PLANT DISEASE CLASSIFICATION USING TRANSFER LEARNING METHODS

MSc Research Project
MSc. Data Analytics

Nikhil Awachat
Student ID: x21100446

School of Computing
National College of Ireland

Supervisor: Dr. Christian Horn

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Nikhil Awachat
Student ID:	x21100446
Programme:	MSc. Data Analytics
Year:	2021 - 2022
Module:	MSc. Research Project
Supervisor:	Dr. Christian Horn
Submission Due Date:	15/08/2022
Project Title:	Plant Disease Classification Using Transfer Learning Methods
Word Count:	7951
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Nikhil Awachat
Date:	14th August 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

PLANT DISEASE CLASSIFICATION USING TRANSFER LEARNING METHODS

Nikhil Awachat
21100446

Abstract

The majority of today's food comprises plant-based diets. Agriculture-based products are easily impacted by a variety of plant diseases. Farmers suffer social, ecological, and financial losses as a result of these infections. It becomes crucial to conduct a thorough and timely analysis of plant diseases. Some illnesses are readily detectable and accessible because they are apparent to human sight. In the past, manual examination of plant diseases was done by specialists in that field. This involves tremendous labor and takes a long time to process. To solve this problem an efficient and effective solution is needed. Processing all types of disease images with absolute accuracy can be simplified extremely thanks to neural networks. This can be implemented using Deep Learning with a combination of some data augmentation techniques. The neural network is helpful in the detection of crop diseases through which the risk factors for the disease can be reduced. In this research work, it is found that using VGG16 with data augmentation on the problem gives remarkable results in classifying plant disease, as the final accuracy of the VGG16 model achieved was 99.67%, and the average precision is 74%, average recall is 75%, and average f1-score is 74%. All the results and evaluations of the various models are compared with the previous research on the same dataset.

Keywords Transfer Learning, CNN, VGG16, Crop Diseases

1 Introduction

Plant disease may be defined as the abnormal physiological functioning of the plant which can be the result of the continued methods of irrigation which is harmful and makes them weak and makes them infectious and harmful to people. The two types of plant diseases that are caused are fungal and the other one is fungal-like organisms. The disease found in the plant is caused by pathogens, which are the bacteria that affect the plants. Different types of pathogens that are included in the cause of bacteria in the plants are viruses, fungus, viroid, or parasitic flowering plants. The disease that is caused in the plant is a combination of the different viruses that are formed together and make the plant infected (Farh et al.; 2018). Plant disease is generally identified by using methods that include microscopic evaluation whose features help in identifying the pathogens. Microbiological techniques are also used in the identification process.

Crop disease is the major issue that is faced by the farmers and after identification of the issue, several steps naturally are taken for removing the issues. Technological improvement has made it possible to develop new innovative techniques and machines that can identify the issues. Crop disease is the biggest threat not only to food but at the same time, has negative impacts on the lives of farmers who are very small and whose livelihood depends on the healthy crops that are sold in the market (Sharma and Chetani; 2017). Pesticides have been used for protecting the crops from getting infected but this has not given a good result for a long time. So, with the improved knowledge and skills, there have been clinics which are developed by the people where the plants can be taken for the clear identification of the disease that is inside it so that proper steps and measures could be taken. The use of smartphones has also contributed to identifying the diseases because they have high-resolution displays and HD cameras that gives a proper view of the information which is generated. The high performance of the cameras and the HD displays has made it easier in identifying a situation and diseases. The proper training has been given to the people in understanding the domains and helped in end-to-end learning. Neural networks provide an image of a diseased plant. Deep learning helps in identifying pathogens that are found in drought-prone regions, and the control of plant disease with germicidal seeds will reduce the disease in plants.

Therefore, the detection of diseases in crops is a crucial issue that offers several advantages for keeping a close watch on vast fields of crops. Wheat disease damages the inner layer of the leaflets, which might have an impact on both yield and standard. The fastest and most accurate technique to identify the illness type is to manage these wheat problems, followed by the prompt implementation of the necessary remedial measures. The identification of diseases is precise, fast, and convenient using machine learning algorithms and computerized methods for image processing. Improvement and enrichment of plant protection are opportunities provided by innovations in computer intelligence. In this study, we detect crop illnesses using several Transfer learning methods. Due to its sequential multilayer feature extraction capability, the convolution layer may extract characteristics that are more precise and pertinent. Deep learning algorithms will be used to detect crop diseases. Machine learning has some disadvantages as well because it needs a huge amount of datasets for training the models. Performance suffers when data is not big enough to handle adequate photos. Transfer learning provides a number of benefits, including the fact that it doesn't require a huge amount of datasets for training the models. Transfer learning enhances learning by transferring knowledge from previous knowledge, equivalent task to the current one at hand. (Hassan et al.; 2021) Many researchers applied the transfer learning method to detect diseases.

1.1 Motivation

The main motivation of this research was to check how the process of detecting crop disease can be made simple for farmers as they face a variety of problems when dealing with crop diseases and prevention as normally are not aware of what disease is affecting what plant, making it harder to implement a simple solution. The assignment was useful in giving a clear picture of the issues which are faced by the farmers because it is not known by people that there are many issues which are faced by the farmers for producing the crops as a variety of diseases are harming the quality of the food. This gave information about how these issues can be solved by using the latest technologies which have been made available to the market. This made people aware of the importance of plants and how they have been degraded because of the bacteria that are harming them to grow and also to give a healthy and safe environment for people to live in. So, it helped make us understand that the food that people get to eat easily comes to them after many problems that small farmers face and many difficulties in keeping the crops protected and healthy which is the only earning source for them. The Plants need to be protected as it is very important for the life of humans.

1.2 Research Question

How do neural networks based on VGG16 and other transfer learning models using data augmentation predict wheat crop diseases and how efficient are they if new crop data images are used?

1.3 Objective and Aims

1. To analyze how neural networks help predict crop disease, evaluate how VGG16 aids in predicting crop disease, and analyze the effect augmentation have on images.

2. To ensure the use of neural networks which is an important factor in identifying the disease through images? They provide a mapping between the networks and input an image of the diseased plant which might be infected at a high rate (Okafor et al.; 2018). The challenge is about creating a deep network where the edge weights correctly for giving a proper and justified output.

3. To analyze the different techniques and methods which can be implemented for getting better results and also in reducing the cost of buying chemical fertilizers (Yang et al.; 2019).

1.4 Plan of paper

In section 2 the literature review is done where the overall view of the study is understood after going through the facts and the information. The different themes have been introduced for understanding the exact requirements and the use of previous literature in section 2 makes a clear understanding of what needs to be implemented with the showing of the pictures where the issues can be seen through image information. The effects of augmentation on the images are also taken into consideration and the literature gap will help in informing the literature that has been missing in the given topic and how the explanation is done after resolving it.

Section 3 has the Methodology with the Model design and architecture, explaining why and which metrics and parameters were used in the models and their importance. Section 4 has the Implementation and the Evaluation of all the models built in the study by comparing their accuracy, validation accuracy, loss, and validation loss. Section 5 has a Discussion and model comparison of our models on balanced and imbalanced data and also the comparison of results from the previous study on similar data compared with the results of this research.

Last the conclusion about the given topic is made by understanding the exact facts and information that has been derived from the given topic and how it can be improved with different measures and necessary and proper steps.

2 Literature Review

This segment of study is going to discuss and critically analyze the previous and current research papers and the viewpoints of various authors on a similar topic to the topic of the given research. In terms of the crop protection system, it is a vital and required step for modern agriculture to accurately perform the diagnosis of plant disease. With the technological advent, in the modern days, machine learning and AI technology can efficiently diagnose and predict various crop diseases along with malnutrition and the issues of pests using various sensors. The traditional practices of the classification of plant disease are associated with testing in the laboratory or visual observation. However, these traditional practices may be time-consuming and erroneous and, in this scenario, machine learning-based approaches are used currently in a huge volume to perform the tasks of classification of plant-based diseases with the advent of data science. The image-based machine learning approach is generally used using a variety of algorithms such as Depp Convolutional Neural Network and Support Vector Machine. In these processes, the diseased images are generally captured from the diseased plants, and they are extracted through various image processing techniques (Shrivastava et al.; 2019). Then, they are fed into different classification techniques in terms of Neural Networks, Decision Trees, and SVM for the classification of the exact disease.

2.1 Previous Literature

Various previous research studies have been found to provide key insights into the prediction of crop diseases and their classification using machine learning approaches. A study by (Neforawati et al.; 2019) has revealed that the Convolutional Neural Network can detect plant diseases along with the prediction of the rate of growth of plantation through the extraction of features and analysis of those features in an automated manner. It is a part of precision agriculture and the usage of technological advent in this area can lead to the detection of early onset of disease so that the farmers can get alert and can apply the right farm input resources. It can result in the maximization of the production process by being more informed about crop health by the detection of the disease. This study has demonstrated that it can be possible to detect and differentiate sugar beet disease in terms of the vegetation spectral indices using a support vector machine with a classification accuracy of 97%.

Another study by (Sabrol and Kumar; 2016) demonstrated the detection of Tomato plant disease using Neural Network models by the extraction of natural outdoor images through soft computing techniques. Various diseases of plants by various agents such as bacteria, viruses, and fungi can lead to a negative impact on plant growth and crop production every year. The proper identification and classification method of plant disease can lead to the limitation of agricultural and economic loss by helping to take actions to prevent such losses. The computerized automated diagnosis method can lead to the better detection of the plant disease based on the symptoms reducing the dependency on experts. Pattern recognition methods can be used in this sector using Support Vector Machines, Neural Networks, and Adaptive Inference systems. The plant disease images, in this context, can show various features including color features, shape features, and texture features.

According to another study by (Ferentinos; 2018), it is evident that the deep Neural based recognition model of a plant disease classifier can work using image classification techniques exploiting the processing techniques of common digital images combing different methods for better feature extraction. For the detection of diverse plant diseases using CNN, it is required to rain and fine-tune the neural network so that it can fit accurately to the database of the plant leaves resulting in the generation of a simplified and developed approach enabling easy differentiation between the healthy eaves and diseased leaves.

2.2 How do neural networks predict crop disease?

The neural network and its technology are very helpful in terms of the diagnosis of the various plant diseases, and prediction of the plant diseases and are also used for the detection of malnutrition in the agricultural sector. According to (Abiodun et al.; 2018), neural network applications are comparable with real-world scenarios. It helps connect with the artificial neural networks through which the emerging trends and the application are been furnished (Abiodun et al.; 2018). Plant diseases and pest detection is one of the important aspects when the field of machine vision is been enhanced. The neural network is been adjusted where the images are been taken and then it is been classified through the help of the collection of the plant images. According to (Liu and Wang; 2021) in comparison to the definite classifications, the detection and segmentation tasks are been analyzed through the help of the neural networking of computer image detection. There is the presence of an advanced level of AI technology is one of the promising parts, where greater improvement is been done concerning the prediction of crop diseases, enhanced and various power tools are been engaged. The tools, open-source tools for deep learning are among the effective tools. Under neural learning, the open-source tools which are been engaged include, Tensorflow, Torch/PyTorch, Caffe, Theano and many more. According to (Ramírez-Sánchez et al.; 2020), The open-source tools of the Python toolkit exponentially support various processing. The Python API model of the open-sources tools helps in company line tools which helps in the thin wrapping process (Ramírez-Sánchez et al.; 2020). The various tools namely PyTorch and Tensorflows are adjusted with the Android, IOS, Windows, and Linux through which it became more comparable to use and it also supports the third-party libraries towards providing the deep network structure creation. Through the help of the tools, the detection of crop diseases can be enhanced.

There is the presence of various materials and model approaches through which the identification and prediction of crop diseases can be enhanced and the disease risk can be dynamically reduced. According to (Newlands; 2018), Integrated Modeling Approach is one of the important frameworks where it is designed to take the account majoring aspects and is also considered to involve the operational model forecasting concerning crop disease identification. This kind of approach helps in a combination of the data concerning the host and environment and through the help of the model through which the capture of various dynamic diseases under the assumption can be done (Newlands; 2018). Integrated pest management is one of the essential parts of the approach where crop control management are been engaged and the inclusion of the technology is helpful in social and economic environmental management (Rezaei et al.; 2020). The management also incorporated the neural networking process through which the management can able to detect and identify the pest control, the diseases and the risk factors can be reduced.

2.3 In what ways does VGG16 aid in predicting crop disease?

The CGIAR wheat crop disease, which contains three categories of images which include healthy, leaf rust, and stem rust, was used in paper by (Sood et al.; 2022) to classify wheat diseases using a neural network VGG16 transfer learning algorithm. By adjusting the different parameters such as epoch size, batch, and activation functions, neural network models gave them good performance. With 0.01 as the initial learning rate that decelerated to 0.0001, the developed approach showed an excellent evaluation overall accuracy of 99.54%. When results were analyzed on various metrics, they observed that despite these algorithms' great precision, they remained ineffective for classifying leaf and stem rust. This is because several photos in this data had multiple diseases present; for example, one image included traits from both stem rust and leaf rust.

According to (Achanta et al.; 2012), there are various ways in which VGG16 can help in predicting plant diseases. VGG16 is a neutral network that is 16 layers deep. The detection of plant diseases cannot be seen with the human eyes. The leaves that are affected can be detected with the help of this system automatically. This classifies the diseases of the plants by showing the images of the leaves by a particular technique. The diseases that are detected can be very much dangerous for the soil and can cause serious problems to the farmers and can also hamper the agricultural process. Plant diseases are a major threat to the agricultural sectors, and the identification of these diseases is beneficial for the crop-growing industries. VGG16 comes very effectively for the identification of the diseases of the plant's leaves. This also allows for aiding the plants on time. With the use of this many crops that are consumed by humans daily can be stopped from getting written or affected. With the consumption of the affected crops, human health gets affected.

There are certain advantages of using this technique such as VGG16 allows for automatic detection of plant diseases and recognition of the diseases; it also gives an accurate answer to the required problems. The accuracy level of the diagnosis is also high and the cost of operation is also reduced. The affected leaves are also detected automatically, saving the leaves means saving the plant. This process is the simplest of all the other processes and it does not require any special training. With the automatic detection of the leaves helps to understand the disease of the plants. The images of the leaves are also shown larger and use more than a million parameters. Thus, it can be said that it has a huge network of parameters. VGG16 also allows for rapid quality screening of the leaves(Genaev et al.; 2020).

Detection of diseases in a plant by the naked eye requires a continuous monitoring process and expert observations (Arya et al.; 2018). The number of cultivated crops is huge in number and even the diseases slip through the eyes of experienced pathologists and agronomists. Therefore, the VGG16 model of neural networks helps in predicting and forecasting diseases of plants (Dhaka et al.; 2021). Identification of issues promptly helps in halting the quantitative and qualitative loss of plant matter and thus saving millions. Manual identification of diseases in plants requires a lot of time and is also expensive. Thus, the process of identification based on automation impacts the production quality greatly. The different diseases are identified by seeing the symptoms that show themselves on the leaves. The diseases are mainly caused by unusual temperatures, abnormal humidity, and varying soil moisture. The diseases of the plants can be classified according to three factors one being bacterial, viral, and fungal disease. Extraction of various factors by segmentation of leaves and then finding the region affected by the disease through a complex background imaging technique is challenging. These may cause unreliable results and false predictions due to human error. Therefore, deep learning techniques like the VGG16 which uses neural networks have recently gained popularity in the agricultural sector. Different architectures like the convolutional neural networks like the VGG16 help in the automatic extraction of plant features.

The VGG16 model automatically extracts the features like the color of the plant sections, leaves, texture, edge, and many more. However, a neural network made up of complex structures poses the negative effect of consuming high memory and computational resources. The Visual Geometry Group or VGG16 is based on several deep learning models like the ResNet, Inception, Xception, and DenseNet. Among all these models VGG is simpler and less complex to use. The VGG model has been developed by (Bhatt et al.; 2019), who included a network of many convolution layers. Pooling layers are also incorporated with many different filters. The two models that VGG provides that are VGG16 and VGG19 have 16 layers and 19 layers respectively. The parameters that the VGG network can generate are over 140 million. The training of the network is based on a huge dataset of 1000 categories.

2.4 What effect does augmentation have on images?

Originally, people relied mostly on their senses of sight to identify plant diseases. It is frequently hard, tedious, and inaccurate. To enhance disease identification processes, deep learning techniques linked to crop pictures were developed. CNN has gained popularity and has shown to be quite efficient. CNN has resulted in good and accurate classification results, but the problem of small dataset persists. With the focused action direction and create the standard networking system with augmentation process with prediction accuracy and maintain focused address the different disease from the plants. Experiments in this context follow the idea of enhanced support plans and cater to constructive directions so that concentrative assistance can be supportively engaged (Aboneh et al.; 2021).

It is extremely usual to augment the original data with computerized generated images, so this technique is applied throughout the model framework for data that is relatively small. (Urva; 2021) presents that synthesized face images and the composing information directions here connect with automatic action justifications with the targeted action criteria. State-of-art-method with the following dataset management criteria generates an overall 7% margin and this also merges the training set by accessing the current domain and its overall performance improvement by 2%. () proposed the rendering pipeline and considered realistic approaches which follow the accuracy action system prospect and generate proper direction towards real images and its assistance.

2.5 Literature gap

This literature review has been done by critically analyzing the viewpoints of various authors about the necessity of plant disease detection and how machine learning and AI technologies can be used effectively in this context in terms of image processing and classification. However, the literature review did not focus on the disease detection and classification systems and methods in different plants separately and different processes of detection in potatoes, tomatoes, pomegranates, and others. Moreover, this literature review found that not many discussed the various diseased conditions of the plants separately and did not focus on the various symptoms of plant disease such as black spot, crown gall, bacterial will, or others.

From the aforementioned discussion and analysis, it can be concluded that the early identification of the pant disease is one of the vital tasks of the agricultural background to increase the growth rate of the crop plants and improve the production yield. The Transfer learning approach as a subcategory of Deep learning can be effectively used in this prediction system as a part of precision farming as it has been found to show more accuracy as compared to the traditional methods of plant disease classification. CNN in terms of image processing and classification can be efficiently used as well-known DL architectures, as shown in Table 2.

3 Methodology

This section of the methodology shows details about the dataset used in this research and all the steps used to clean, augment, and preprocess the dataset. This research has been conducted by following the steps of the KDD (Knowledge Discovery in Databases) by following all the stages of the KDD process.

3.1 Knowledge Discovery in Databases (KDD)

Obtaining relevant, non-trivial data from large datasets is an iterative, multi-stage process defined as the KDD process in general. The user is given several options at every step of the process, all of which have the possibility of significantly influencing the research's conclusion. The early steps of a KDD process are highlighted in this technique, which is presented as an action plan. It also demonstrates how careful preparation may result in productive and efficiently managed work.

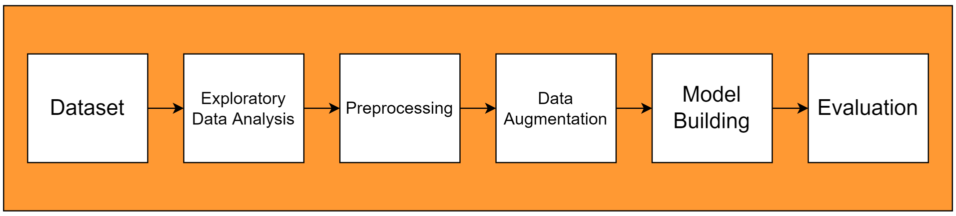


Figure 1: KDD process with all the stages

The next sections include detailed descriptions of all the different factors used at each step, from data collection and specification through implementation.

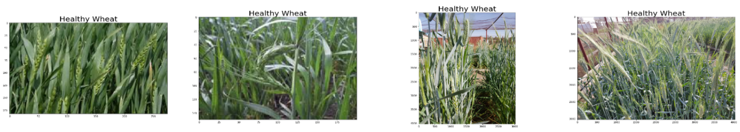
3.2 Dataset Acquisition

The data used in this research is taken from Kaggle and is the same dataset used in the reference paper (Sood et al.; 2022) on which this work is based. The dataset contains 1486 images in total, which are split into training and test folders. The train data has 876 images distributed within 3 classes. The images are available in good quality but are in different shapes and sizes. The original Data for the images is obtained from many resources. Almost the large majority of the data was acquired in Tanzania and Ethiopia and collected by the International Maize and Wheat Improvement Center.

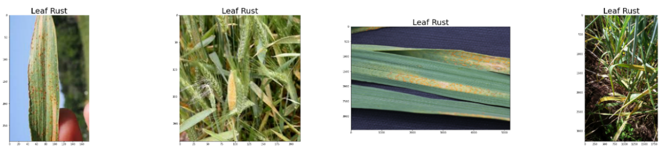
3.3 Exploratory Data Analysis

An EDA (Exploratory Data Analysis) is a detailed analysis designed to investigate a data set's underlying structure. It is significant for research because it shows patterns, trends, and correlations that are not intuitively clear. There are three classes present in the dataset i.e., Healthy wheat, Leaf rust, and Steam rust but they are not equally distributed. The Leaf rust and Steam rust class have almost twice the number of images as the Healthy wheat images, which may cause an imbalance. The pictures in this data contain unique characteristics including being entirely colored, having a variety of shapes, having distinct orientations, having varying quality, and being taken at various resolutions. A sample of each class in the dataset is shown below.

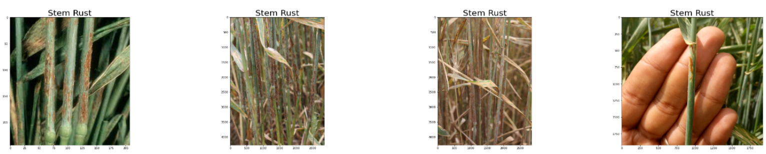
3.3.1 Healthy Wheat - 142 Images



3.3.2 Leaf Rust - 358 Images



3.3.3 Stem Rust - 376 Images



3.4 Data Preprocessing

This step is important in building good deep learning models as the data needs to be in the correct format, shape, and size to be used further in the process. The images also need to be in the correct folder and use the correct filename, as this denotes the specific label or class of the wheat. As we saw in the EDA step, the images were of different shapes and sizes, and to solve this issue, all the images were rescaled.

As the dataset doesn't have a different validation folder with our three classes, a step was added to create one with first creating two folders for train and test. Then the original dataset was split into a ratio of 75/25 into train and test. This step is necessary as this validation/ test data will be used in model evaluation which will show us if the model is performing accurately or not.

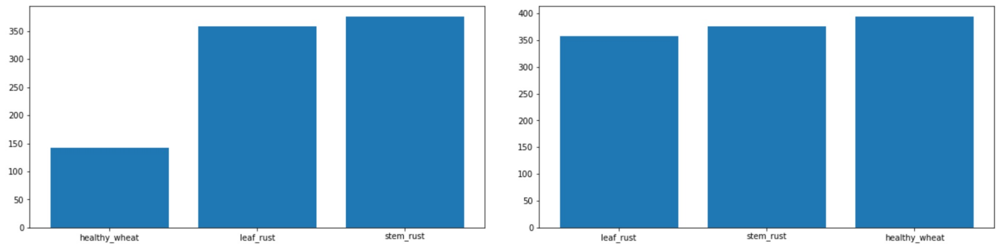


Figure 2: Dataset Distributions for imbalanced and balanced

The models were run on two types of datasets i.e., balanced and imbalanced, as the original dataset is imbalanced it was balanced by distributing the images equally in all the classes.

3.5 Data Augmentation

To make sure that the model doesn't overfit, several data transformation and augmentation techniques were used on the training data to get good accuracy on the validation test set. All the images are resized to a ratio of 224 * 224 as the original images are all in different dimensions, increasing the difficulty of training a model and which would have also required a considerable amount of computation power. By dividing the photos by 255 during the normalization process, the picture will be reproduced and use a scale of 0–1.

3.6 Model Building

This step includes building the model based on the different specifications. The four models built in this research are VGG16, CNN, AlexNet, and GoogleNet. To decrease the complexity and computation power needed for the models some parameters are made non-trainable which decreases the total number of usable parameters. All the model specifications and other details are shown in the design specification and implementation sections of the report.

3.7 Evaluation

To check the performance of the models, various metrics are used such as Accuracy, Loss, Sensitivity, and Specificity. The ability of the algorithm accurately categorize wheat diseases is called sensitivity, commonly referred to by the true-positive rate. The objective of this work is to categorize photos into certain diseases, hence models with good sensitivity are required. A loss function analyzes how effectively the neural network simulates the training data by comparing the target and expected correct output. Accuracy is the percentage of accurate predictions made by the algorithm.

4 Design Specification

This section gives more detailed explanations of the construction of every approach and the design of each model because this research used four different models. A typical

VGG16 model is made up of many layer types that enable the algorithm to learn and retrieve the characteristics of those categories.

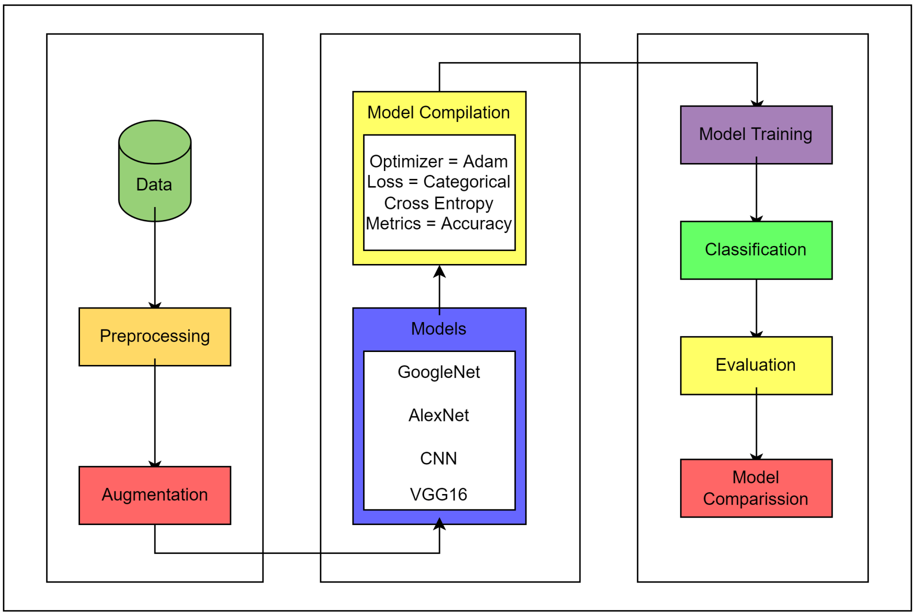


Figure 3: Model Architecture and Specification

Figure 3 displays the algorithm's design and architecture. The model has been broken down into three sections, the first of which includes a section for data preprocessing, transformations, and augmentation. The 2nd layer includes sections called Models and model compilation. The third section includes the categorized outputs from the models, model training, evaluation, and model comparisons.

The first section of the architecture includes the dataset, preprocessing, and augmentation. In the preprocessing step the original dataset was split into a ratio of 75/25 into train and test. This step is necessary as this validation/ test data will be used in model evaluation which will show us if the model is performing accurately or not. In the data augmentation step, all the images are resized to a ratio of 224 * 224 as the original images are all in different dimensions, increasing the difficulty of training a model and which would have also required a considerable amount of computation power. The models were run on two types of datasets i.e., balanced and imbalanced, as the original dataset is imbalanced it was balanced by distributing the images equally in all the classes.

The second section of the architecture includes building the model based on the different specifications. The four models built in this research are VGG16, CNN, AlexNet, and GoogleNet. To decrease the complexity and computation power needed for the models some parameters are made non-trainable which decreases the total number of usable parameters. The number of non-trainable parameters is higher in number as compared to the trainable parameters. The model compilation step has the optimizer set to Adam because it creates an optimization technique that can solve sparse gradients in noisy situations by combining the best features of both the RMSProp and Adaboost algorithms. The Loss function is set to Categorical Cross Entropy because this is a multi-class classification problem and evaluates how well a classification model performs when it produces a probability value ranging from 0 and 1. The metric is set to Accuracy because indicates the model's overall performance and is helpful if all categories are equally important.

The last section of the architecture includes model training, classification, Evaluation, and Model Comparisons. To train and test the model, the original data is split into a ratio of 75:25 as the train and validation set. In the classification step, the results are classified into three classes. In the Evaluation Step results of all the different models are evaluated using various graphs. First, we check the accuracy, validation accuracy, loss, and validation loss of all the models along with the same in a progression graph comparing all the values. Then we check the confusion matrix of each model to check the Sensitivity and Specificity values which give the true positive and negative values. Finally, the comparison of all the models is done to check which model is performing better than which model and in which metric.

5 Implementation and Evaluation

This section shows the implementation of all the models that were built for the research, with all the explanation of the model parameters and the final results, by giving a side-by-side evaluation of each model. Some parameters are made non-trainable to reduce the complexity and compute power required for the models, reducing the overall number of usable parameters.

5.1 Environment Setup

To train the models, a system with a Ryzen 5500U CPU and 8 GB of RAM is used. The system GPU is AMD Radeon 6 graphics card. The primary programming language used for building models is Python version 3.10.6 written on the Jupyter Notebook as the IDE. The Keras API and TensorFlow were used to create all the models. The optimizer is set to Adam, the loss function to Categorical Cross Entropy, the metric to Accuracy, and the learning rate to 0.01 in all the models. The Epoch size is set to 80 for all the models and the batch size for the imbalanced data is set to 30 images per epoch, and for the balanced data batch size is set to 35 images per epoch.

5.2 VGG16 Model

5.2.1 Implementation of the VGG16 Model

Based on convolutional neural networks, the first model that is implemented is called VGG16, which has 16 layers in total. We can also load a pre-trained version of the VGG16 which has been trained on the ImageNet database, which has around a million images. The Sixteen in VGG16 stands for sixteen weighted layers. VGG16 includes thirteen convolution layers, five max-pooling block, and three dense fully connected block which has one SoftMax block with the total classes i.e., three, a flatten layer, and a total of 21 layers make up VGG16, although only 10 of them are weighted layers, also known as trainable parameter layers.

All the layers are imported from the TensorFlow library using Keras. First, the model is built using the default configuration to see how the model performs. Then after iterating over various configurations, it is noted that the VGG16 with less trainable parameters gives the optimal results.

5.2.2 Imbalanced Dataset Results

The average accuracy and validation accuracy for VGG16 algorithm on the imbalanced dataset is 99.67% and 77.52% respectively. The average loss and validation loss for VGG16 algorithm on the imbalanced dataset are 0.23% and 0.89% respectively.

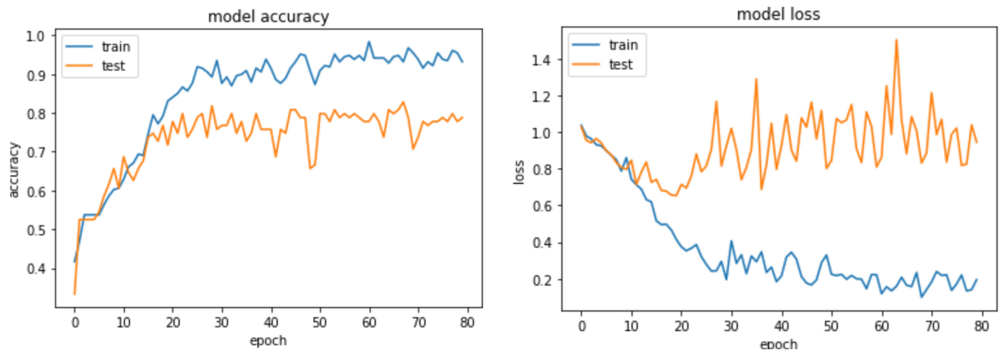


Figure 4: VGG16 model accuracy and loss results on Imbalanced Data

5.2.3 Balanced Dataset Results

The average accuracy and validation accuracy for VGG16 algorithm for the balanced dataset is 98.62% and 76.43% respectively. The average loss and validation loss for VGG16 algorithm for the balanced dataset are 0.13% and 0.83% respectively.

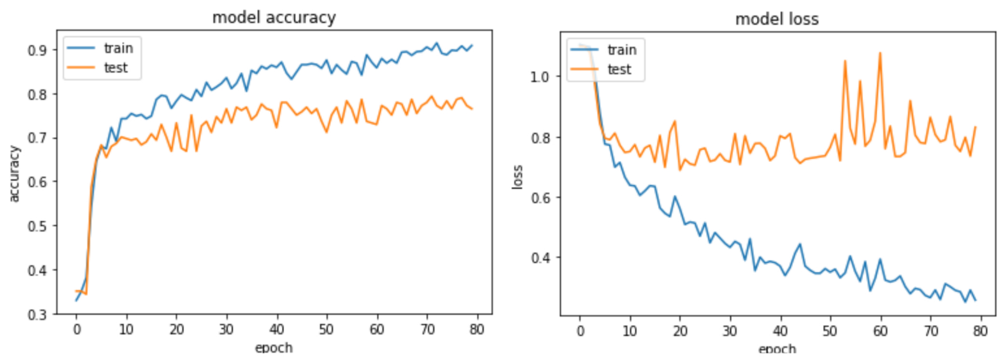


Figure 5: VGG16 model accuracy and loss results on balanced Data

5.3 CNN Model

5.3.1 Implementation of the CNN Model

The second model that is implemented is called CNN, which has 10 layers in total. The CNN model includes three convolution block, three max-pooling block, and two dense fully connected block; a SoftMax Dense block with the total classes, i.e., three, and a dropout layer, making a total of 10 layers make up the CNN model, and all the layers are weighted layers, also known as trainable parameter layers.

The CNN model was built by iterating the optimizer from 'rmsprop' to 'Adam', and it was noticed that the CNN algorithm performs better with the 'rmsprop' optimizer with metrics set to 'Accuracy'. The final accuracy of the CNN model is a little less compared to VGG16 algorithm, and the validation accuracy is also a little less.

5.3.2 Imbalanced Dataset Results

The average accuracy and validation accuracy for CNN algorithm on the imbalanced dataset is 99.55% and 71.01% respectively. The average loss and validation loss for the CNN algorithm on the imbalanced dataset are 0.01% and 1.82% respectively.

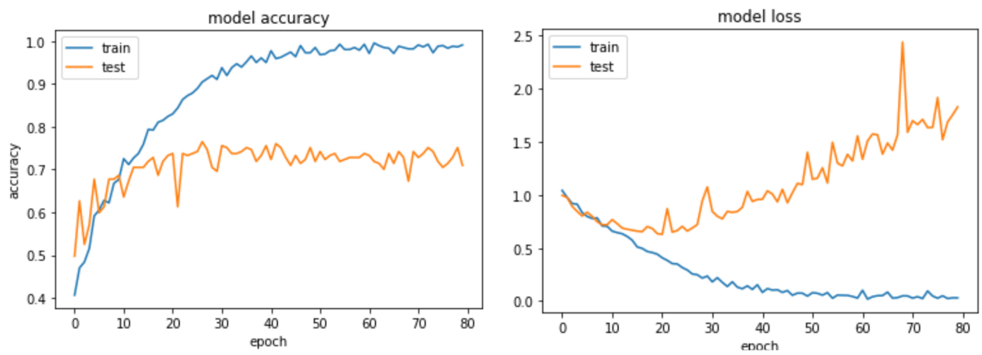


Figure 6: CNN model accuracy and loss results on Imbalanced Data

5.3.3 Balanced Dataset Results

The average accuracy and validation accuracy for the CNN algorithm for the balanced dataset is 98.23% and 73.57% respectively. The average loss and validation loss for CNN algorithm for the balanced dataset are 0.07% and 1.22% respectively.

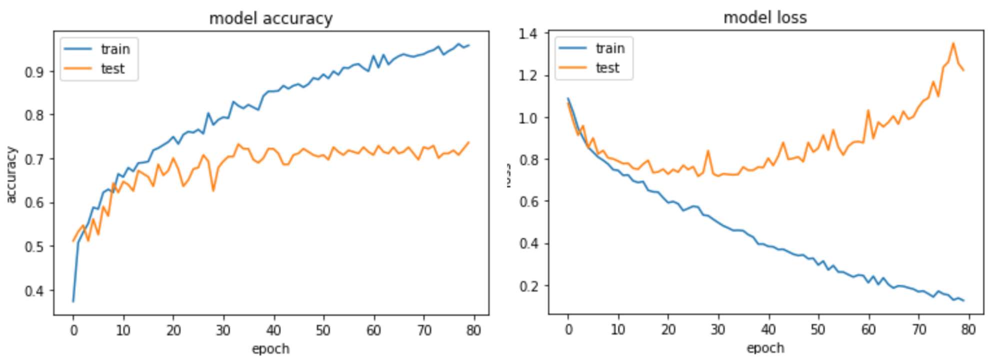


Figure 7: CNN model accuracy and loss results on balanced Data

5.4 AlexNet Model

5.4.1 Implementation of the AlexNet Model

The third model that is implemented is called AlexNet, which has 19 layers in total. The AlexNet model includes five convolution block, three max-pooling block, five batch normalization block, and two dense fully connected block; a SoftMax Dense block with the total classes, i.e., three, and a dropout layer, making a total of 19 layers make up the AlexNet model, and all the layers are weighted layers expect the Dense layers, also known as trainable parameter layers.

The AlexNet model was built by setting the optimizer to 'Adam', and it was noticed that the AlexNet model performs better with the learning rate been fixed to '0.001' with metrics set to 'Accuracy'. The final accuracy of the AlexNet model was a lot less stable as compared to the VGG16 model.

5.4.2 Imbalanced Dataset Results

The average accuracy and validation accuracy for the AlexNet algorithm for the imbalanced dataset is 82.14% and 67.82% respectively. The average loss and validation loss for the AlexNet algorithm for the imbalanced dataset are 0.64% and 1.16% respectively.

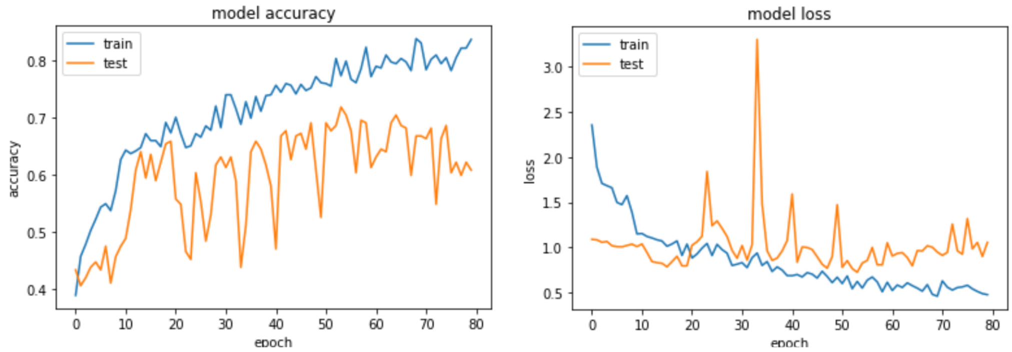


Figure 8: AlexNet model accuracy and loss results on Imbalanced Data

5.4.3 Balanced Dataset Results

The average accuracy and validation accuracy for the AlexNet algorithm for the balanced dataset is 80.66% and 65.71% respectively. The average loss and validation loss for the AlexNet algorithm for the balanced dataset are 0.42% and 1.03% respectively.

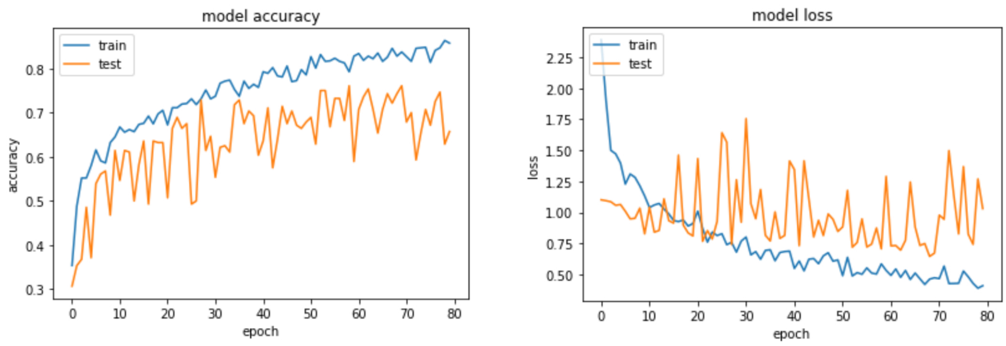


Figure 9: AlexNet model accuracy and loss results on balanced Data

5.5 GoogleNet Model

5.5.1 Implementation of the GoogleNet Model

The fourth model that is implemented is called GoogleNet, which has 22 layers in total. The GoogleNet model includes five convolution block, four max-pooling block, nine inception block, and three dense fully connected block; a SoftMax Dense block with the total classes, i.e., three, and a dropout layer, making a total of 22 layers make up the GoogleNet model, and all the layers are weighted layers expect the Dense layers, also known as trainable parameter layers.

The GoogleNet model was built by setting the optimizer to 'Adam', and it was noticed that the GoogleNet model performs better with the learning rate been fixed to '0.001' with metrics set to 'Accuracy'. The final accuracy for the GoogleNet model is less stable compared to that of the VGG16 model.

5.5.2 Imbalanced Dataset Results

The average accuracy and validation accuracy for the GoogleNet algorithm for the imbalanced dataset is 98.34% and 76.53% respectively. The average loss and validation loss for the GoogleNet algorithm for the imbalanced dataset are 0.17% and 1.54% respectively.

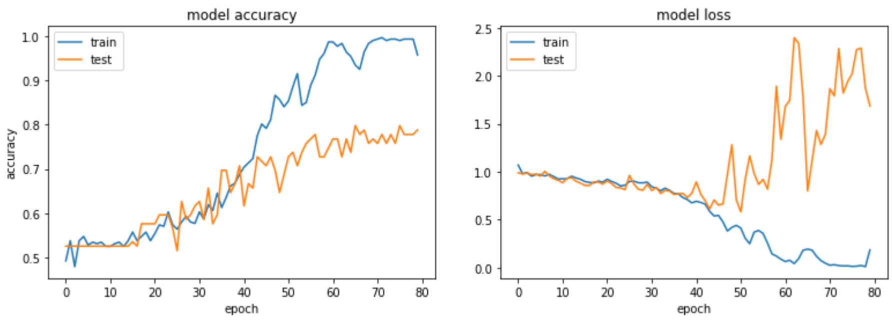


Figure 10: GoogleNet model accuracy and loss results on Imbalanced Data

5.5.3 Balanced Dataset Results

The average accuracy and validation accuracy for the GoogleNet algorithm for the balanced dataset is 98.25% and 69.38% respectively. The average loss and validation loss for the GoogleNet algorithm for the balanced dataset are 0.21% and 1.42% respectively.

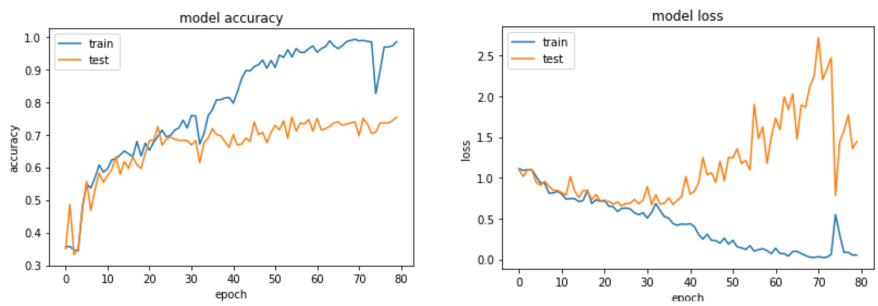


Figure 11: GoogleNet model accuracy and loss results on balanced Data

6 Discussion and Model Comparisons

In this section, all results from different models are compared with each other to check which models perform the best with similar parameters.

Model	Imbalanced Dataset		Balanced Dataset	
	Train	Test	Train	Test
VGG16	99.67%	77.52%	98.62%	76.43%
CNN	99.55%	71.01%	98.23%	73.57%
AlexNet	82.14%	67.82%	80.66%	65.71%
GoogleNet	98.34%	76.54%	98.25%	69.38%

Table 1: Comparing all the model accuracy on balanced and Imbalanced data

The above table shows the accuracy of all the different models built based on both the balanced and imbalanced datasets. The training accuracy of all the models in the imbalanced dataset is almost similar and shows a similar kind of trend. The validation accuracy on the other hand changes a little bit depending on the model. As shown in the above image, the VGG16 model performs the best out of all the models built on the wheat crop disease dataset giving the highest accuracy and validation accuracy. The VGG16 and CNN models take the most amount of time and computation power, taking almost 5.3 hours each. The AlexNet and GoogleNet model take less time and computational power than the other two, taking almost 2.5 hours each.

Author	Dataset	Dataset Same/Different	Accuracy	Model
(Sood et al.; 2022)	CGIAR Crop Diseae	Same	99.54%	VGG16
(Aboneh et al.; 2021)	Wheat Crop Disease	Similar	96.48%	VGG16
			95.65%	Inception V3
			99.30%	VGG19
(Pandian et al.; 2019)	Plant Leaf Disease	Different	87.03%	VGG16
(Bhatt et al.; 2019)	Wheat Crop Disease	Similar	96.60%	VGG19
(Genaev et al.; 2020)	CGIAR Crop Diseae	Same	98.10%	Densenet201
This Research	CGIAR Crop Diseae	Same	99.67%	VGG16
			99.55%	CNN
			82.14%	AlexNet
			98.34%	GoogleNet

Table 2: Comparing all the results from previous studies to our model

The above table shows the model comparison between our model and the models built in various research works. Some of the research work used the same dataset used in this research i.e., the CGIAR wheat crop disease dataset, while some of them had a similar dataset of wheat diseases but taken from a different source.

The below image of the classification report includes the values such as precision, recall, and f1-score for the VGG16 model test data. The average precision is 74%, average recall is 75%, and average f1-score is 74%.

The confusion matrix in the below image shows that the true positive for all the classes is way higher than the true negative for our validation test dataset, which shows

	precision	recall	f1-score	support
0	0.82	0.91	0.86	65
1	0.65	0.58	0.62	48
2	0.73	0.70	0.71	57
accuracy			0.75	170
macro avg	0.73	0.73	0.73	170
weighted avg	0.74	0.75	0.74	170

Figure 12: Classification Report of the final VGG16 model Test Data

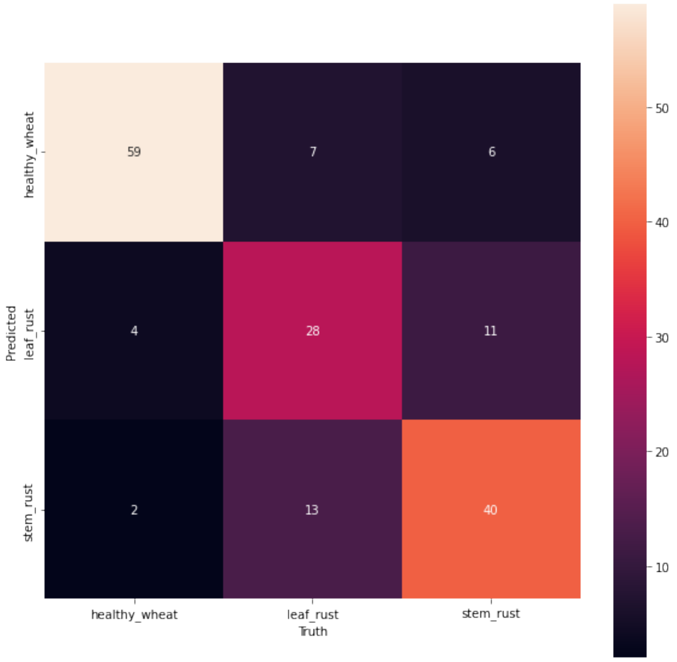


Figure 13: Confusion Matrix of the final VGG16 model Test Data

that the final VGG16 model is showing good performance. As we can see many of them used VGG16 as their primary model for the dataset and got similar kinds of results. The reference article used in the research by (Sood et al.; 2022) got an accuracy of 99.54% and the final accuracy of our model was 99.67% which indicated our research was accurate and was able to get similar results.

7 Conclusion and Future Work

In this research, to classify the crop diseases, a deep learning methodology was used by using the VGG16, CNN, AlexNet, and GoogleNet algorithms, based on transfer learning methods. The model framework and architecture were based on the computational and performance needs of the research. All the models were trained using the same dataset, i.e., the CGIAR wheat crop disease dataset, which had 1486 images in total split between train and test in three classes. The final VGG16 model achieved an accuracy of 99.67% and a validation accuracy of 77.52%. The average precision is 74%, average recall is 75%, and average f1-score is 74%. The accuracy achieved is similar to that of the reference paper (Sood et al.; 2022) on which this research is closely based.

The initial aim of this research was achieved as the final model was able to achieve good results, but it had some limitations too. Some limitations countered during this research were that the validation accuracy was not able to increase as per the training accuracy and the computational power and time needed to train some models were high even after decreasing the trainable parameters and reducing the model complexity. While the first limitation was seen in other previous research based on the same dataset, the computational time can be reduced in future studies by developing new less complex algorithms or a combination of algorithms for image classifications.

In the future, the results of this research can be implemented to create a mobile application that would help farmers check their crops in real-time, to check if their crops are prone to any disease or not. Doing this will ensure real-time monitoring of the crops, making the overall process faster for the farmers which will indeed some their money and time in crop production.

8 Acknowledgment

I want to offer my sincerest gratitude to my mentor and supervisor Dr.Christian Horn for helping me throughout this challenging work. I was able to achieve the results and present this research because of all the guidance and meaningful suggestions he was able to give me on all the aspects of the study. Because of his expertise in the data analytics field, I was able to grasp some good concepts which formed the basis of this research. I would like to thank him and everyone who provided guidance during the work which helped me achieve these results.

References

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. and Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey, *Heliyon* **4**(11): e00938.

Aboneh, T., Rorissa, A., Srinivasagan, R. and Gemechu, A. (2021). Computer vision framework for wheat disease identification and classification using jetson gpu infrastructure, *Technologies* **9**(3): 47.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods, *IEEE transactions on pattern analysis and machine intelligence* **34**(11): 2274–2282.

Arya, M., Anjali, K. and Unni, D. (2018). Detection of unhealthy plant leaves using image processing and genetic algorithm with arduino, *2018 International Conference on Power, Signals, Control and Computation (EPSCICON)*, IEEE, pp. 1–5.

Bhatt, P., Sarangi, S., Shivhare, A., Singh, D. and Pappula, S. (2019). Identification of diseases in corn leaves using convolutional neural networks and boosting., *ICPRAM*, pp. 894–899.

Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Ijaz, M. F. and Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases, *Sensors* **21**(14): 4749.

Farh, M. E.-A., Kim, Y.-J., Kim, Y.-J. and Yang, D.-C. (2018). Cylindrocarpon destructans/ilyonectria radicicola-species complex: Causative agent of ginseng root-rot disease and rusty symptoms, *Journal of ginseng research* **42**(1): 9–15.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis, *Computers and electronics in agriculture* **145**: 311–318.

Genaev, M., Ekaterina, S. and Afonnikov, D. (2020). Application of neural networks to image recognition of wheat rust diseases, *2020 Cognitive Sciences, Genomics and Bioinformatics (CSGB)*, IEEE, pp. 40–42.

Hassan, S. M., Maji, A. K., Jasiński, M., Leonowicz, Z. and Jasińska, E. (2021). Identification of plant-leaf diseases using cnn and transfer-learning approach, *Electronics* **10**(12): 1388.

Liu, J. and Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review, *Plant Methods* **17**(1): 1–18.

Neforawati, I., Herman, N. S. and Mohd, O. (2019). Precision agriculture classification using convolutional neural networks for paddy growth level, *Journal of Physics: Conference Series*, Vol. 1193, IOP Publishing, p. 012026.

Newlands, N. K. (2018). Model-based forecasting of agricultural crop disease risk at the regional scale, integrating airborne inoculum, environmental, and satellite-based monitoring data, *Frontiers in Environmental Science* **6**: 63.

Okafor, E., Schomaker, L. and Wiering, M. A. (2018). An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals, *Journal of Information and Telecommunication* **2**(4): 465–491.

Ramírez-Sánchez, G., Zaragoza-Bernabeu, J., Bañón, M. and Ortiz-Rojas, S. (2020). Bifixer and bicleaner: two open-source tools to clean your parallel data, *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, pp. 291–298.

Rezaei, R., Safa, L. and Ganjkhanloo, M. M. (2020). Understanding farmers' ecological conservation behavior regarding the use of integrated pest management-an application of the technology acceptance model, *Global Ecology and Conservation* **22**: e00941.

Sabrol, H. and Kumar, S. (2016). Fuzzy and neural network based tomato plant disease classification using natural outdoor images, *Indian J. Sci. Technol* **9**(44): 1–8.

Sharma, A. and Chetani, R. (2017). A review on the effect of organic and chemical fertilizers on plants, *Int. J. Res. Appl. Sci. Eng. Technol* **5**: 677–680.

Shrivastava, V. K., Pradhan, M. K., Minz, S. and Thakur, M. P. (2019). Rice plant disease classification using transfer learning of deep convolution neural network, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* **3**(6): 631–635.

Sood, S., Singh, H. and Jindal, S. (2022). Rust disease classification using deep learning based algorithm: The case of wheat.

Urva, P. (2021). Detection and classification of grain crops and legumes disease: A survey, *Sparklinglight Transactions on Artificial Intelligence and Quantum Computing (STAIQC)* **1**(1): 41–55.

Yang, J., Zhao, Z., Zhang, H. and Shi, Y. (2019). Data augmentation for x-ray prohibited item images using generative adversarial networks, *IEEE Access* **7**: 28894–28902.