

3.8 Implementation Requirements

- PC / Desktop.
- Internet Connection.
- Google Collaboratory.
- Python Environment / Deep Learning.

CHAPTER 4
EXPERIMENTAL RESULTS & DISCUSSION

4.1 Experimental Setup

The experiment was conducted using a GPU card Tesla T4 and a 64-bit operating system on a Google Collaboratory. Using the TensorFlow 2.8.2 backend and Python 3.7.13, Keras 2.8.0's deep learning framework created the CNN-based model.

4.2 Introduction

A "Sunflower dataset" was utilized in the construction of our model. This dataset was obtained from Data in Short [30], and it consisted of a total of 1668 pictures of data organized into four categories. The classroom was slowly taken over by gray mold, downy mildew, leaf scars, and flies. Each category included a unique set of symptoms, which aided in the process of decision-making.

4.3 Training Accuracy

In this section, we reviewed the experimental analysis, model analysis, and comparative analysis of the proposed system with prior related research.

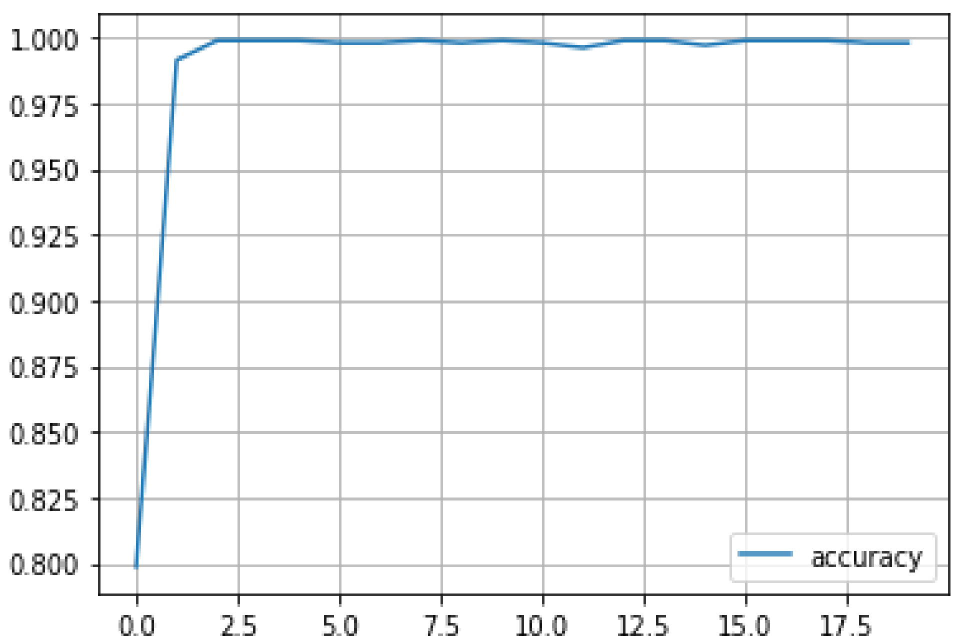

Figure-4.3.1: Training accuracy of ResNet50.

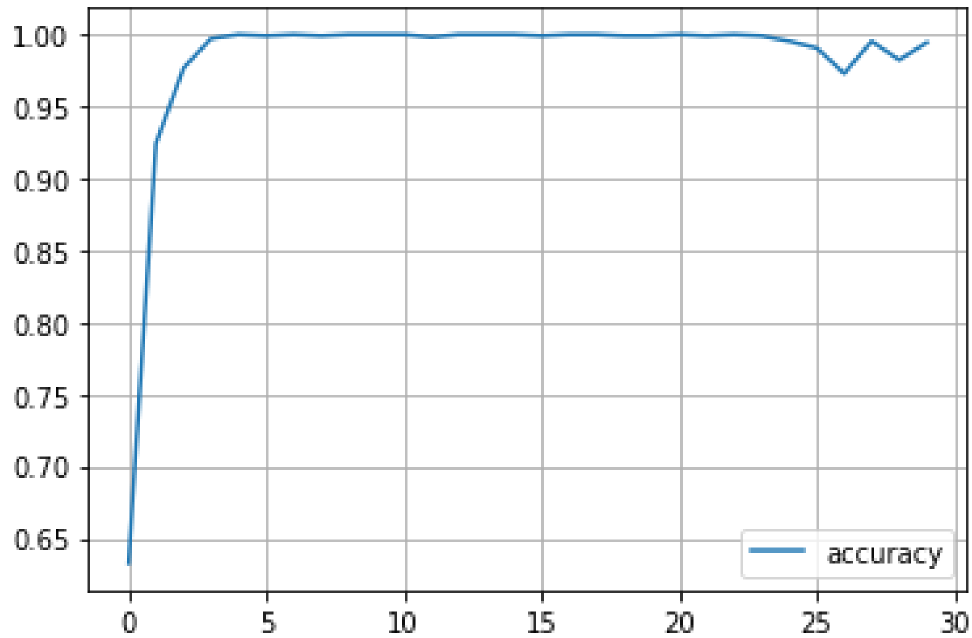

Figure-4.3.2: Training accuracy of VGG16.

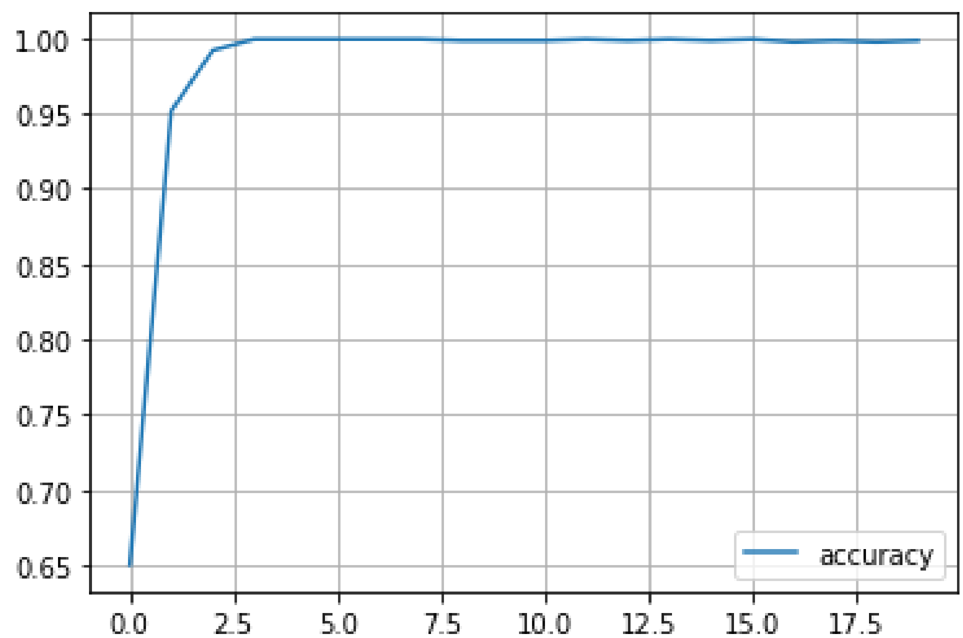

Figure-4.3.3: Training accuracy of MobileNet V2.

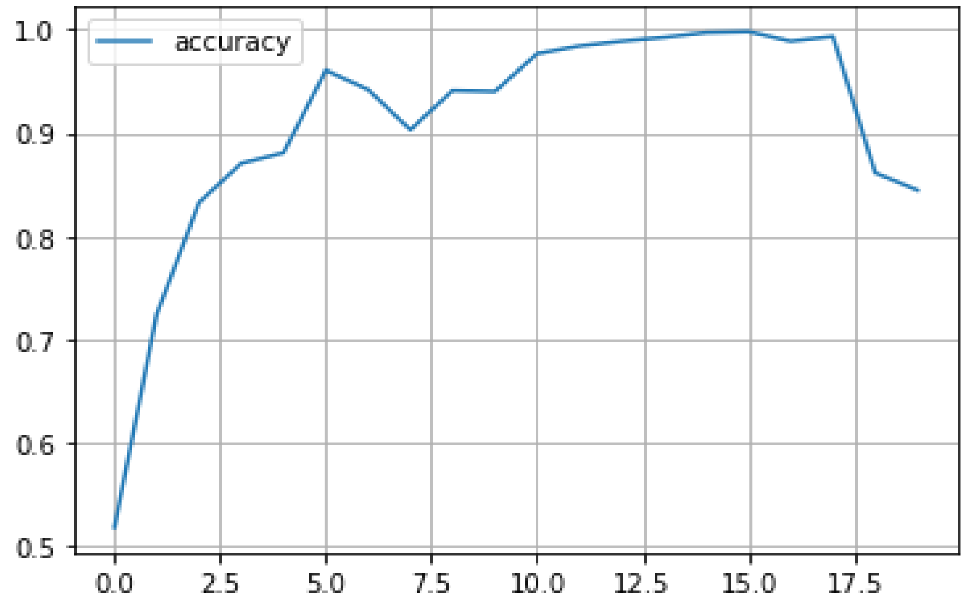

Figure-4.3.4: Training accuracy of InceptionV3.

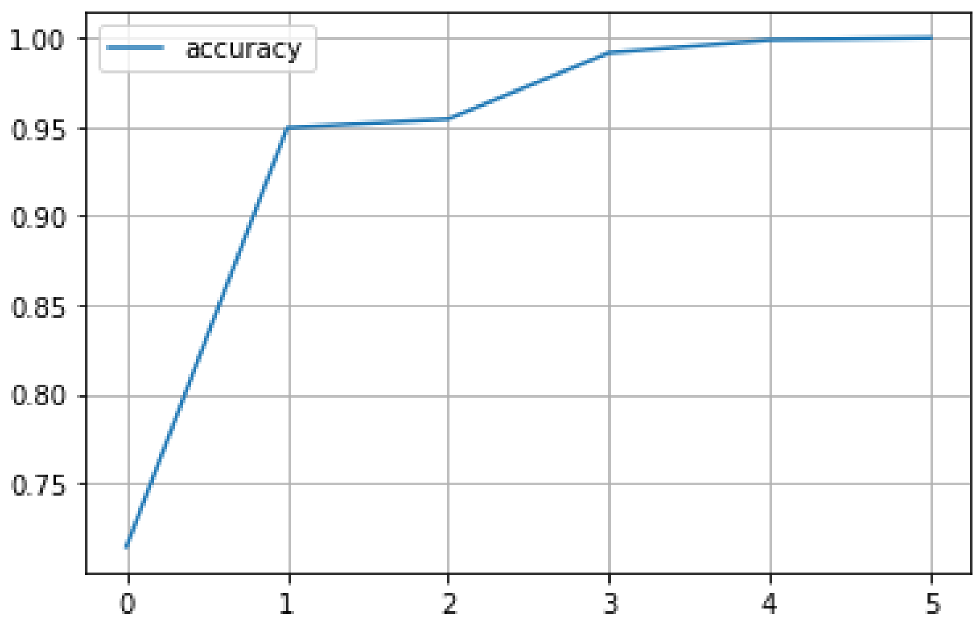

Figure-4.3.5: Training accuracy of Xception.

This study is about a disease that affects sunflowers. So, the suggested system will use four distinct varieties of data. Sunflowers are healthy four times and diseased three times. as a result of compiling data from a variety of classes. The final model is fed both the training data used to train the system and the validation data used to test its accuracy.

We use a callback function to store the most precise model. We have used various models in our tests, including ResNet50, VGG16, MobileNetV2, Inception v3, and Xception. Python, TensorFlow, and Keras were used to create this whole system. Figures 7, 8, 9, 10, and 11 show each model's epoch-by-epoch accuracy and loss. Figures 7, 8, 9, 10, and 11 show that the ResNet50, VGG16, MobileNetV2 Inception v3, and Xception models achieve an accuracy of 97.88%, 96.47%, 94.36%, 78.16%, and 76.76%, respectively. The most accurate result was obtained by ResNet50, while the least accurate result was obtained from Xception.

4.4 Confusion Matrix

Performance of each classifier are measured using different performance evaluation matrix.

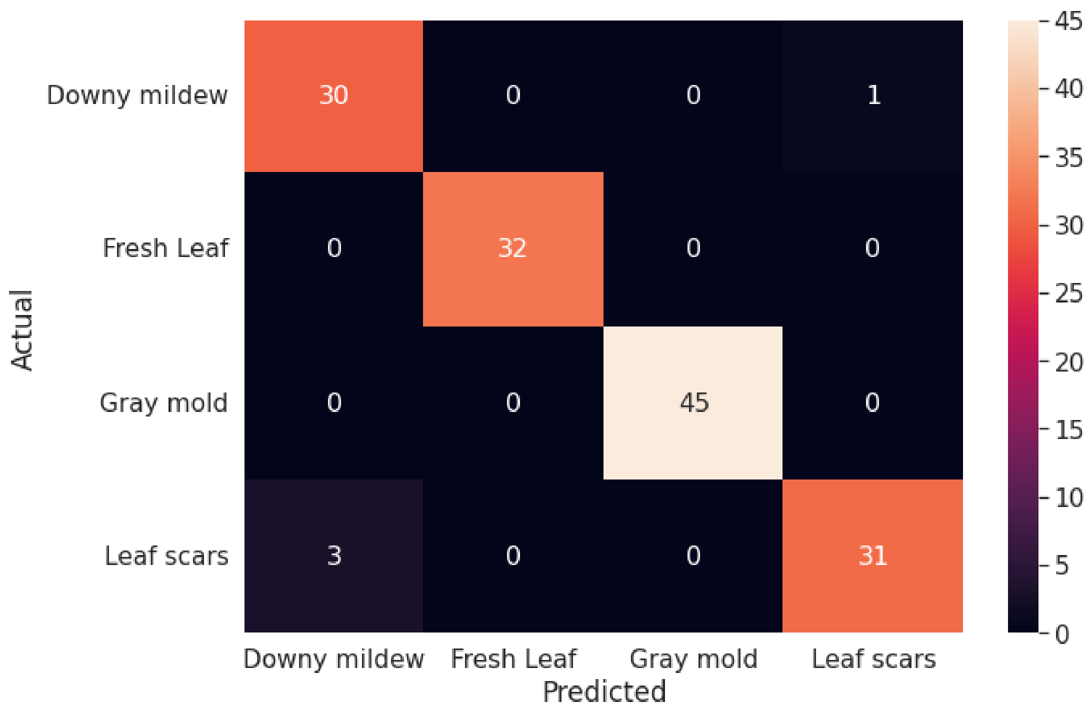

Figure-4.4.1: ResNet50 Classifier performance visualization.

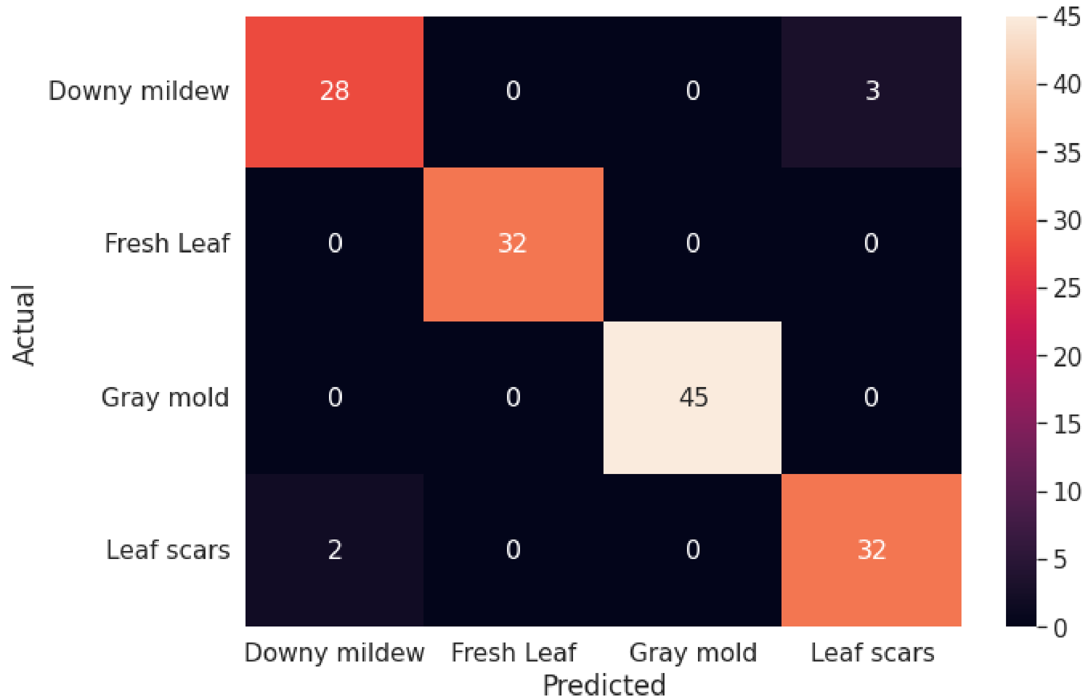

Figure-4.4.2: VGG16 Classifier performance visualization.

Figure-4.4.3: MobileNetV2 Classifier performance visualization.

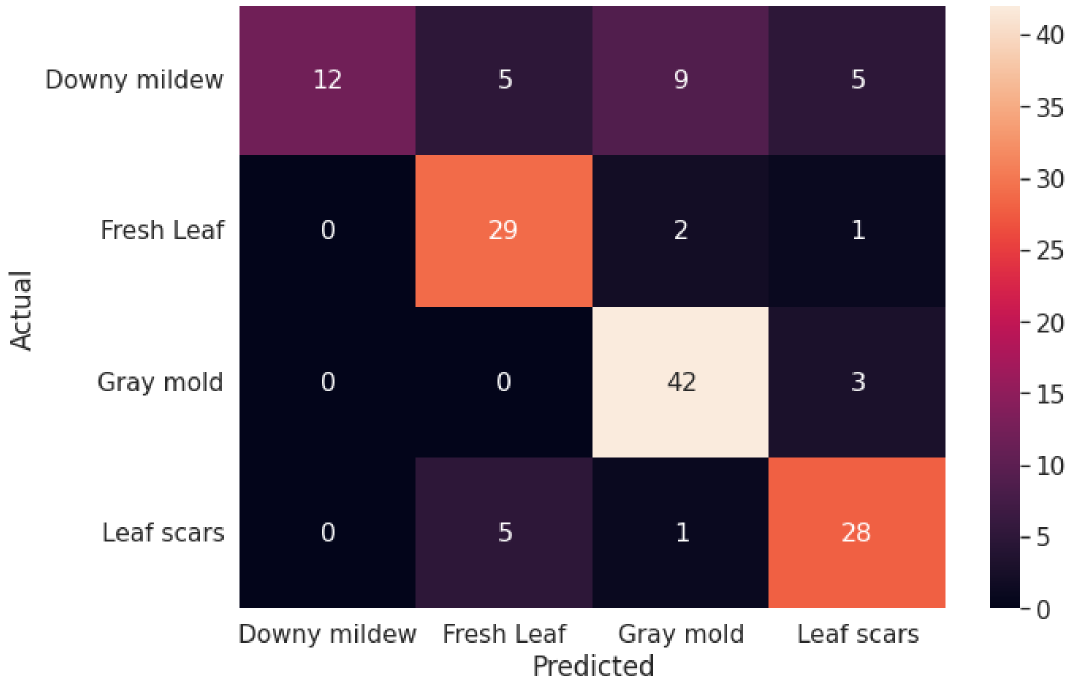

Figure-4.4.4: InceptionV3 Classifier performance visualization.

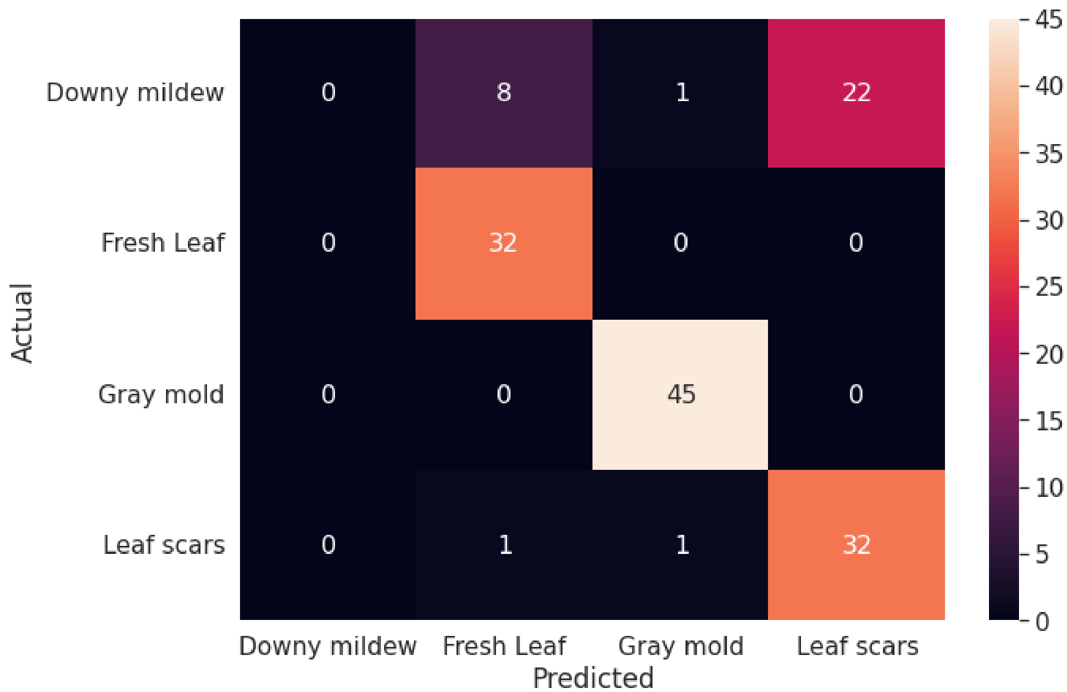

Figure-4.4.5: Xception Classifier performance visualization.

The confusion matrix produced by the models may be seen in each Figure 12, Figure 13, Figure 14, Figure 15, and Figure 16, in that order. The accuracy rates for the models ResNet50, VGG16, MobileNetV2, Inception v3, and Xception, were calculated to be 97.88%, 96.47%, 94.36%, 78.16%, and 76.76%, respectively, when using the confusion matrices shown in Figures 12, 13, 14, 15 and 17, respectively. Based on the confusion matrix presented in Figure 12, it has been determined that the accuracy of the suggested system based on the ResNet50 model is as high as 97.88%.

4.5 Performance Evaluation

Experiments were conducted with a few more state-of-the-art pre-trained models before ResNet50 was chosen as the model for transfer learning. Accuracy levels of 96.47% and 94.36% are provided to us by the VGG16 and MobileNetV2 trained models, respectively. Nevertheless, we were anticipating a system with greater precision. As a result, the MobileNet CNN model has been put through its paces, resulting in an accuracy of 97.88%. In the end, the ResNet50 model that had already been trained was used. As was said earlier, this model has the highest accuracy (97.88% of the time).

TABLE-4.5.1: CLASS BASED PERFORMANCE EVALUATION METRICS FOR RESNET50

Classifier	Disease Name	Accuracy	Precision	Recall	F1 Score	Model Accuracy
ResNet50	Downy mildew	97.88	0.91	1	0.95	97.88
	Fresh Leaf	100	1	1	1	
	Gray mold	100	1	1	1	
	Leaf scars	97.88	1	0.91	0.95	

TABLE-4.5.2: CLASS BASED PERFORMANCE EVALUATION METRICS FOR VGG16

Classifier	Disease Name	Accuracy	Precision	Recall	F1 Score	Model Accuracy
VGG16	Downy mildew	96.47	0.93	0.9	0.92	96.47
	Fresh Leaf	100	1	1	1	
	Gray mold	100	1	1	1	
	Leaf scars	96.47	0.91	0.94	0.93	

TABLE-4.5.3: CLASS BASED PERFORMANCE EVALUATION METRICS FOR MOBILENETV2

Classifier	Disease Name	Accuracy	Precision	Recall	F1 Score	Model Accuracy
MobileNet V2	Downy mildew	97.18	0.94	0.94	0.94	94.36
	Fresh Leaf	97.88	0.94	0.97	0.95	
	Gray mold	97.88	1	0.93	0.97	
	Leaf scars	95.77	0.89	0.94	0.91	

TABLE-4.5.4: CLASS BASED PERFORMANCE EVALUATION METRICS FOR INCEPTION V3

Classifier	Disease Name	Accuracy	Precision	Recall	F1 Score	Model Accuracy
Inception V3	Downy mildew	86.61	1	0.39	0.56	78.16
	Fresh Leaf	90.84	0.74	0.91	0.82	
	Gray mold	89.43	0.78	0.93	0.85	
	Leaf scars	89.43	0.76	0.82	0.79	

TABLE-4.5.5: CLASS BASED PERFORMANCE EVALUATION METRICS FOR XCEPTION

Classifier	Disease Name	Accuracy	Precision	Recall	F1 Score	Model Accuracy
Xception	Downy mildew	78.16	0	0	0	76.76
	Fresh Leaf	93.66	0.78	1	0.88	
	Gray mold	98.59	0.96	1	0.98	
	Leaf scars	83.09	0.59	0.94	0.73	

The results of ResNet50, VGG16, MobileNetV2, Inception v3, and Xception are displayed in Tables II, III, IV, V, and VI, respectively. These tables are organized according to the disease being studied.

In Table II, we can see that the ResNet50 model achieved a similar level of performance across diseases, with the healthy class yielding 100% accuracy and the Downy mildew and Leaf scars classes both yielding 97.88% accuracy. Downy mildew and leaf scars have the greatest FPR at 95%.

Table III, which shows how well the VGG16 model works for each illness, shows that the VGG16 model is best at identifying healthy gray mold and worst at identifying downy mildew.

According to Table III, which highlights the VGG16 model's performance by illness, the VGG16 model is most accurate at identifying healthy gray mold and least accurate at identifying downy mildew.

Table V displays the results of the Inception v3 model in terms of accuracy concerning different diseases. The Fresh class achieves the best accuracy (90.84%), while the Leaf Scars knot gives the lowest (89.43%). Gray and gold classes have the most precision, at 0.75, while fresh classes have the lowest, at 0.74.

Table VI provides an overview of the disease-specific performance of the Xception model. The accuracy rate of 98.59% is the best for the class "gray mold," while the accuracy rate of 78.16% is the lowest for the type of "downy mildew."

Comparing these tables reveals that the accuracy rate provided by the model ResNet50 is the highest overall, coming in at 97.88%. In contrast, the accuracy rate supplied by the model Xception is the lowest, coming in at 76.76%.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The potential social impact of our project is significant. Flowers are widely used in many environments such as oil and bee politicization and many more. Our research has shown that by diagnosing and addressing the leaf disease of sunflowers in a timely manner, it is possible to improve the survival and health of many plants. This will ultimately have a positive effect on society.

5.2 Impact on Environment

As previously mentioned, "SunNet: A Deep Learning Approach to Detect Sunflower Disease" is built upon Deep learning principles. The environmental impact of our project is minimal, as our direct carbon footprint is much smaller compared to activities such as using planes or burning coal. Therefore, we can confidently assert that our project will not have a detrimental impact on the environment.

5.3 Ethical Aspects

Since our research primarily involves plants, which are typically considered inanimate objects without feelings of humiliation or privacy, we can be assured that no ethical dilemmas have arisen during our work. As such, we can proceed with our research without any concerns about causing harm or disrespect to living beings.

5.4 Sustainability Plan

In terms of suitability there are many criteria in this arena. The main goal of our research is to reduce the economic and aesthetic damage caused by sunflower diseases. So, throughout our research, we have developed a way so that people can detect diseases and start working to cure them. Through it, the risk of disease spreading possibility can decrease, and also people can follow the right approaches at the right time. And it is very important for good and healthy monitoring in order to manage disease infections.

CHAPTER 6
CONCLUSION AND FUTURE STUDY

6.1 Summary of the Study

Sunflowers are a significant and vital element of daily life since they provide a source of nutritious oil as well as a variety of products that humans require. It is unfortunate that many varieties of sunflowers are prone to diseases that might reduce the value of the oil they produce. Sunflowers are the most common source of an alternative to other types of oil. In order to address this issue, we have implemented algorithms for the purpose of identifying these diseases, and we are working to discover strategies to recover from them. We have high hopes that if we continue our research, we may one day be able to eradicate these illnesses and protect the wellbeing of sunflower plants.

6.2 Future Study

Based on this research, an automated mobile application system has been established. This system has been given the capability to detect plant leaf diseases and deliver information regarding them in the near future.

- A dataset that has been filtered and is now available for use in the future by anyone has been developed. This will make it possible for anyone else to make further progress on this report. Because this information has been arranged, each researcher can benefit from the photos for the progression of his study.
- It's possible that in further research, more plant categories will be grouped together, and the emphasis will shift to subclassification.

The functionality of the algorithm can be enhanced by the incorporation of additional novel components and the subsequent evaluation of those components. Experiments should be utilized to train different sorts of leaves with different diseases so that the application can be used on as many different kinds of plants as is practically possible.

6.3 Conclusion

This article gives a way to look for diseases in sunflowers. The ResNet50 model is accurate 97.88% of the time, which lets us find the disease that sunflowers have at a very high rate of productivity. In spite of the fact that, we have done a lot of work to make sure that our results are as accurate as possible, even though we only have a small amount of source data. Bangladesh's agriculture sector might benefit from this kind of project. Most farmers around the world can't tell what diseases are on leaves without the help of the internet or experts. Because of this, they can't grow crops that are good enough to eat. After it has been put into place, we have high hopes that our method will help farmers figure out exactly how healthy they are. Engineers can also create things for the Internet of Things (IoT). Devices that can remove damaged leaves and fruits from sunflower fields on their own. IoT devices will be built on top of the paradigm we've set up.

REFERENCE

[1] Advantage of Sunflower, Available Online: https://www.britannica.com/plant/sunflower-plant, (Last Access: 23-04-2021).

[2] Sunflower producing countries, Available Online: https://www.worldatlas.com/articles/the-top-sunflower-seed-producingcountries-in-the-world, (Last Access: 23-04-2021).

[3] Rajbangsi, A., Biswas, A.A., Biswas, J., Shakil, R., Akhter, B. and Barman, M.R., 2021. Sunflower diseases recognition using computer vision-based approach. In *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 1-5). IEEE.

[4] Zhang, S., Huang, W. and Zhang, C., 2019. Three-channel convolutional neural networks for vegetable leaf disease recognition. *Cognitive Systems Research*, *53*, pp.31-41.

[5] Beedi, P. and Goel, P., 2021. Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. *Artificial Intelligence in Agriculture*, *5*, pp.90-101.

[6] Feeling, F., Olsson, A. and Alonso-Fernandez, F., 2018, November. Fruit and vegetable identification using machine learning for retail applications. In *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)* (pp. 9-15). IEEE.

[7] Bantam, R.A., Ali, A., Naeem, S., Jamal, F., Elgar, M. and Chesnee, C., 2020. Discrimination of sunflower seeds using multispectral and texture dataset in combination with region selection and supervised classification methods. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(11), p.113142.

[8] Supriya, A.S., Machine Learning Approach for Classifying Paddy Crop Diseases.

[9] Ramesh, S. and Vydeki, D., 2019. Application of machine learning in detection of blast disease in South Indian rice crops. *J. Phytol*, *11*(1), pp.31-37.

[10] Malik, A., Vaidya, G., Jagota, V., Eswaran, S., Sirohi, A., Batra, I., Rakhra, M. and Asenso, E., 2022. Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach. *Journal of Food Quality*, *2022*.

[11] Mohammed Zabeeulla, A.N. and Shastry, C., 2022. Automation of Leaf Disease Prediction Framework based on Machine Learning and Deep Learning in different Crop Species. *JOURNAL OF ALGEBRAIC STATISTICS*, *13*(3), pp.3098-3113.

[12] Dawod, R.G. and Dobre, C., 2022. Automatic Segmentation and Classification System for Foliar Diseases in Sunflower. *Sustainability*, *14*(18), p.11312.

[13] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D. and Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, *2016*.

[14] Brahimi, M., Boukhalfa, K. and Moussaoui, A., 2017. Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence*, *31*(4), pp.299-315.

[15] Ghosh, G. and Chakravarty, S., 2020. Grapes Leaf Disease Detection Using Convolutional Neural Network. *International Journal of Modern Agriculture*, *9*(3), pp.1058-1068.

[16] Tm, P., Pranathi, A., SaiAshritha, K., Chittaragi, N.B. and Koolagudi, S.G., 2018, August. Tomato leaf disease detection using convolutional neural networks. In *2018 eleventh international conference on contemporary computing (IC3)* (pp. 1-5). IEEE.

[17] Zhang, S., Huang, W. and Zhang, C., 2019. Three-channel convolutional neural networks for vegetable leaf disease recognition. *Cognitive Systems Research*, *53*, pp.31-41.

[18] Md. A. Islam, Md. S. Islam, Md. S. Hossen, M. U. Emon, M. S. Keya, and A. Habib, “Machine Learning based Image Classification of Papaya Disease Recognition,” in 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, pp. 1353–1360, 2020.

[19] Sunflower Diseases Dataset, Available Online: https://www.sciencedirect.com/science/article/pii/S2352340922002542, (Last Access: 22-06-2022).

[20] Ghosh, P., Azam, S., Hasib, K.M., Karim, A., Jonkman, M. and Anwar, A., 2021, July. A performance based study on deep learning algorithms in the effective prediction of breast cancer. In *2021 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

[21] Nikhitha, M., Sri, S.R. and Maheswari, B.U., 2019, June. Fruit recognition and grade of disease detection using inception v3 model. In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1040-1043). IEEE.

[22] Ramkumar, M.O., Catharin, S.S., Ramachandran, V. and Sakthikumar, A., 2021. Cercospora identification in spinach leaves through resnet-50 based image processing. In *Journal of Physics: Conference Series* (Vol. 1717, No. 1, p. 012046). IOP Publishing.

[23] Akter, S., Shamrat, F.J.M., Chakraborty, S., Karim, A. and Azam, S., 2021. COVID-19 detection using deep learning algorithm on chest X-ray images. *Biology*, *10*(11), p.1174.

[24] Sevi, M. and Aydin, İ., 2020, October. COVID-19 detection using deep learning methods. In *2020 International conference on data analytics for business and industry: way towards a sustainable economy (ICDABI)* (pp. 1-6). IEEE.

[25] Howard, A., Zhmoginov, A., Chen, L.C., Sandler, M. and Zhu, M., 2018. Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation.

[26] ImageNet Description, Available Online: http://www.image-net.org, (Last Access: 16-02-2022).

[27] Rajbongshi, A., Biswas, A.A., Biswas, J., Shakil, R., Akter, B. and Barman, M.R., 2021. Sunflower diseases recognition using computer vision-based approach. In *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 1-5). IEEE.

[28] Shakil, R., Akter, B., Shamrat, F.J.M., Jahan, N., Hasan, S. and Khater, A., 2022, October. Systematic Analysis of Several Deep Learning Approaches for COVID-19 Detection Using X-ray Images. In *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1301-1307). IEEE.

[29] Shakil, R., Akter, B., Faisal, F., Chowdhury, T.R., Roy, T. and Khater, A., 2022, March. A Promising Prediction of Diabetes Using a Deep Learning Approach. In *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 923-927). IEEE.

[30] P. Kulkarni, A. Karwande, T. Kolhe, S. Kamble, and M. Wyawahare, "Plant Disease Detection Using Image Processing and Machine Learning," p. 13.

Plagiarism Report

ORIGINALITY REPORT

17%	12%	11%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
2	Submitted to Daffodil International University Student Paper	2%
3	www.hindawi.com Internet Source	1%
4	Aditya Rajbongshi, Al Amin Biswas, Jahanur Biswas, Rashiduzzaman Shakil, Bonna Akter, Mala Rani Barman. "Sunflower Diseases Recognition Using Computer Vision-Based Approach", 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), 2021 Publication	1%
5	www.researchgate.net Internet Source	1%
6	www.mdpi.com Internet Source	1%
7	"ICDSMLA 2020", Springer Science and Business Media LLC, 2022	<1%