A TRANSFER LEARNING BASED TECHNIQUES: VEGETABLES DISEASE RECOGNITION AND CLASSIFICATION

BY

TANJILA JAHAN
ID: 191-15-2354

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

Supervised By

Musfiqur Rahman
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Dr. S. M. Aminul Haque
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Project/internship titled **"A Transfer Learning Based Techniques: Vegetables Disease Recognition and Classification"**, submitted by Tanjila Jahan, ID No: 191-15-2354 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 *30/01/2023*.

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. Mohammad Shamsul Arefin
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Ms. Sharmin Akter
Lecturer (Senior Scale)
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

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

DECLARATION

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

Supervised by:

28.01.23

Musfiqur Rahman
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Dr. S. M. Aminul Haque
Associate Professor and Associate Head
Department of CSE
Daffodil International University

Submitted by:

Tanjila Jahan

Tanjila Jahan
ID: 191-15-2354
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 **Musfiqur Rahman**, **Senior Lecturer**, Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Transfer 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 Musfiqur Rahman, Dr. S. M. Aminul Haque, and Professor Dr. Touhid Bhuiyan**,** 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

A worldwide palatable food is vegetables cultivated by the farmer and it is most important for a healthy body. There are three different vegetables such as Malabar leaves, Bean Leaves and Papaya Leaves and fruits have been used in this approach because of the big source of vitamins, calcium, iron etc. Total 887 vegetables data have been used for classifying between disease affected and Healthy vegetables using most transfer learning techniques such as MobileNet-V2, InceptionResNet-V2, Deep CNN, NasNetLarge and ResNet152-V2 pre-trained model. InceptionResNet-V2 performs better than others applied pretrained models with 93.60% accuracy with the lowest loss value and there have been multiple activation functions as elu and ReLu, and for the dense layer mainly used sigmoid and softmax function for classification of the model. The proposed approach is a more acceptable model compared to existing models based on the accuracy, recall, precision and f1 Score.

Keywords–Deep CNN, Vegetable Disease, Transfer Learning, Bean, Papaya, Malabar, MobilenetV2, ResNet152V2, NasNetLarge, InceptionResNetV2.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1 Introduction

Vegetables are known as the most popular and delicious foods worldwide with the various vitamins, nutrients, calcium and scientifically proven as the healthy foods for each and every stage in a human life.

There are mainly three different types of vegetables discussed and classified by disease and health and these vegetables are Malabar, Bean, Papaya. The dataset(Malabar, Bean, Papaya) have been collected by own capturing, from published conference's dataset and iStock websites to make it unique.

A healthy vegetable is too importants for a healthy life but now a days the farmers are suffering for various disease of vegetables which damaged their vegetables planting and continuously they need to faced the loss of vegetables planting and that's why they are demotivated to planting vegetables and the lack of vegetables are impacting on worldwide humans. This is why this proposed model would be a most convenient solution which will motivate the farmers to plant vegetables. Because this proposed model is able to classify between healthy and disease affected vegetables with the lowest classify and recognition duration.

Various research articles have been already available using the both supervised and unsupervised learning techniques based on the label and unlabeled data with the different performanced. The proposed method is designed based on supervised learning because of label data to get the acceptable performance and also compared among these applied algorithms. It is pretty clear that the designed model is much better to provide a strong model with acceptable performance and because of Supervised and label data. The purpose of label data defines the preprocess of data labeling as it is required for the proposed models.

The proposed approaches are mainly implemented by Convolutional Neural Network based keras pre-trained model as MobileNet-V2, InceptionResNet-V2, Deep CNN, NasNetLarge and ResNet152-V2 because of fast and most acceptable accuracy with lowest loss value. It compared among the mentioned model’s accuracy, precision, recall

and f1 score as a comparative model to highlight the better performance and more acceptable performanced model.

There have been various research articles to detect and recognize between vegetables disease and health that are already available online but there are different types of vegetables used and the first unique reason is the different type of dataset with the combination of Malabar, Bean and Papaya.

1.2 Motivation

Vegetable disease is one of the main reasons for economic losses and is a threat to agriculture. Image processing as a deep learning algorithm is developed for detecting plant disease by identifying the image color feature of the leaves area. Therefore, diagnosis is one of the most important aspects of a plant pathologist's training. Without proper identification of the disease and the disease-causing agent, disease control measures can be a waste of time and money and can lead to further plant losses.

1.3 Rationale of the Study

There are various vegetables disease related articles already available online with different types of machine learning algorithms but for implementing the proposed approaches is selected those vegetables which are delicious food and normally people fixed it in their daily life. The Deep CNN based deep learning models have been selected because of the imageNet dataset which is so fast and mostly provides acceptable accuracy. It is important to say that this implemented model is defined to help those farmers who cultivate the Bean, Malabar and Papaya to define the disease and will take action against the disease.

1.4 Research Questions

1. What is the main reason to propose a vegetable disease approach?
2. What is the unique entity of this approach?
3. How to implement these applied Deep CNN based models?
4. What is the success key of vegetables cultivated farmers to follow this approach?

1.5 Expected Output

- A acceptable accuracy with minimum loss value.
- Multiple Deep CNN based Keras model will be applied to find a best performed model with compare to all applied models.
- Three different vegetables(Bean, Malabar, Papaya) is collected to make a unique dataset.
- There have been used differents activation function, dense layer(GeLu, SeLu,ReLu, elu, Softmax, Sigmoid, etc) in the applied model to find the better performance.

1.6 Project Management and Finance

After selecting the proposed model systems, the desired dataset is collected from different resources like published conference datasets, iStock and by own capturing. Then I have run the Keras pre-trained model in google collab with the minimum time of GPU services and total 50 numbers of epochs have been used to train the proposed model. I didn't need to pay to collect the dataset and I have run the proposed model with respect to Keras pre-trained model without cost.

1.7 Report Layout

- Introduction
- Background
- Research Methodology
- Experimental Result and Discussion
- Impact on Society, Environment and Sustainability
- Summary, Conclusion, Recommendations and Implication for Future Research
- Reference

CHAPTER 2

Background

2.1 Preliminaries/Terminologies

A popular palatable food is vegetables affected by different types of disease at the early stage, middle stage and that's why the vegetable farmers are continuously losing their interest to cultivate the vegetables. There are lots of articles already available online but this proposed approach will be a more acceptable article based on the result of applying different Deep CNN based Transfer learning algorithms, and will find the best specific algorithms comparing among them with multiple types of vegetables such as Bean, Malabar and Papaya. The purpose to propose this approach is not only to help the vegetables cultivate farmer's society but also the scientific society to get the idea about the disease of the desired vegetable at different stages.

2.2 Related Works

There are very few amounts of national and international conference and journal papers are mentioned here,

Habib et al.[1] proposed an agromedical expert system based on machine vision.In this study, they take an in-depth exploratory method to identifying papaya disease using machine vision. The disease-attacked region of the acquired image is segmented using the K-means clustering algorithm, and then the necessary characteristics are retrieved to let the support vector machine categorize the diseases. The classification accuracy rate has increased to around 90%.

Ashiqul et al.[2] proposed a machine learning-based intelligent system that can identify papaya diseases has been described in the proposed study work. The main goal of this study is to examine several algorithms for recognizing papaya disorders, diagnose the condition by capturing an image, and categorize the patients according to their conditions using an intelligent system. The data was gathered from Mendeley. This study uses CNN obtained a good accuracy in CNN (98.4%) and other techniques including random forest, SVC, SVC and k-means clustering.

Vinod et al.[3] proposed a study demonstrated the ability of the Residual Network model (ResNet34) to recognize and categorize disease from photographs of leaves. The proposed ResNet34 model demonstrated viability by achieving a 99.40% accuracy. ResNet model are also compared with that of four other techniques- SVM, K-NN, Decision Tree and Logistic Regression. The dataset has 15200 photos with 14 different crops that are separated into 38 different classes. However,To enable the network to recognize and categorize a greater variety of diseases and plant species, the dataset can be expanded by adding more and more photos as data points.

Sagar et al.[4] proposed categorization and prediction methods for papaya disease via convolutional neural networks.According to Keras API, CNN model is utilized.Fully connected, where categorization is finished and every step is based on deep learning, makes this model dependable. It used a 200x200 RGB image with a set size as its input.The data sets are gathered from both kaggle and actual photographs. With an accuracy of 91%, their model has done somewhat better than other models that have been presented.

Adhao et al.[5] proposed the Machine Learning regression techniques for cotton leaf disease detection and controlling using IoT based with the overall classification accuracy as 83.26% using Support Vector Machine. Cotton five Leaf disease as the Bacterial Blight, Alternaria, Gray Mildew, Cereospra and Fusari-um wilt are monitoring for making the cotton leaf disease detection techniques.

Mayen et al.[6] proposed a advanced computer vision techniques to recognize the scab on Malabar nightshade including MATLAB with multiclass SVM classifier. This work presents a straightforward and computationally efficient method for proving the existence of leaf sickness and conducting reviews using image processing and computer vision.There are nightshade diseases are detected with 82.15% accuracy and also used Bayes Net, Trees.J8, Function logistic algorithm with less accuracy.

Surampalli et al.[7] designed a acceptable technique to detect Tomato Plant Leaf Disease utilizing image processing methods based on picture segmentation, clustering, and open-source algorithms, such that all contributing to the development of a leaf illness with a focus on tomato plants.The purpose of the paper is to develop a methodology for

categorizing tomato leaf illnesses and offering the best remedy to treat them. This has been successfully accomplished using open source programming language Python, cutting-edge specialized algorithms, and image processing techniques.The suggested technique maps input image pixel intensities and compares them with images from the trained dataset using the CNN algorithm for hierarchical feature extraction.The proposed approach has a 98% accuracy rate.

Sandeep et al.[8] designed a machine learning based leaf disease classification and detection is achieved different accuracy for different attributes as for Alternaria alternate, bacterial blight, Anthracnose, Cercospora leaf spot and healthy leaf with 96.7102%, 95.8042%, 96.0042%, 95.1772%, 98.3871% accuracy. There is also used k-means clustering for segmentation, GLCM (Gray Level Co-occurrence Matrix) for feature extraction and Support Vector Machine(SVM) for classification.

Raida et al.[9] have been proposed a transfer learning approaches for disease detection of plant leaves using the pre-trained model Mobilenetv2 as feature extraction method with respect to CNN(Convolutional Neural Network) achieved 90.80% validation accuracy and classification into two categories as disease affected and disease free leaves. They use real image capturing dataset and Their dataset has 7800 photos, with 4976 images for the train set, 1102 images for the test set, and 1742 images for the validation set.

ELHOUCINE et al.[10] developed an automatic model to categorize and identify the type of disease based on MoblieNet, beans leaf images, and based on an effective network architecture, in order to produce precise models that can be readily categorize the disease into their classes. The best experimental result was obtained when their model was trained using the Adam optimizer, learning rate of 0.001 and batch size of 32 with training and validation set accuracy of 100% and 98.49% respectively, and the model also achieved an accuracy of 92.97% on 128 test data. Similarly, in this study, the minimum training accuracy and validation accuracy were 98.The dataset utilized in this study is a public dataset presented by Tensorflow and was selected from GitHub. Experts from the National Crops Resources Research Institute (NaCRRI) in Uganda annotated this public dataset by identifying which disease was manifested in each image.

Soner KIZILOLUK [11] proposed automatic disease classification from images of potato, cotton, bean, and banana leaves has been achieved using standard and trained versions of well-known CNN models DarkNet-19, GoogleNet, Inception-v3, Resnet-18, and ShuffleNet.Transfer learning considerably boosts the performance of CNN models, according to experimental findings.Four distinct datasets on potatoes, cotton, beans, and bananas that were retrieved from Kaggle were used in this investigation.It can be shown that Pretrained-DarkNet-19 performed the best, with a mean accuracy of 94.54 and a maximum accuracy of 97.1 for training validation and a mean accuracy of 94.22 and a maximum accuracy of 96.9 for testing. Furthermore, it is clear that the transfer learning approach has improved the performance of conventional CNN models by 20%.

Husnul et al.[12] has proposed a comprehensive diagnostic method for tomato and potato plant leaf diseases with a graphic layout of preventive measures utilizing image processing and CNN. On Kaggle dataset of the potato and tomato leaf, image processing techniques are applied through the operations of data pre-processing, augmentation, and data extraction to explore the symptoms of diseased leaves.This study suggests utilizing image processing and two well-known convolutional neural network (CNN) models, AlexNet and ResNet-50, to detect plant leaf illness and take preventive steps.The experimental findings demonstrate the effectiveness of the suggested approach, achieving, for the categorization of the healthy-unhealthy leaves and leaf illnesses, respectively, overall ResNet50 accuracy scores of 97% and 96.1% and overall AlexNet accuracy scores of 96.5% and 95.3%.

Aliyu Muhammad Abdu, et. al.[13] have proposed an Automated vegetable disease classification and identification based on the specific lesion of features and specifically used the CCV(Color Coherence Vector) and LPB(Local Binary Pattern) methods to identify the vegetable disease. There need to write a huge code and it takes more space complexity just because of uses the methods instead of uses Keras based neural networks.

Naimur Rashid Methun, et. al.[14] proposed a CNN based model to recognize the Carrot disease and there have four different types of disease and healthy carrots are also used to make a dataset. To make the between the existing model and this proposed model is

defined accuracy, precision, recall, and f1 score but there is no AUC value and ROC curve to define the strongest of the proposed model's performance. In this proposed model, Inception V3 provides better accuracy is 97.4%.

Meenakshi Sharma et. al.[15] proposed a review of a based Machine Learning Approach to the classification and recognition of vegetable disease which helps to think about the existing model and their performance. there also discussed the labeled data and unlabeled data with respect to supervised and unsupervised learning techniques. It also defines the deep learning model, especially CNN, and ANN.

2.3 Comparative Analysis and Summary

Table-2.1: Comparison Between Existing and Proposed Model

Key of Comparison	Existing Model	Proposed Model
Purpose of Study	Used to Recognition and classification of vegetables disease using the Machine Learning and mostly used Deep Learning.	Used to Recognition and classification of vegetables disease using Deep CNN based Transfer Learning.
Dataset	Both Single vegetable and multiple vegetables leaf and fruits are used to make the dataset.	Multiple vegetables have been used to make a dataset including 887 datas as Bean, Malabar and Papaya.
Model Uniqueness	Machine Learning and Deep Learning models.	Deep CNN based Transfer Learning with the different types of layer of activation function and dense layer used to find comparatively better performance.

Process of Implementation	Collected Data,Preprocess, Feature Extraction, Train Model.	Collected Data, Preprocess, Checking multiple size of data and convert into single size, Feature Extraction, Train Model

The existing models available online are implemented based on Machine Learning, Deep Learning to recognize the vegetables disease but Deep CNN based Transfer Learning have been used to propose models. The dataset is totally collected by own capture and collecting from google searching is made a unique dataset with various size of data and for that firstly it converted into a single size of data, then feature extraction and finally trained model using differents layers to get the better performance from the proposed model is a lengthy process from the normal way to implemented Keras pretrained model. Proposed model Accuracy and Loss is 93.60% and 0.41 which is pretty good combinely comparative with existing model.

2.4 Scope of the Problem

There have some reasons to propose this vegetable disease recognition approach not only for designing an article but there have some common procedures which resolve some issues of future researchers and farmers are given below,

- As a comparative study, researcher can find the best one in future for their approaches
- Transfer Learning based article are able to recognitions same type of image dataset's disease
- There have used multiple layers of CNN are reuse able with the requirements
- In future, there have another opportunity to proposed a comparative study using the existing models.

2.5 Challenges

- To study vegetable and transfer learning based article
- To make a unique dataset with Bean, Malabar and Papaya

- To implement the Deep CNN based transfer learning model because of getting the better performance by checking different layers
- To understand the writing steps to follow the standard format
- To find the uniqueness of the proposed model comparative with the existing model

CHAPTER 3

Research Methodology

3.1 Research Subject and Instrumentation

The proposed approach is designed to recognize and classify healthy vegetables and disease vegetables mainly selected based on the public demands like calcium, iron and vitamin and also selected after measuring the deliciousness of the vegetables. So I finally selected the Bean, Malabar and Papaya from the large amount of vegetables because of its calcium, iron, vitamin and deliciousness. After selecting the dataset, although the dataset is collected by own capturing, collected from different sites like google searching and iStock, but it was quite difficult to collect the perfect dataset. Then I needed to resize all of the image into a single size because all of the images were different sizes and finally it was totally prepared for training the desired model.

3.2 Data Collection Procedure/Dataset Utilized

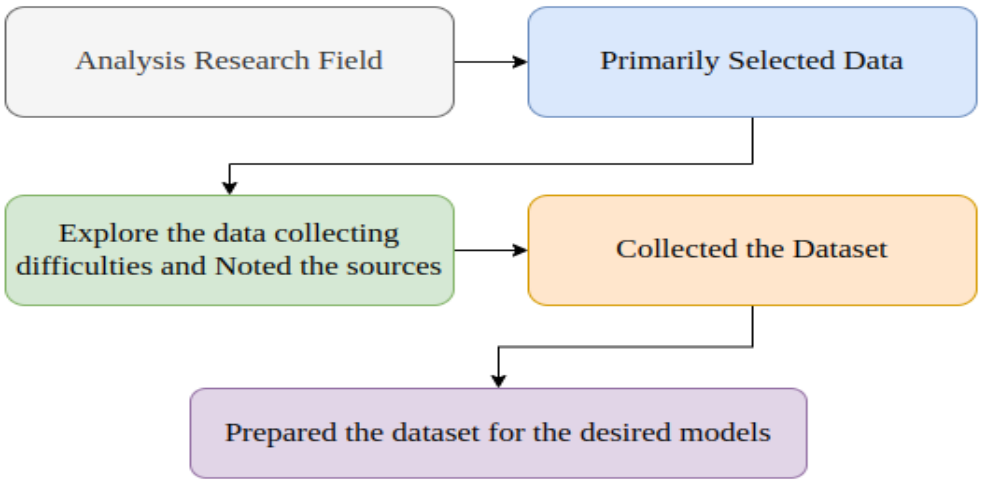


Figure 3.1: Data Collection Procedure

From the figure of Data Collection Procedure is defined the analysis of the research field based on demand of the data worldwide. After primarily selecting the data for the proposed method as vegetables and kind of vegetables based on calcium, iron and vitamin because it helps a human to be healthy and also thought about the deliciousness of these vegetables. After exploring the way of collecting the dataset, it was quite

difficult to make a unique dataset and then the resources were collected like published conference's dataset, iStock and by own capturing are the main steps to make it a totally unique dataset. A total of 887 data was collected and made a vegetables dataset with Bean Disease Leaves, Fresh Leaves, Malabar Disease Leaves, Malabar Fresh Leaves, Papaya Disease Leaves and Fruits and Papaya Fresh Leaves and Fruits. After collecting the dataset, it was split into three parts like training, validation and testing is mentioned in table-Dataset Distribution.

Table 3.1: Dataset Distribution

Attributes Name	**Training**	**Validation**	**Testing**
Bean Healthy Leaves	195	37	74
Bean Unhealthy Leaves	83	14	29
Malabar Healthy Leaves	76	20	19
Malabar Unhealthy Leaves	73	20	12
Papaya Healthy Fruits and Leaves	34	9	6
Papaya Unhealthy Fruits and Leaves	150	19	16

The dataset distribution is completed based on individually total number of dataset of the desired objects like Bean Healthy Leaves, Bean Unhealthy Leaves, Malabar Healthy Leaves, Malabar Unhealthy Leaves, Papaya Healthy Fruits and Leaves and Papaya Unhealthy Fruits and Leaves.

3.3 Statistical Analysis

Due to existing research article, it's clear that there are a huge amount of research article have been published to helps the farmer and they also classify the vegetables disease like Cucumber, Pumpkin, Peppers, malabar, Tomato, Potato,Bean, Eggplant, Papaya etc and maximum time they takes the single type of vegetables to classify the disease and also multiple types of vegetables. But the proposed vegetables as Bean, Malabar, Papaya

were selected based on public demand in daily life and also think about the calcium, iron, vitamin and deliciousness. Then it thought about the farmer because they lost their profits cultivating these vegetables and it would be too helpful in the agriculture world to classify the disease and early disease of these vegetables so that they can take attempts to destroy the disease and make healthy vegetables for people and for nations.

3.4 Proposed Methodology/Applied Mechanism

3.4.1. MobileNet-V2

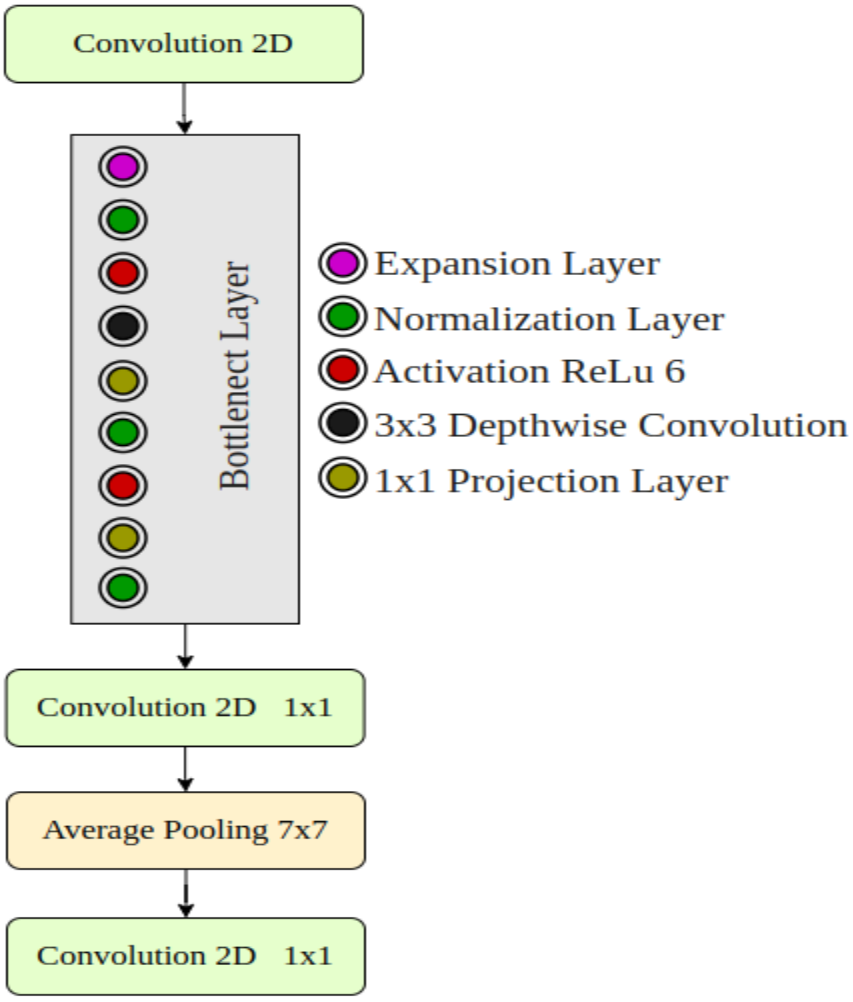


Figure 3.2: Architecture of MobileNet-V2

MobileNet-V2 (Javierto et al., 2021) mainly defines the inverted residual block with the bottlenecking features and the bottleneck layer contains a few previous layers based on compared nodes. The input size is required that must be getter than 32 x 32 and provides better performance. MobileNet-V2 is mainly initialized by the keras pre-trained model as tf.keras.applications.mobilenet_v2.preprocess_input and mobilenet_v2.preprocess_input is used to scale the input image between 1 and -1. The classifier_activation mainly defines the output layer of the model and suggests either softmax or sigmoid function.

3.4.2. InceptionResNet-V2

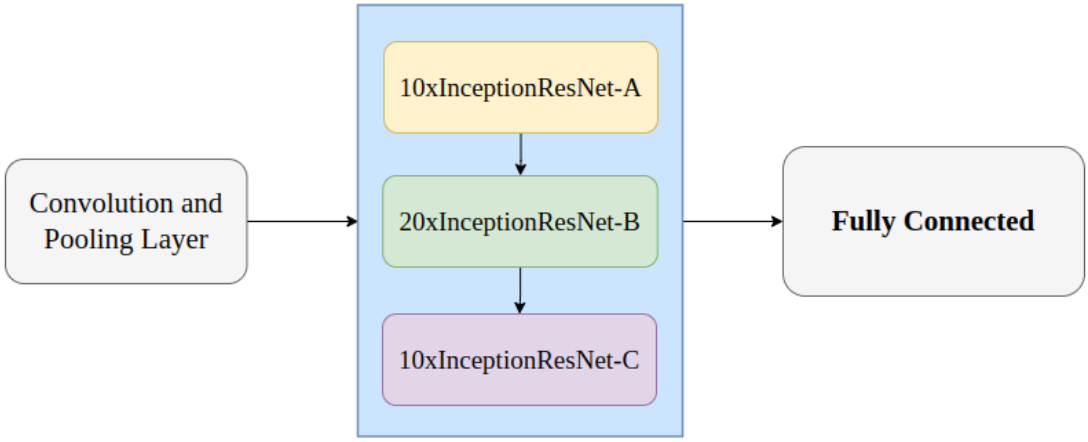


Figure 3.3: Architecture of InceptionResNet-V2

InceptionResNet-V2 (Naveenkumar et al., 2021) is a popular classification solution for the ImageNet database that contains a million images based on a CNN(Convolutional Neural Network) and there are 164 network layers and also defines 1000 objects category of image classification. tf.keras.applications.inception_resnet_v2.preprocess_input is used to initialize the model in required idle and inception_resnet_v2.preprocess_input is used for scaling the input image like -1 and 1. The define output layer needs to call the softmax or sigmoid function of classifier_activation and it can make the classification of the desired approaches.

3.4.3. ResNet152-V2

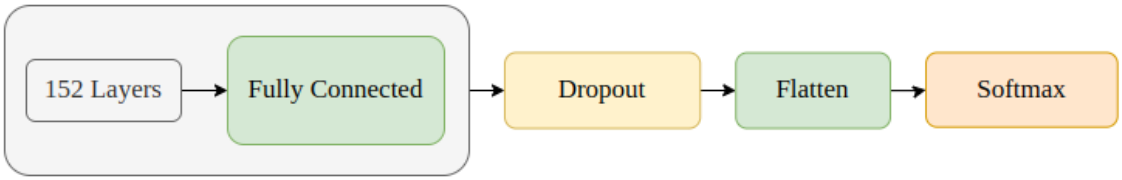


Figure 3.4: Architecture of ResNet152-V2

The ResNet152v2 (Elshennawy & Ibrahim, 2022) is able to obtain a successful performance with respect to the ImageNet dataset and this model mainly defines residual connections, the step of improving the gradient flow, allowing the deeper model and suggesting hundreds of layer of the models to train. For initializing the ResNet152v2 model, we need to call tf.keras.applications.resnet_v2.preprocess_input and resnet_v2.preprocess_input is used to scale the input image size between -1 and 1. It never suggests smaller than 32 sizes of images and although mostly suggests the softmax function as classifier_activation but it also suggests the sigmoid function to get the output of the desired model.

3.4.4. Deep Convolutional Neural Network(Deep CNN)

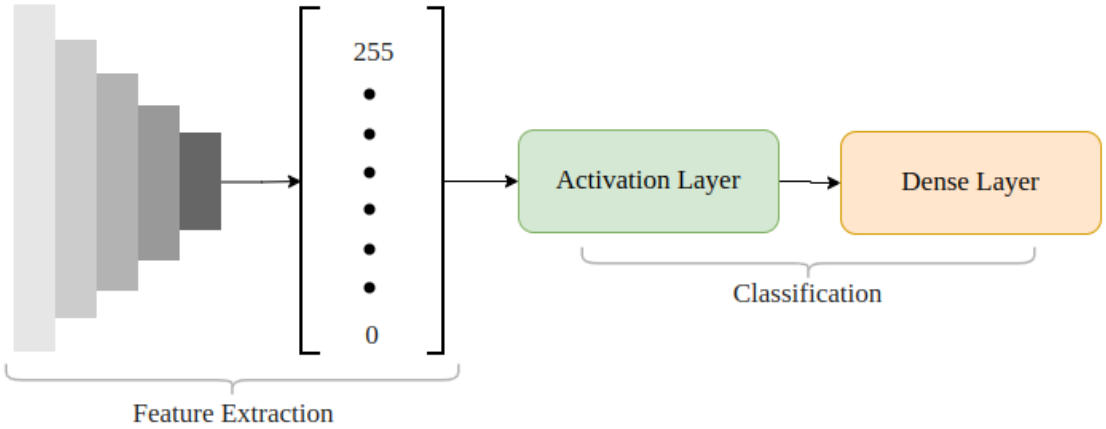


Figure 3.5: Architecture of Deep Convolutional Neural Network

Deep CNN (Sivakumar et al., 2021) defines with some set of layers like the Convolution Layer, ReLu, maxPooling and Fully Connected Layer which exactly behaves like the

neurons of the human brain. There are multiple image size have been used as target image sizes like 224x224x3 and other types of shapes.

$$Y = f(\sum_{i=1}^{n} x_i * w_i + b) \quad (1)$$

From the equations of (1) f is activation function, x defines the set of input, w defines the weight, b means bias and Y defines the output of the proposed architecture.

3.4.5. NASNetLarge

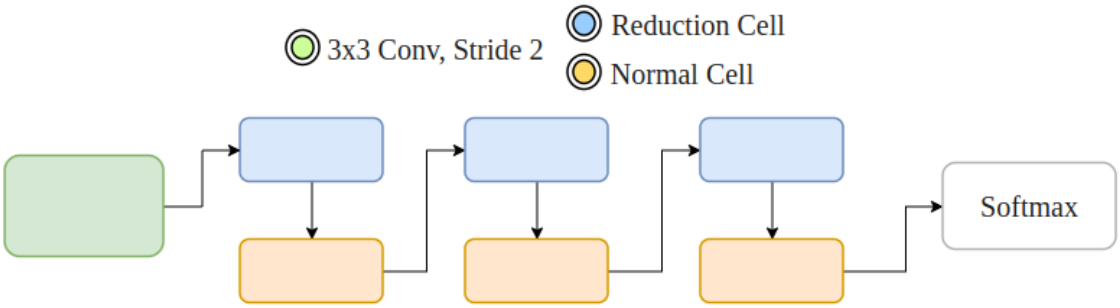


Figure 3.6: Architecture of NASNetLarge

NASNetLarge (Mehmood et al., 2022, 4368) is trained with respect to the ImageNet database for millions of image based on a (CNN) Convolutional Neural Network and it is able to classify 1000 object categories of the desired images dataset and the proposed model prefer to select a imageNet dataset. There have been used tf.keras.applications.nasnet.preprocess_input to initialize the Keras pre-trained model and mostly suggest using the input size 224x224x3. With the similarities of other Keras pre-trained models, it also suggests using the softmax function as classifier_activation to get the output in the output layer of the pre-trained model.

3.5 Implementation Requirements

There are some requirements to implement the desired proposed approach based on dataset like image, video, binary etc type of datas as given below:

1. After defining the dataset, note the proposed dataset collection resources if.
2. Collected the dataset and splitted into Training(69%) , Validation(13%) and Testing(18%)
3. In the google colab, firstly imported the dataset with linking google drive
4. Installed the needed python packages like opencv, tensorflow, keras matplotlib, numpy
5. Preprocess the proposed model
6. Trained the model using Training data with respect to validation
7. Evaluated the training model based on testing data and found the confusion matrix

CHAPTER 4

Experimental Result and Discussion

4.1 Experimental Setup

In the Experimental Setup is a way to explore the whole process of pre-trained models to classify between disease and healthy vegetables. After collecting the vegetables data it defines the preprocess of the models like it installs required packages, resizes all of the images data based on models, and defines the keras pretrained models.

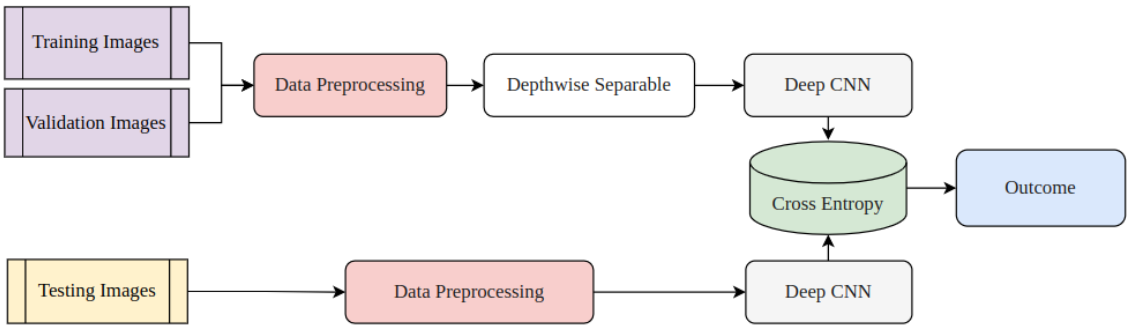


Figure 4.1: Experimental Setup

After the preprocessing, it is the step of training the model based on Deep CNN and in the training steps, there have been used training and validation data to train the model in an efficient way to get a better performance. Right now, the model is ready to cross entropy and that's why in the model, the testing data is also preprocessed with respect to the trained model and finally it is used for cross entropy and in the section of outcome it successfully classifies and recognizes the vegetables disease.

4.2 Experimental Results & Analysis

The proposed model defines the preprocess of collected data and resized based on applied keras pre-trained model like MobileNet-V2, Inception ResNet-V2, Deep CNN, NasNetLarge and ResNet152-V2.

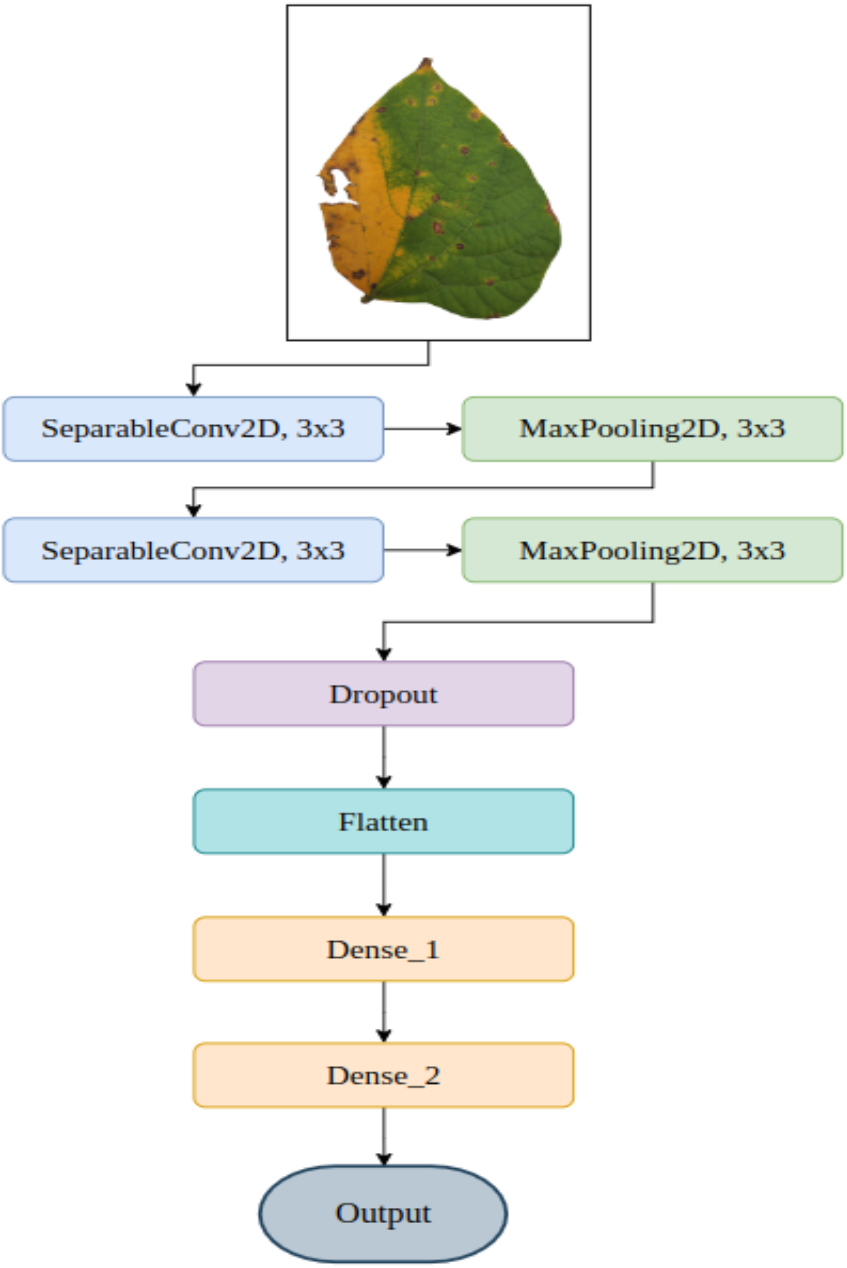


Figure 4.2: Proposed Model

In a transfer learning model, the Convolutional layer is known as the learnable layer which mainly increases the classification size of the input images, SeparableConv2d is used in the proposed approach for fast computation of the learnable layer. It mainly performs based on pointwise convolution on depthwise spatial with the mixing result.

$$N_H = N_W = \frac{n+2p-f}{S} + 1 \quad (2)$$

where S=1, p=0, f=3

MaxPooling2D is mainly used for reducing the desired spatial resolution of the input images.

$$ReduceSpatialResolution = \frac{\frac{n+2p-f}{S}}{2} \quad (3)$$

where S=2, f=2.

Then there have been the Dropout to overfit the model, Flatten is used to convert into 1-dimensional, dense is used to fit the images and the last dense is mainly used to classify the model.

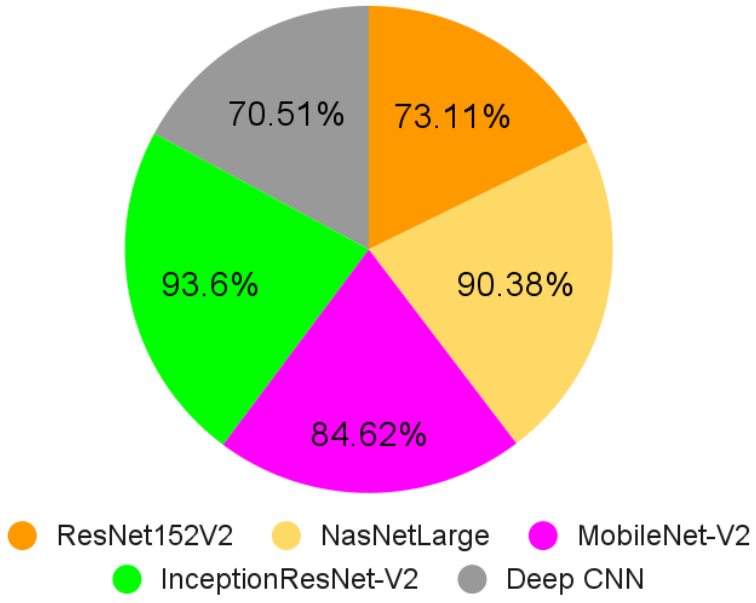
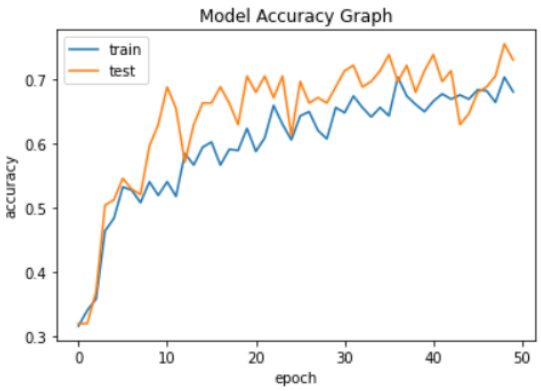
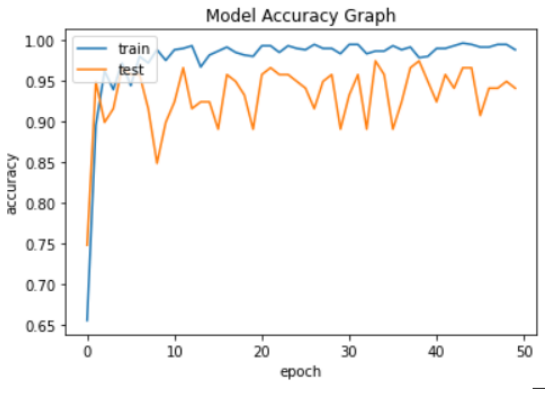


Figure 4.3: Model Accuracy Chart

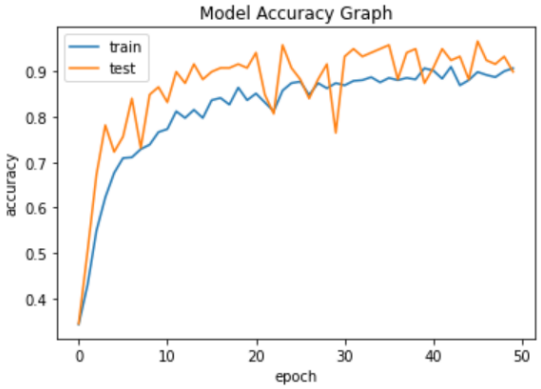
There have been differences in accuracy for different models given in Fig-Model Accuracy and InceptionResNet-V2 provides better accuracy compared to other applied algorithms. It also provides the model accuracy and loss graph with respect to total epochs as given below,



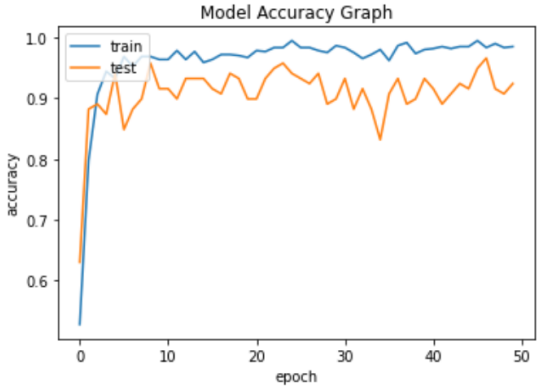
(a) ResNet152V2



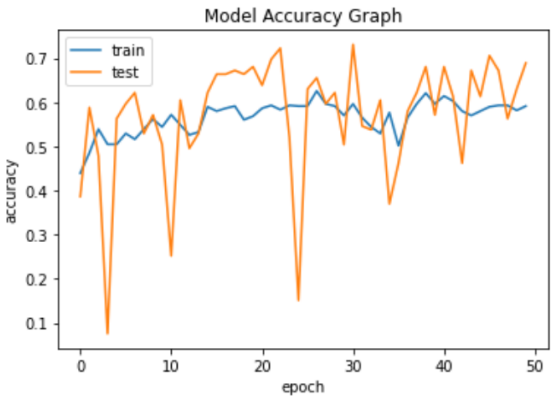
(b) NasNetLarge



(c) MobileNet-V2



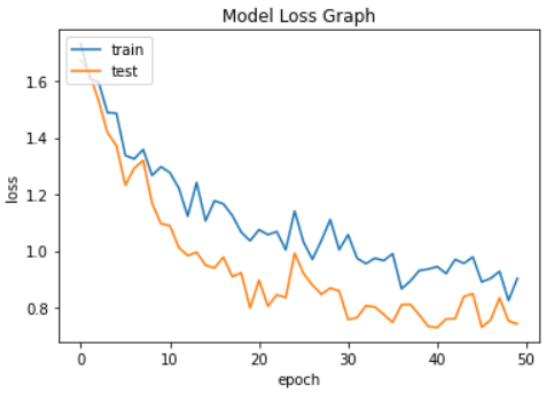
(d) InceptionResNet-V2



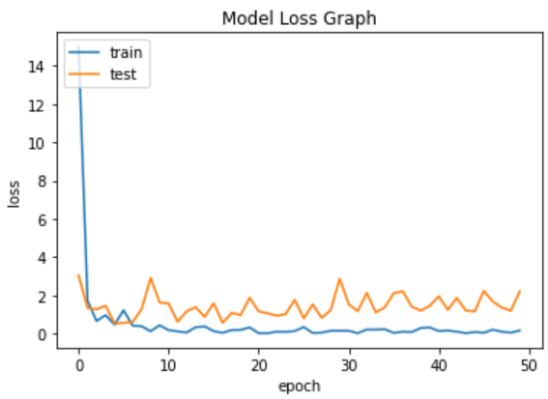
(e) Deep Convolutional Neural Network

Figure 4.4: Model Accuracy Graph

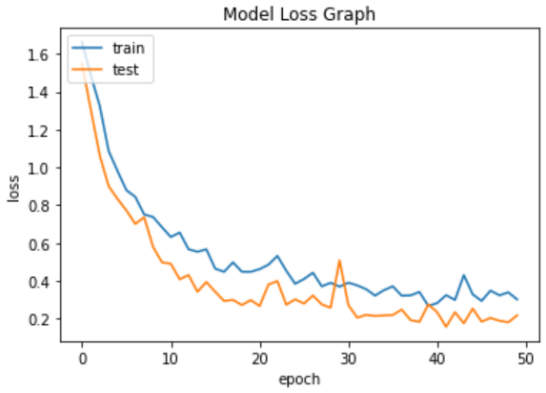
The model accuracy is mainly designed based on the total number of epochs and the accuracy graphs showed better performance with respect to increasing the number epochs it increased the training accuracy.



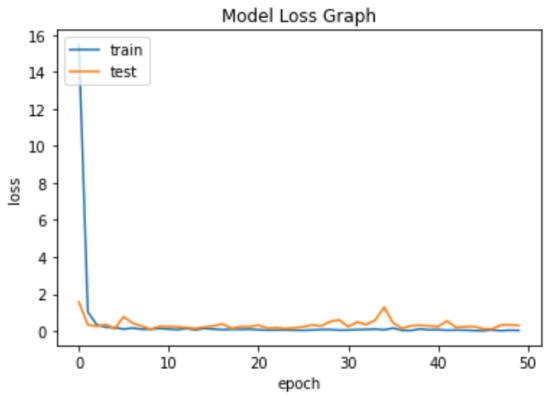
(a) ResNet152V2



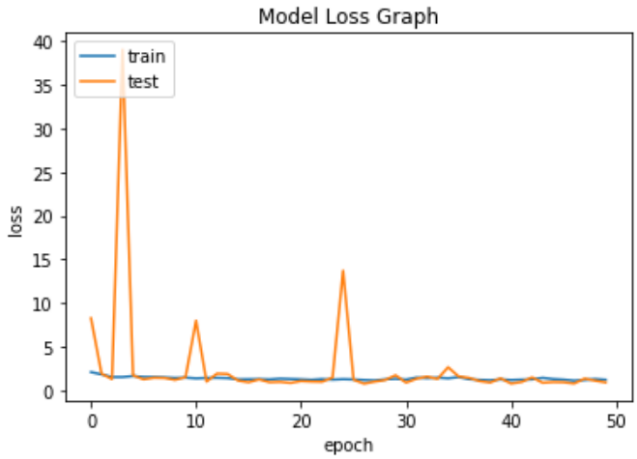
(b) NasNetLarge



(c) MobileNet-V2



(d) InceptionResNet-V2



(e) Deep Convolutional Neural Network

Figure 4.5: Model Loss Graph

Compared to the model's accuracy graph, it defines a strong performance of the proposed model because of decreasing the loss values when it increases the number of epochs.

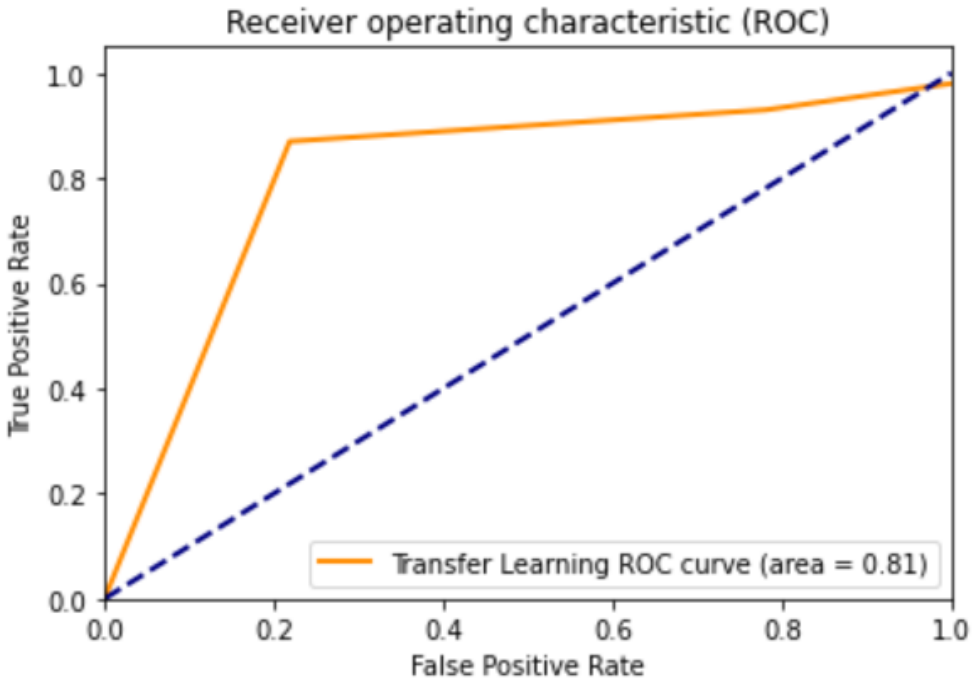
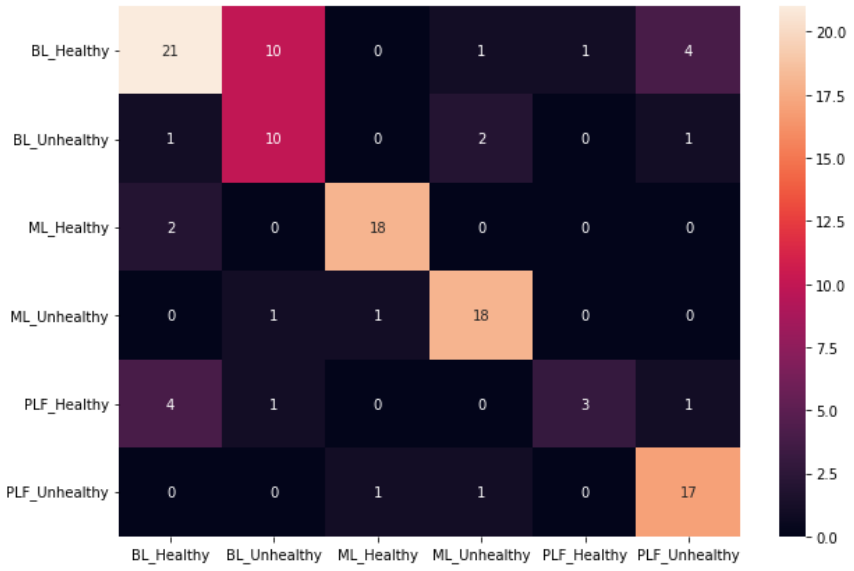
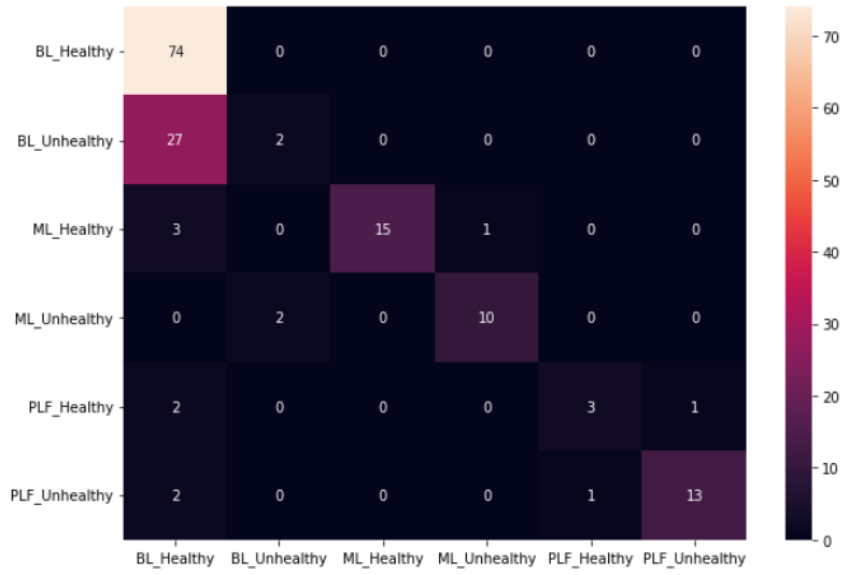


Figure 4.6: ROC Curve of InceptionResNet-V2

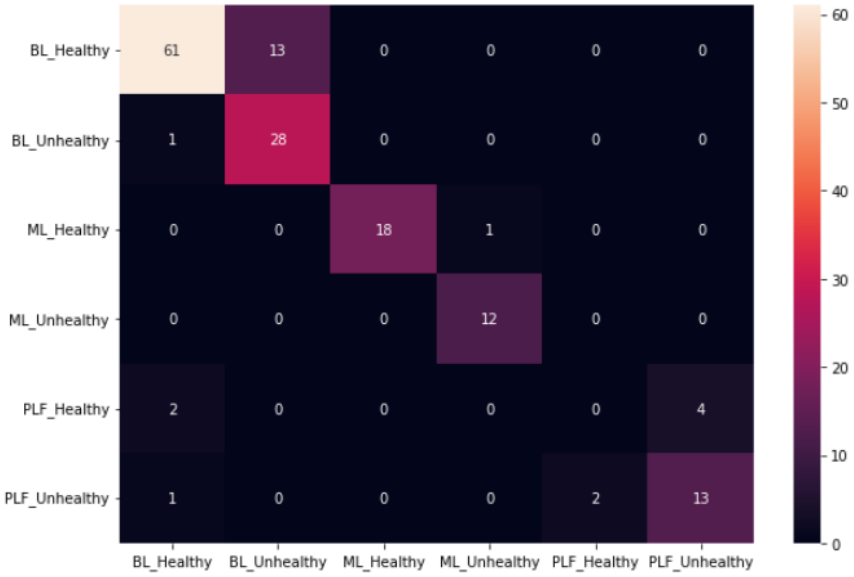
As a predictable vegetable disease approach, there have a probability of the desired observation belongs to the applied class as Bean Disease Leaves, Fresh Leaves, Malabar Disease Leaves, Malabar Fresh Leaves, Papaya Disease Leaves and Fruits and Papaya Fresh Leaves and Fruits. ROC Curve mainly plots it into two parameters like True Positive Rate and False Positive Rate which are mentioned in equations (6) and (7). This is the curve of InceptionResNet-V2 because of its best performance and the classification thresholds defined a acceptable performance as increasing both of True Positive and False Positive.



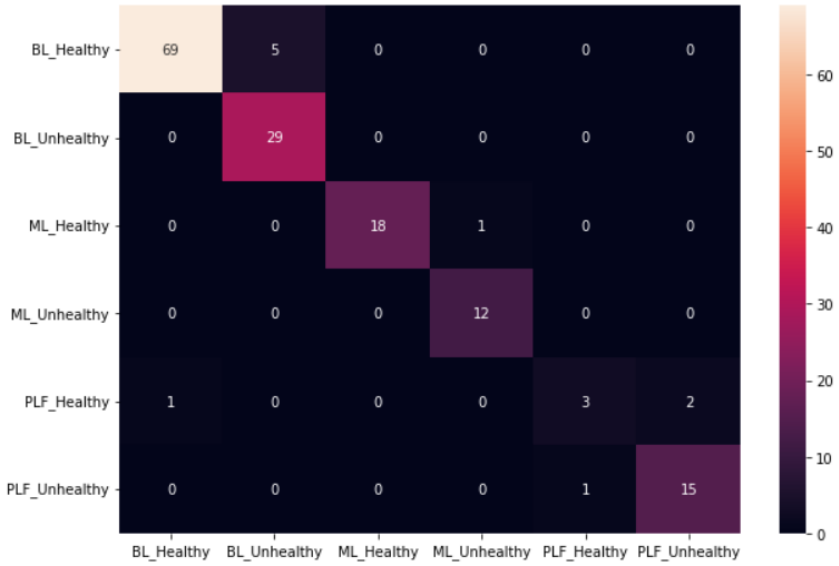
(a) ResNet152V2



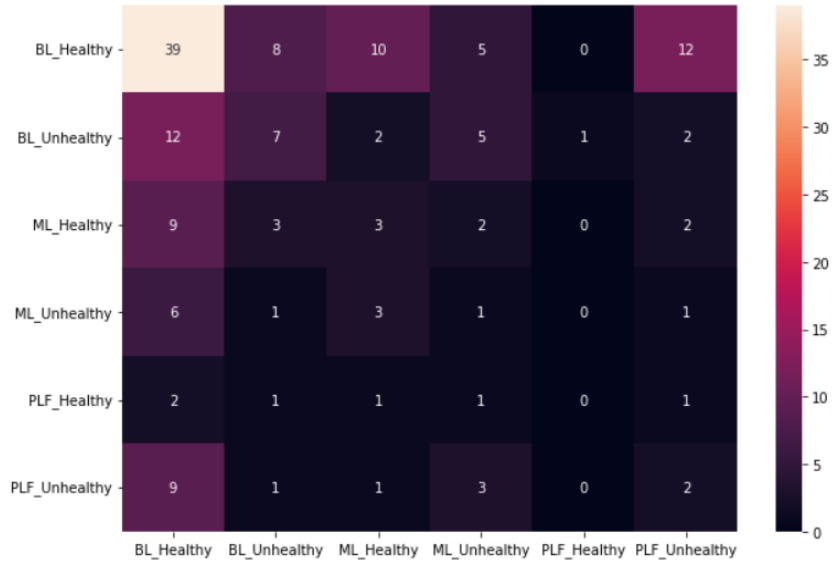
(b) NasNetLarge



(c) MobileNet-V2



(d) InceptionResNet-V2



(e) Deep Convolutional Neural Network

Figure 4.7: Confusion Matrix

Confusion matrix is the one of the better way to measure the performance of the applied models and there are some equations of Accuracy, Precision, Recall and f1 Score are given below for calculating the acceptances of the model,

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (4)$$

$$Precision(\%) = \frac{TP}{TP + FP} \times 100 \quad (5)$$

$$Sensitivity(Recall)(\%) = \frac{TP}{TP + FN} \times 100 \quad (6)$$

$$Specificity(\%) = \frac{TN}{TN + FP} \times 100 \quad (7)$$

$$f1\ Score(\%) = 2 * \frac{Recall * Precision}{Recall + Precision} \times 100 \quad (8)$$

TP:True Positive, TN: True Negative, FP:False Positive and FN:False Negative.

The above mentioned equations as (4),(5),(6),(8) have been used on the confusion matrix of ResNet152V2, NasNetLarge, NasNetLarge, MobileNet-V2 and Deep CNN which calculates different type of Accuracy, Precision, Recall and f1 Score are given in table.

Table-4.1: Performance Measurement of the Proposed Model

Algorithm	***Accuracy(%)***	***Loss values***	***Precision(%)***	***Recall(%)***	***F1 Score(%)***
ResNet152V2	73.11	0.74	74	72	70
NasNetLarge	90.38	2.92	80	67	69
MobileNet-V2	84.62	0.44	72	76	73
InceptionResNet-V2	93.60	0.41	90	89	89
Deep CNN	70.51	0.91	19	19	19

CNN based transfer learning models provide an acceptable performance as mentioned in the table and from the table InceptionResNet-V2 provides a most acceptable performance with better accuracy and lowest loss value.

4.3 Discussion

This proposed approach mainly defines the vegetables disease classify and recognition system based on transfer learning with respect to Deep CNN and it is mainly implemented to help the farmers and those scientists who want to explore more. There

have been used different types of keras pretrained and Deep CNN which provides different acceptable accuracy and pretrained model, InceptionResNet-V2 provides better accuracy as 93.60% with the lowest loss value. It is important to say that all of the pretrained model and Deep CNN were finalized after checking multiple layers, different sizes of images when it preprocessed the dataset to get better accuracy from the designed model. As a comparative approach, it also showed an acceptable model compared to existing models based on the accuracy, precision, recall and AUC values and also used the loss value to explore the acceptance. There have also been shown the confusion matrix to visualize the better.

CHAPTER 5

Impact On Society, Environment And Sustainability

5.1 Impact on Society

The ESRC (Economic and Social Research Council) has been defined as a demonstrative contribution of research in society and the economy. ESRC mainly talks about the impact on society like how it could make life easier and the uses of these techniques and also ensure that it would be much helpful economically in society and worldwide. There have been used the Deep CNN based keras pretrained model which provides the accuracy fast and takes a little bit memory, GPU to perform in proposed method and economically, it must be helpful because of the uses of this proposed method can be defines the vegetables disease and that's why farmers will be beneficially. In the early of the vegetables are affected by disease and most of the time farmers are failed to classify and recognize but after uses this proposed methods it would be easy to define the disease affected or not which also protect from time consume and economically helps to farmers.

5.2 Impact on Environment

There is no bad impact on the environment using this proposed approach. This is mainly designed to help those scientists and farmers who are working to develop these technologies which are used for improving the cultivation of agriculture ecosystem and also helps farmers from killing their times to recognize bean, malabar and papaya disease.

5.3 Ethical Aspects

There have some principles which is followed to designed the proposed approach as given below to collected the differents data like bean, malabar, papaya to make a uniques vegetables dataset,

- Bean disease and healthy leaves are collected by own capturing
- Malabar[citation_number] disease, healthy leaves and papaya[citation_number] disease leaves, healthy leaves, disease fruits, healthy fruits are collected from

existing articles by changes some attributes are mentioned in the reference section.

- Implemented differents algorithms to compared among them
- Used different layers and image target size to make it unique.
- used fixed epochs to trained the model

5.4 Sustainability Plan

The main purpose to designed this vegetables disease classification and recognition based on transfer learning is given below,

- This will help to recognize theses mentioned disease in the early of the vegetables
- To helps those farmers who cultivates these vegetables
- Helps those researcher and scientists who want get comparison result is known throughout this approach
- There is no any jeopardized by using the proposed model
- As the comparative model is provides the best and fast provides the better performance

CHAPTER 6

Summary, Conclusion, Recommendation And Implication For Future Research

6.1 Summary of the Study

The Vegetables disease classification and recognition approach is mainly designed based on Transfer Learning techniques with respect to Deep CNN. There have been used three different types of vegetables such as bean, malabar and papaya, totally unique dataset is collected from existing articles and by own capturing. The proposed model is mainly a comparative model among Deep CNN, MobileNet-V2 ResNet152-V2, NasNetLarge and InceptionResNet-V2 specifically provides better performance with 93.60% accuracy. The main reason to propose this model is to help those farmers who are cultivating vegetables like bean, malabar and papaya to detect the disease or healthy leaves and fruits and also be able to help the scientist and researcher who works on these vegetables.

6.2 Conclusions

Nowadays worldwide vegetables cultivation farmers are losing their profits because of affecting the vegetables by various diseases and that's why some are already taking a break to cultivate vegetables which is harmful for all.The proposed approach will be one of the best solutions to find the disease in any stage of cultivating vegetables with better performance. In this method there have been used bean, malabar and papaya to make vegetables and used five different transfer learning based algorithms as Deep CNN, MobileNet-V2 ResNet152-V2, NasNetLarge and specificity InceptionResNet-V2 provides the better accuracy as 93.60% as also a comparative model.

6.3 Implication for Further Study

The proposed approach will impact on vegetables cultivating disease recognition and classification methods with acceptable accuracy by comparing the performance with existing articles. The research article is mainly designed to help the farmer to recognize the vegetables disease and the researcher must be able to get the knowledge about the vegetable data as bean, malabar, papaya and also they will compare the result which will

help them for better scope. As a comparative approach, it will provide five different algorithms as Deep CNN, MobileNet-V2 ResNet152-V2, NasNetLarge and InceptionResNet-V2 performance with the accuracy, precision, recall, f1 score. There have been a lot of research article is search to define the proposed model like vegetable, tree, humans disease etc but for daily life a food is too important to alive and vegetable is too much healthy that's vegetable related research have been selected and bean, malabar, papaya is known as delicious foods worldwide.

Reference:

[1] Habib, M.T., Majumder, A., Jakaria, A.Z.M., Akter, M., Uddin, M.S. and Ahmed, F., 2020. Machine vision based papaya disease recognition. Journal of King Saud University-Computer and Information Sciences, 32(3), pp.300-309.

[2] Islam, M.A., Islam, M.S., Hossen, M.S., Emon, M.U., Keya, M.S. and Habib, A., 2020, November. Machine learning based image classification of papaya disease recognition. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1353-1360). IEEE.

[3] Kumar, V., Arora, H. and Sisodia, J., 2020, July. Resnet-based approach for detection and classification of plant leaf diseases. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 495-502). IEEE.

[4] Hossen, M.S., Haque, I., Islam, M.S., Ahmed, M.T., Nime, M.J. and Islam, M.A., 2020, December. Deep learning based classification of papaya disease recognition. In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (pp. 945-951). IEEE.

[5] Sarangdhar, A.A. and Pawar, V.R., 2017, April. Machine learning regression technique for cotton leaf disease detection and controlling using IoT. In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 2, pp. 449-454). IEEE.

[6] Mojumdar, M.U. and Chakraborty, N.R., 2020, July. A computer vision technique to detect scab on malabar nightshade. In 2020 11th international conference on computing, communication and networking technologies (ICCCNT) (pp. 1-4). IEEE.

[7] Ashok, S., Kishore, G., Rajesh, V., Suchitra, S., Sophia, S.G. and Pavithra, B., 2020, June. Tomato leaf disease detection using deep learning techniques. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 979-983). IEEE.

[8] Kumar, S., Prasad, K.M.V.V., Srilekha, A., Suman, T., Rao, B.P. and Krishna, J.N.V., 2020, October. Leaf disease detection and classification based on machine learning. In 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE) (pp. 361-365). IEEE.

[9] Moyazzoma, R., Hossain, M.A.A., Anuz, M.H. and Sattar, A., 2021, January. Transfer learning approach for plant leaf disease detection using CNN with pre-trained feature extraction method Mobilnetv2. In 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) (pp. 526-529). IEEE.

[10] Elfatimi, E., Eryigit, R. and Elfatimi, L., 2022. Beans leaf diseases classification using mobilenet models. IEEE Access, 10, pp.9471-9482.

[11] KIZILOLUK, S., 2021. Comparison of Standard and Pretrained CNN Models for Potato, Cotton, Bean and Banana Disease Detection. NATURENGS, 2(2), pp.86-99.

[12] Ajra, H., Nahar, M.K., Sarkar, L. and Islam, M.S., 2020, December. Disease Detection of Plant Leaf using Image Processing and CNN with Preventive Measures. In 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE) (pp. 1-6). IEEE.

[13] Abdu, A.M., Mokji, M.M. and Sheikh, U.U., 2020. Automatic vegetable disease identification approach using individual lesion features. *Computers and Electronics in Agriculture*, *176*, p.105660.

[14] Methun, N.R., Yasmin, R., Begum, N., Rajbongshi, A. and Islam, M.E., 2021. Carrot Disease Recognition using Deep Learning Approach for Sustainable Agriculture. *International Journal of Advanced Computer Science and Applications*, *12*(9).

[15] Sharma, M., Kaushik, P. and Chawade, A., 2021. Frontiers in the Solicitation of Machine Learning Approaches in Vegetable Science Research. *Sustainability*, *13*(15), p.8600.

[16] Javierto, D.P.P., Martin, J.D.Z. and Villaverde, J.F., 2021, November. Robusta Coffee Leaf Detection based on YOLOv3-MobileNetv2 model. In *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (pp. 1-6). IEEE.

[17] Sivakumar, P., Mohan, N.S.R., Kavya, P. and Teja, P.V.S., 2021, November. Leaf Disease Identification: Enhanced Cotton Leaf Disease Identification Using Deep CNN Models. In *2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT)* (pp. 22-26). IEEE.

[18] Naveenkumar, M., Srithar, S., Kumar, B.R., Alagumuthukrishnan, S. and Baskaran, P., 2021, November. InceptionResNetV2 for Plant Leaf Disease Classification. In *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 1161-1167). IEEE.

[19] Mehmood, M., Alshammari, N., Alanazi, S.A., Basharat, A., Ahmad, F., Sajjad, M. and Junaid, K., 2022. Improved Colorization and Classification of Intracranial Tumor Expanse in MRI Images via Hybrid Scheme of Pix2Pix-cGANs and NASNet-Large. *Journal of King Saud University-Computer and Information Sciences*.

[20] Karpagam, M. and Maheshwari, A., 2022. Novel Computer Vision approach for identifying diseased Tomato leaves by Classifying Leaf Images for Pest Identification using ResNet152V2 Architecture.

[21] Shutterstock, available at <<https://www.shutterstock.com/image-photo/basella-alba-malabar-spinach-isolated-on-228195319>>, last accessed on 10.09.2022.

[22] iStock, available at <<https://www.istockphoto.com/search/search-by-asset?assetid=622316024&assettype=image>>, last accessed on 11.09.2022.

A Transfer Learning Based Techniques: Vegetables Disease Recognition and Classification_191-15- 2354_Tanjila-Jahan

ORIGINALITY REPORT

18%	15%	10%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
2	Submitted to Daffodil International University Student Paper	2%
3	assets.researchsquare.com Internet Source	2%
4	www.coursehero.com Internet Source	1%
5	www.researchgate.net Internet Source	1%
6	Elhoucine Elfatimi, Recep Eryigit, Lahcen Elfatimi. "Beans Leaf Diseases Classification Using MobileNet Models", IEEE Access, 2022 Publication	1%
7	Husnul Ajra, Mst. Khairun Nahar, Lipika Sarkar, Md. Shohidul Islam. "Disease Detection of Plant Leaf using Image Processing and CNN with Preventive Measures", 2020 Emerging Technology in	1%