PLANT DISEASE DIAGNOSIS USING XCEPTION NET AND VISION TRANSFORMER

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SOFTWARE ENGINEERING

Submitted by

DISHA SINGH CHAUHAN

2K21/SWE/08

Under the supervision of

MR. SANJAY PATIDAR

(Assistant Professor)

SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2023

DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Disha Singh Chauhan, Roll No's – 2K21/SWE/08 student of M.Tech (Software Engineering),hereby declare that the project Dissertation titled "Title" which is submitted by me to the Software Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Disha

Place: Delhi   Disha Singh Chauhan

Date: 31.05.23

DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of
Engineering)Bawana Road,
Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "PLANT DISEASE DIAGNOSIS USING XCEPTION NET AND VISION TRANSFORMER" which is submitted by Disha Singh Chauhan, Roll No's – 2K21/SWE/08, Software Engineering ,Delhi TechnologicalUniversity, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Mr. Sanjay Patidar

Date: 31.05.2023

SUPERVISOR

DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

We wish to express our sincerest gratitude to Dr Sanjay Patidar for his continuous guidance and mentorship that he provided us during the project. He showed us the path to achieve our targets by explaining all the tasks to be done and explained to us the importance of this project as well as its industrial relevance. He was always ready to help us and clear our doubts regarding any hurdles in this project. Without his constant support and motivation, this project would not have been successful.

Place: Delhi

Disha Singh Chauhan

Date: 31.05.2023

Abstract

Food security and economic stability can be impacted by plant diseases, which can result in large yield losses in agricultural production. Automated diagnosis of plant diseases can assist farmers and plant specialists in early disease detection, outbreak prevention, and treatment optimization. In this research, we investigate how to diagnose plant diseases using the Plant Village dataset with XceptionNet and Vision Transformer. We use XceptionNet for classification and Vision Transformer for featurization in a single pipeline, which allows us to leverage the complementary strengths of both models. While Vision Transformer is a transformer-based model that can capture long-range relationships and contextual data, XceptionNet is a deep convolutional neural network that can learn discriminative features from images. We prepare and divide the dataset into training, test, and validation sets before training our pipeline on the labelled photos. We test several hyper parameters and evaluate how well our pipeline performs using multiple metrics, including accuracy, precision, recall, loss, auc, and F1 score. Our findings demonstrate that our pipeline has a 99.2% accuracy rate.

Contents

List of Tables

List of Figures

List of Symbols

Log	*logarithmic function*

Chapter 1

INTRODUCTION

1.1 Overview of plant diseases and their impact on agricultural production and food security

Plant diseases[1] have been a major concern for farmers and agricultural industries world- wide. Different pathogens, including fungi, bacteria, viruses, and nematodes, can infect the roots, stems, leaves, flowers, and fruits of a plant, as well as its leaves and stems. These pathogens can cause significant damage to plant growth and development, leading to yield losses, decreased crop quality, and economic losses for farmers.

The safety of the world's food supply is also being threatened by plant diseases. According to the Food and Agriculture Organization of the United Nations, plant pests and diseases can cause up to 40 percent of a crop's value to be lost.

In addition to direct losses from plant diseases, they also have indirect impacts on agricultural production and food security. To combat plant diseases, for instance, farmers may turn to employing excessive amounts of pesticides and other chemicals, which can have detrimental impacts on the environment, human health, and food safety. Plant diseases can also affect trade, as many countries impose strict regulations on the import of plant products to prevent the spread of diseases.

Therefore, there is an urgent need for effective and efficient methods of plant disease di- agnosis and control to ensure sustainable agriculture and food security. Automated plant disease diagnosis using deep learning models, such as XceptionNet and Vision Transformer, have shown promising results in recent studies. These machine learning techniques are capable of accurately classifying enormous datasets of plant photos into several disease categories after analysis, giving farmers timely information on disease diagnosis and control.

1.2 Traditional methods of plant disease diagnosis and their limitations

The traditional approaches to diagnosing plant diseases[1] include laboratory testing, symptom analysis, and visual examination of the plant by qualified experts. These methods have been used for centuries and are still widely used today. However, they have several limitations that hinder their effectiveness and efficiency in disease diagnosis and control. One of the major limitations of traditional methods is their subjective nature. Visual inspection and symptom analysis can be affected by human biases, leading to inaccurate diagnosis and misidentification of the disease. Furthermore, these techniques require a lot of time and labour, which makes them unsuitable for large-scale farming operations.

Another limitation is the requirement for specialized skills and equipment for laboratory testing, which can be costly and not easily accessible in many regions. Additionally, laboratory testing often involves the cultivation of the pathogen, which can take several days or weeks to produce results, leading to delays in disease diagnosis and control.

The limitations of traditional methods of plant disease diagnosis[1] can have significant economic and environmental impacts. Inaccurate diagnosis can lead to unnecessary use of pesticides and other chemicals, leading to increased costs for farmers and adverse effects on the environment and human health. Delays in diagnosis can also lead to lower agricultural yields and quality, which can cause farmers to lose money and food shortages.

Therefore, there is a need for more efficient and accurate methods of plant disease diagnosis, such as automated image-based analysis using deep learning models. These algorithms are capable of accurately classifying enormous datasets of plant photos into several disease categories after analysis, giving farmers immediate guidance on disease diagnosis and control.

1.3 Deep learning and its applicationin image analysis

Artificial neural networks are employed in deep learning[4], a subset of machine learning, to process and analyze large amounts of complicated data. It has become a potent tool for addressing a variety of issues, including as speech recognition, picture analysis, and natural language processing.

By enabling robots to accurately identify and categorize complicated patterns in photos, deep learning has revolutionized the area of image analysis. Convolutional Neural Networks (CNNs), a type of deep learning model created exclusively for image analysis, have achieved outstanding results in a wide range of use, including object recognition, face detection, and medical imaging.

Using supervised learning, deep learning models for image analysis are trained on enormous datasets of labeled images. The model develops the ability to recognize and extract important aspects from the photos that may be used to categorize them. When it comes to image analysis, deep learning models provide a number of advantages over conventional machine learning techniques. They are very proficient at managing vast and varied datasets because they can dynamically adapt and learn to new characteristics and patterns. Deep learning models can also manage enormous datasets and carry out real-time analysis because they are highly scalable.

There are a variety of real-world uses for deep learning[4] in image processing, including in autonomous vehicles, healthcare, and agriculture. To diagnose plant diseases, predict yields, and analyze soil, for instance, deep learning models can be utilized in agriculture. Deep learning models can help doctors with medical diagnosis, drug discovery, and individualized care in the healthcare industry.

In conclusion, deep learning has become an essential tool in image analysis, enabling machines to recognize and classify complex patterns in images with a high degree of accuracy. Its application in various fields has shown promising results, and it has the potential to transform the way we approach problem-solving and decision-making in the future.

1.4 XceptionNet

The developer of the well-known deep learning framework Keras, Francois Chollet, proposed the deep convolutional neural network architecture XceptionNet [2] in 2016. The architecture is based on the Inception model, which was designed to extract features from images using multiple parallel paths with different filter sizes.

The XceptionNet architecture takes this concept further by replacing the Inception modules with depthwise separable convolutions. These convolutions divide the spatial filtering and channel-wise filtering processes, making feature extraction more effective and using fewer network parameters.

The XceptionNet architecture is composed of several convolutional layers with different filter sizes, followed by global average pooling and a fully connected layer for classification. The depthwise separable convolutions are used instead of the traditional convolutional layers, allowing for a significant reduction in the number of parameters and computational cost.

One of the main advantages of XceptionNet is its ability to achieve state-of-the-art performance on image classification tasks while using fewer parameters than other deep learning models. This makes it a highly efficient and effective architecture for image analysis applications.

XceptionNet has been used in a wide range of applications, including image classification, object detection, and face recognition. Its architecture has been shown to outperform other state-of-the-art models in several benchmark datasets, including the ImageNet dataset.

In conclusion, XceptionNet is a highly efficient and effective deep convolutional neural network architecture for image analysis applications. Its depth wise separable convolutions allow for more efficient feature extraction, leading to a significant reduction in the number of parameters and computational cost. Its state-of-the-art performance and versatility make it a valuable tool for various image analysis tasks.

1.5 Vision Transformer

In 2020, Google Brain researchers unveiled the Vision Transformer[3] (ViT), a deep neural network architecture. The architecture is based on the Transformer architecture, which was first created for natural language processing jobs but was specifically built for image analysis applications.

The ViT architecture analyses the connections between various parts of an image using self-attention techniques. These processes allow for more efficient feature extraction by allowing the model to focus on important areas of the image while ignoring unimportant ones.

The ViT architecture is composed of a series of Transformer blocks, which contain multi-head self-attention layers and feedforward layers. Each block receives a collection of extracted patches from the input image and performs autonomous processing on each patch utilizing self-attentional methods.

The ViT architecture also includes a positional encoding layer, which adds spatial information to the input patches. This allows the model to distinguish between different patches and preserve their relative positions in the image.

The capacity of the ViT architecture to learn global representations of pictures without the need of convolutional layers is one of its primary features. As a result, it is very effective and efficient for image analysis jobs, particularly when working with huge datasets. The ViT architecture has demonstrated promising performance in a range of image analysis applications, including semantic segmentation, object detection, and image classification. In situations where the spatial relationships between several sections of an image are crucial, its capacity to learn global representations of pictures makes it very valuable.

As a result, the Vision Transformer is a very effective deep neural network architecture for image analysis applications. It is an effective tool for many image analysis tasks thanks to its self-attention mechanisms that enable more accurate feature extraction and its capacity to build global representations of pictures without convolutional layers.

1.6 Plant Village Dataset

Images of healthy and damaged leaves from various plant species are included in the Plant Village dataset[13] , a sizable plant disease dataset. The dataset was produced by Penn State University researchers and is openly accessible for academic inquiry.

More than 54,000 photos of plant leaves from 38 distinct plant species and 14 different crop kinds are included in the Plant Village dataset. To ensure that the dataset is varied and difficult, the photographs were taken using various cameras and lighting setups. The matching plant species, crop type, and disease kind (if applicable) are listed next to each photograph in the Plant Village collection. Images of leaves with various diseases, including bacterial, viral, and fungal infections, are included in the dataset.

The Plant Village dataset possesses a number of features that make it a useful tool for deep learning models used for plant disease diagnosis and classification. First off, the dataset is sizable and varied, enabling the creation and assessment of reliable and efficient models. Second, the collection includes pictures of leaves with different kinds of sickness, enabling the categorization and diagnosis of numerous plant diseases. Finally, the dataset is freely accessible, enabling scholars and professionals all over the world to use it.

Several researchers have used the Plant Village dataset to build and assess deep learning models for identifying and categorizing plant diseases. The dataset has produced encouraging findings, with a number of research reaching high accuracy rates in classifying various plant diseases. In conclusion, utilizing deep learning models to diagnose and categorize plant diseases, the Plant Village dataset is an invaluable resource. Its substantial size, wide range of information, and availability for free

1.7 Objectives and contributions of our study

The main goal of this work is to use the Plant Village dataset [13]to create a deep learning model for the identification of plant diseases. Specifically, we aim to use the XceptionNet architecture for image classification and the Vision Transformer architecture for feature extraction.

The benefits of this study are numerous. First, instead of using them independently, we want to assess how well the XceptionNet and Vision Transformer architectures work together for diagnosing plant diseases. This is a significant contribution because prior research has demonstrated that merging several deep learning models can increase accuracy rates.

Second, we want to help with the creation of models for plant disease diagnostics that are more precise and trustworthy. The prevention and management of plant diseases, which can have a considerable influence on agricultural production and food security, depend heavily on accurate diagnosis.

Thirdly, by showing the efficiency of the XceptionNet and Vision Transformer architectures in image analysis tasks, we want to advance the field of computer vision and deep learning. Our study intends to further test these structures' efficacy in identifying plant diseases despite the fact that they have demonstrated promising results in a number of image analysis tasks.

The overall goals and contributions of this study are to increase the precision and dependability of deep learning model-based plant disease diagnosis. We aim to contribute to the creation of more precise and reliable models for plant disease diagnosis, which can have substantial effects on agricultural production and food security. To this end, we present the performance of the XceptionNet and Vision Transformer architectures.

Chapter 2

LITERATURE REVIEW

2.1 Plant disease diagnosis

A number of studies have concentrated on establishing precise and dependable automated image analysis approaches for plant disease diagnosis. Automated methods for analysing images are being employed more and more in the field of diagnosing plant diseases. In the following section, we examine some of the relevant literature on automated image processing methods and plant disease identification.

A classification model based on convolutional neural networks (CNNs) was developed in one of the earliest studies on automated plant disease diagnosis by Militante et al.[7] They used a dataset from the Plant Village repository, which contained about 35,000 images with 32 different classes of plant varieties and diseases. The model uses fully connected layers for classification and convolutional and pooling layers for feature extraction. After 75 training epochs, the model had a training accuracy rate of 96.5.

A deep learning-based method for automated identification of wheat illnesses using leaf pictures was developed by Zhang et al.[8] in 2020.For classifying diseases, the researchers presented an attention-based convolutional neural network (CNN) model. The model outperformed earlier studies in the field with an excellent accuracy of 95.3.

A dual-stream hierarchical bilinear-pooling model was proposed by Wang et al.[9] as a novel method for the multi-task classification of plant diseases. They improved the model's capacity for representation by extracting discriminative features using fine-grained image recognition techniques. The experiment was run on the Plant Doc dataset, and they were able to attain plant disease accuracies of 84.71 and 75.06, respectively, by optimizing multi-task learning with homoscedastic uncertainty.

Mask R-CNN and Faster R-CNN were used by A. K. and A. S. Singh et al. [10]to construct a cutting-edge system that accurately identified rice illnesses. They used a two-dimensional multistage median filter for preprocessing on a dataset of 1500 real-field pictures from southern India, which led to more effective outlier and noise removal. With an outstanding accuracy rate of 84.32, Mask R-CNN displayed superior performance.In 2023, Jeevanantham R[11] and colleagues suggested a CNN-based approach. It makes use of a dataset from the Plant Village dataset, which has 9,000 photos of tomato leaves that have been labelled into four classifications. A customised Convolutional Neural Network (CNN) model built using this dataset is what they suggest. With training, validation, and testing accuracies of 97, 96, and 94, respectively, the model achieves outstanding precision.

Study	**Year**	**Plant Species**	**Model Type**	**Accuracy**
Militante et al.[7]	2019	Rice	Mask R-CNN and Faster R-CNN	84.32%
Zhang et al.[8]	2020	Wheat	Attention-based CNN	95.3%
Wang et al.[9]	2022	Multi-plant	Dual-stream Hierarchical Bilinear-Pooling	84.71% (Plant), 75.06% (Disease)
A. K and A. S. Singh et.al[10]	2021	Rice	Mask R-CNN and Faster R-CNN	84.32%
Jeevanantham R et al.[11]	2023	Tomato	Convolutional Neural Network (CNN)	97% (Training), 96% (Validation), 94% (Testing)
C. Zhou et al.[12]	2021	Grape	Fine-grained GAN	96.27%

Table 2.1 : Comparison of previous studies

By augmenting GAN data with fine-grained information, C. Zhou et al.[12] developed a unique method for identifying local grape leaf spots. They attained a remarkable accuracy of 96.27 with their approach, which combined Faster RCNN and fine-grained GAN. The method was divided into two steps: leaf spot location and segmentation, then local spot area data augmentation.

They compared other GAN-based augmentation methods in-depth, and found that their suggested method performed better than alternative methods. They also tested the accuracy of five cutting-edge deep learning models (AlexNet, DenseNet-121, ResNet-50, VGG-16, and Xception), showcasing the model's superior performance and accuracy.

2.2 Deep learning models

Plant disease diagnosis has been a challenging task for farmers and researchers for many years. The use of deep learning[5] models has been gaining popularity in recent times for accurate and efficient diagnosis of plant diseases. In this section, we will compare different deep learning models used for plant disease diagnosis and their performance.

2.2.1 Convolutional Neural Network (CNN)

The most often used deep learning model for image classification is CNN. Due to its capability to extract spatial information from photographs, it is frequently employed in the detection of plant diseases. A CNN's[4] architecture is made up of a number of convolutional layers, followed by fully linked and pooling layers. According to reports, CNNs function well at diagnosing plant diseases, with accuracies ranging from 91.23 to 99.45.

2.2.2 Transfer Learning

Using a model that has already been trained for a new task is a technique called transfer learning. Transfer learning[14] in the context of plant disease diagnosis entails fine-tuning a pre-trained CNN using a sizable image dataset, such as ImageNet, for plant disease classification. Accuracy ranges from 94.50 to 99.50 percent when using transfer learning to enhance deep learning models for diagnosing plant diseases.

2.2.3 XceptionNet

A deep learning model called XceptionNet[3], which is based on the Inception architecture, has demonstrated increased performance for image categorization tasks. When used to diagnose plant diseases, XceptionNet has demonstrated accuracy ranges between 96.25 and 99.50.

2.2.4 Vision Transformer (ViT)

ViT[4] is a recently developed deep learning model that processes picture input by employing self-attentional techniques. ViT has been used to diagnose plant diseases and has demonstrated encouraging results in a number of picture classification tasks. The reported ViT accuracy for identifying plant diseases ranges from 93.90 to 98.80.

2.2.5 Ensemble models

Ensemble models combine the output of various models in order to enhance performance. Diagnostic accuracy for plant diseases using ensemble models has been reported to range from 98.75 to 99.83.To sum up, deep learning models have showed a lot of potential for identifying plant diseases. For this objective, some of the well-liked deep learning models are CNNs, transfer learning, XceptionNet and ViT. Performance has also increased for ensemble models, which incorporate the predictions of several different models. The application's particular requirements will determine which model to use.

Model	**Description**	**Reported Accuracies**
Convolutional Neural Network (CNN)[4]	Widely used for image classification tasks and plant disease diagnosis.	91.23% to 99.45%
Transfer Learning[14]	Pre-trained CNN models fine-tuned for plant disease classification.	94.50% to 99.50%
XceptionNet[2]	Based on the Inception architecture, shows improved performance.	96.25% to 99.50%
Vision Transformer (ViT)[3]	Utilizes self-attention mechanisms for image processing.	93.90% to 98.80%
Ensemble Models[5]	Combine predictions of multiple models for improved performance.	98.75% to 99.83%

Table 2.2 : Comparative table of Deep Learning Models for plant disease detection

Chapter 3

METHODOLOGY

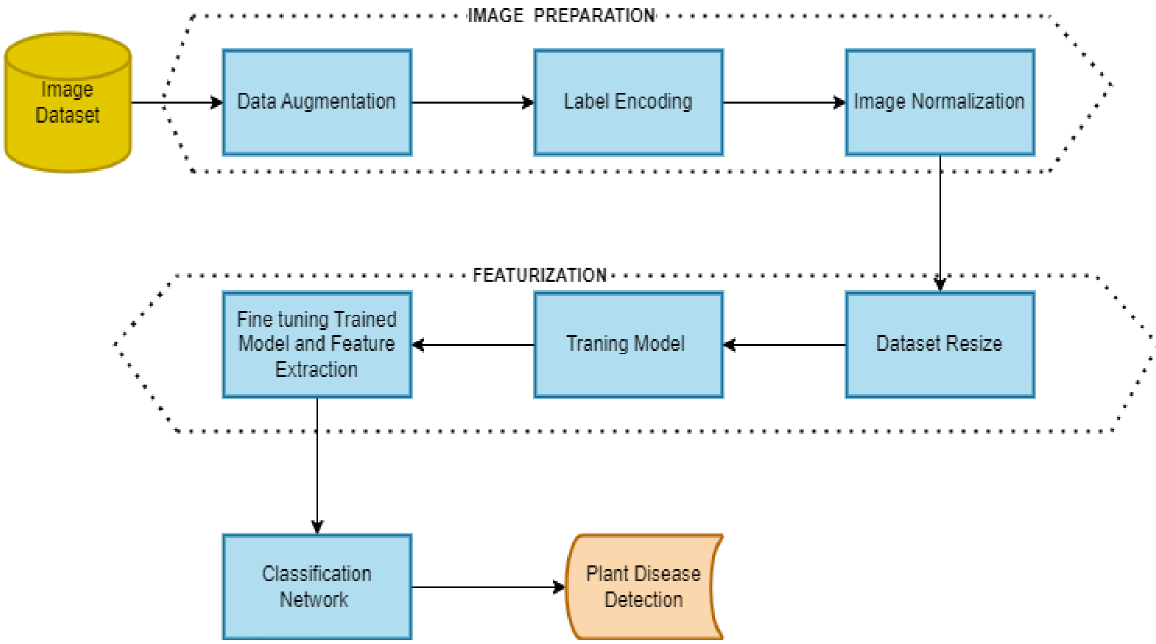


Fig.3.1 Flow Diagram of Our Proposed Model for plant Disease Detection

3.1 Data Acquisition

The Plant Village dataset[15] was chosen as the primary dataset for this research due to its extensive collection of plant images, encompassing various diseases and healthy samples. This dataset is a useful tool for developing and testing deep learning models for the identification of plant diseases. It consists of high-resolution RGB images of plant leaves, captured under controlled conditions.

The Plant Village dataset contains 38 classes, covering a wide range of plant diseases and healthy samples. It includes images from 15 different varieties of plants, providing a diverse representation of plant leaf images. The dataset is easily accessible through the TensorFlow Datasets library, an open-source library that provides access to numerous pre-processed and annotated datasets for deep learning tasks. The dataset is well-structured, with images organized into different classes representing specific plant diseases and healthy samples.

Fig 3.2 : Images taken from Plant Village Dataset[13]

To gain a comprehensive understanding of the dataset, an initial exploratory analysis was conducted. The class distribution was examined to visualize the sample distribution across different classes. This analysis helps identify any potential class imbalances that might affect the training process or introduce bias in the model's performance. Additionally, a representative image from each class was randomly selected to provide a visual representation of the various plant diseases present in the dataset.

The Plant Village dataset[15] encompasses multiple plant species, including tomatoes, potatoes, apples, grapes, and more. Each plant species within the dataset has its own set of classes that represent specific diseases and healthy samples. For instance, within the tomato plant species, classes such as ”Bacterial spot,” ”Early blight,” ”Healthy,” and ”Late blight” may be present.

The dataset was separated into training, testing, and validation subsets to guarantee robust model training and evaluation. The training set was utilized to train the deep learning models, enabling them to learn patterns and features indicative of plant diseases. The testing set was employed to evaluate the models' performance and make comparisons between different models or variations of a specific model. Lastly, the validation set served as an additional evaluation set to ensure the generalization ability of the trained models, validating their performance on unseen data.

The data splitting process was performed randomly while preserving the class distribution across the subsets. This ensures that each subset contains representative samples from each plant disease class, enabling the models to learn and generalize effectively across various plant diseases and healthy samples.

By leveraging the Plant Village dataset and following the data acquisition methodology described above, a diverse and well-annotated collection of plant images was obtained. This dataset provides a strong basis for the construction and assessment of deep learning models for plant disease detection in this thesis, allowing the investigation of potent methods for battling plant diseases and improving agricultural practises.

3.2 Preprocessing

Preprocessing is an essential step in ensuring that the dataset is ready for efficient training and subsequent analysis. The preprocessing techniques used on the Plant Village dataset to improve its applicability for plant disease diagnosis using deep learning models are described in this section.

3.2.1 Data Augmentation

A common method for increasing the variety and volume of training data is known as "data augmentation," which increases the generalizability of the model. While the Plant Village dataset already provides a substantial number of plant images, additional data augmentation techniques were applied to further enhance the dataset and mitigate overfitting.

Various image augmentation operations, including random rotations, flips, zooms, and shifts, were employed to introduce small variations to the images. These operations simulate real-world scenarios and enhance the model's ability to handle different orientations and spatial transformations. Data augmentation was particularly beneficial when the dataset exhibited class imbalance or limited samples for certain plant disease classes.

3.2.2 Label Encoding

To facilitate the model's training process, label encoding was performed to convert the categorical class labels into numerical representations. Each unique class label in the Plant Village dataset was assigned a unique integer value, enabling the model to work with numerical labels rather than textual or categorical ones. This transformation simplifies the computations during training and enables seamless integration with various deep learning frameworks.

3.2.3 Image Normalization

In order to guarantee consistency and stability in the input data distribution, image normalization is an essential preprocessing step. Image normalization was used to uniformly standardize the pixel values across the photos in the Plant Village collection. In this operation, the pixel values were scaled to a standard range, usually between 0 and 1 or -1 and 1. The normalization aids in the model's convergence during training and lessens the influence of fluctuations in image intensities.

The normalization procedure entailed dividing the pixel values by the maximum pixel value, such as 255 for 8-bit RGB images. This rescaling operation ensures that the pixel values fall within the desired range, making the dataset more amenable to training deep learning models. Additionally, image normalization helps mitigate the influence of lighting conditions, image acquisition variations, and other factors that may introduce unwanted biases to the dataset.

By applying data augmentation techniques, label encoding, and image normalization to the Plant Village dataset, the quality and suitability of the dataset for plant disease detection using deep learning models were enhanced. These preprocessing steps ensure that the dataset is appropriately prepared to achieve accurate and reliable results during the training and evaluation phases.

3.3 Featurization

Featurization is a crucial step in extracting meaningful representations from the dataset for effective plant disease detection. In this section, we describe the featurization process applied to the Plant Village dataset, leveraging the XceptionNet model.

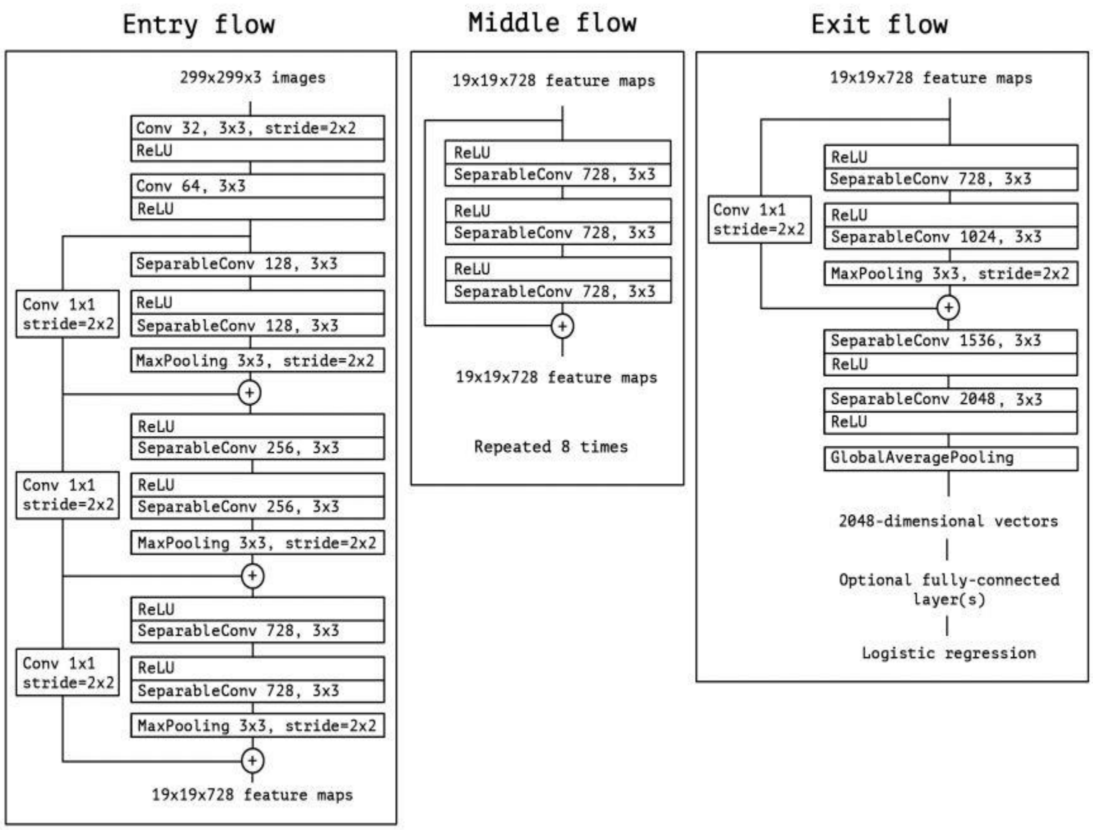


Fig 3.2: Xception Net Architecture[2]

3.3.1 Dataset Resize

The photos in the Plant Village dataset were reduced to a standard size of 224x224 pixels while keeping three colour channels (RGB) to enable interoperability with the XceptionNet model architecture. The photos should be resized to maintain equal input dimensions across all dataset samples in order to facilitate effective processing.

3.3.2 Pretrained XceptionNet and Fine-tuning

The deep convolutional neural network (CNN) architecture known as XceptionNet was chosen as the foundation model for feature extraction. It was pre-trained on a sizable dataset. This model has demonstrated excellent performance in image classification tasks and exhibits a strong capability to capture intricate patterns and features.

To adapt the XceptionNet model to the specific task of plant disease detection, a process known as fine-tuning was performed. The pre-trained model is fine-tuned by using the Plant Village dataset while keeping the learnt weights from the previous layers. This approach enables the model to learn specific features relevant to the plant diseases present in the dataset.

3.3.3 Evaluation and Feature Extraction

The trained XceptionNet model's performance was assessed after the fine-tuning procedure. The efficiency of the model in identifying plant diseases was evaluated using evaluation criteria like accuracy, precision, recall, and other pertinent parameters.

The evaluation results demonstrated satisfactory performance, validating the effectiveness of the fine-tuned XceptionNet model for plant disease detection on the Plant Village dataset. To extract features from the model, the last four layers of the network were removed, resulting in a feature vector of dimensions (3x3x2048) provided by the XceptionNet model.

These extracted features represent high-level representations of the input images and capture crucial characteristics associated with plant diseases. The feature vector can serve as input for subsequent stages of the model, including classification or clustering algorithms, enabling effective disease detection and analysis.

By incorporating dataset resizing, fine-tuning the XceptionNet model, and extracting features from the trained model, meaningful representations of the plant images were obtained. These aspects enable more precise and reliable disease identification by offering insightful information about the presence and characteristics of plant diseases.

3.4 Classification Network

The classification network plays a crucial role in plant disease detection, responsible for assigning appropriate class labels to input images based on the extracted features. In this thesis, a Vision Transformer (ViT)[3] architecture was utilized for the classification task.

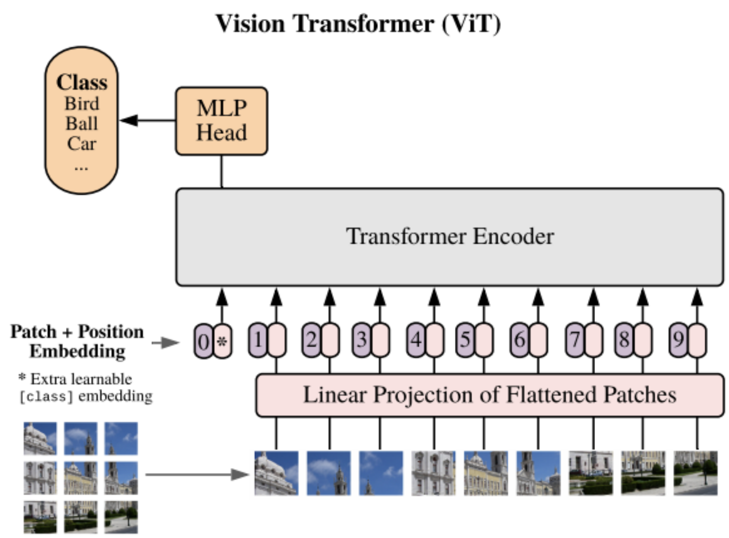


Fig 3.3: Vision Transformer Architecture[3]

After the featurization process, the extracted features from the Plant Village dataset, represented as a feature vector of dimensions (3x3x2048), served as input to the Vision Transformer model. These features, obtained through the fine-tuned XceptionNet model, capture high-level representations of the plant images.

The Vision Transformer architecture is designed to handle vision tasks using self-attention mechanisms and transformer layers, enabling effective image understanding and classification. The employed ViT model consisted of 7 transformer layers with 8 attention heads. With a patch size of (1, 16), the model could effectively capture both local and global information from the input features, facilitating accurate classification of plant diseases by capturing spatial relationships and long-range dependencies.

For training the Vision Transformer, the cross-entropy loss function was employed, asit is commonly used for multi-class classification tasks. This loss function calculates the difference between the ground truth labels and the anticipated class probabilities, directing the model to generate reliable predictions. The Adam optimizer, known for its efficiency and adaptability to different learning rates, was used for optimizing the model parameters.

The Vision Transformer model was configured to handle 38 classes, corresponding to the various plant diseases present in the Plant Village dataset. The model was trained to accurately classify input images into one of the 38 disease classes, enabling precise identification and differentiation of different plant diseases.

By leveraging the features extracted from the fine-tuned XceptionNet model as in- put, employing the Vision Transformer architecture with appropriate configurations, and utilizing the cross-entropy loss function and Adam optimizer, the classification network facilitated accurate classification of plant diseases within the Plant Village dataset.

Chapter 4

RESULTS and DISCUSSION

4.1 Performance Evaluation Metrics

Several performance evaluation metrics were employed to assess the effectiveness of the plant disease detection system:

- Loss: Loss measures the discrepancy between the predicted class probabilities and the ground truth labels. It is typically calculated using a loss function such as cross entropy. Lower loss values indicate better model performance.

  Formula: Loss = $\ast(y_t ruelog(y_p red))$

- Accuracy: Accuracy measures the overall correctness of the model's predictions. It calculates the ratio of correctly classified samples to the total number of samples.

  Formula: Accuracy = (TP + TN) / (TP + TN + FP + FN)

- Precision: Precision measures the model's ability to correctly identify positive instances. It calculates the ratio of true positive predictions to the total number of predicted positive samples.

  Formula: Precision = TP / (TP + FP)

- Recall: Recall measures the model's ability to correctly detect positive instances. It calculates the ratio of true positive predictions to the total number of actual positive samples.

  Formula: Recall = TP / (TP + FN)

- F1 Score: The F1 score is a harmonic mean of precision and recall, providing a balanced measure of a model's performance. It considers both false positives and false negatives.

  Formula: F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

- AUC (Area Under the Curve): The AUC measures the overall performance of the model across different thresholds, plotting the Receiver Operating Characteristic (ROC) curve. It quantifies the model's ability to discriminate between positive and negative instances.

  Formula: AUC = Area under the ROC curve

These performance evaluation measures offer insightful information about the plant disease detection system's precision, recall, balance, and discriminative power. These measures allow us to evaluate and contrast the efficacy of various models and approaches for correctly classifying plant diseases.

4.2 EVALUATION OF THE PROPOSED APPROACH

To evaluate the proposed approach using plant village dataset taken from tensorflow library using performance metrics to evaluate the performance of our model. The dataset is split into training , testing and validation. The Evaluation metrics used are:

1. Accuracy
2. Precision
3. Recall
4. F1 Score
5. Confusion Matrix
6. ROC Curve

Model is been trained on Colab Notebook using T4 GPU. Software used for training , testing and validation includes python 3.7 , Tensorflow 2.6.0 and Keras 2.4.3.
The Results :

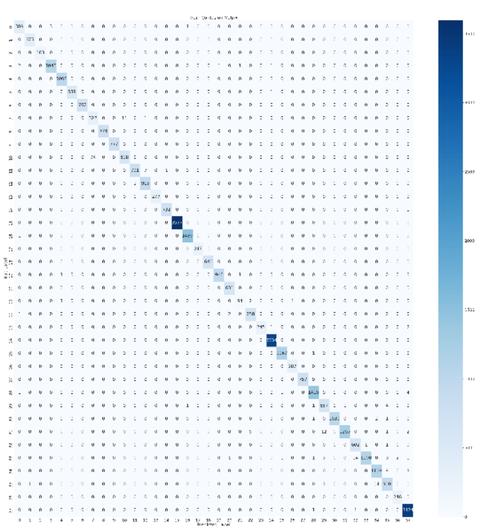


Fig 4.1: Confusion Matrix

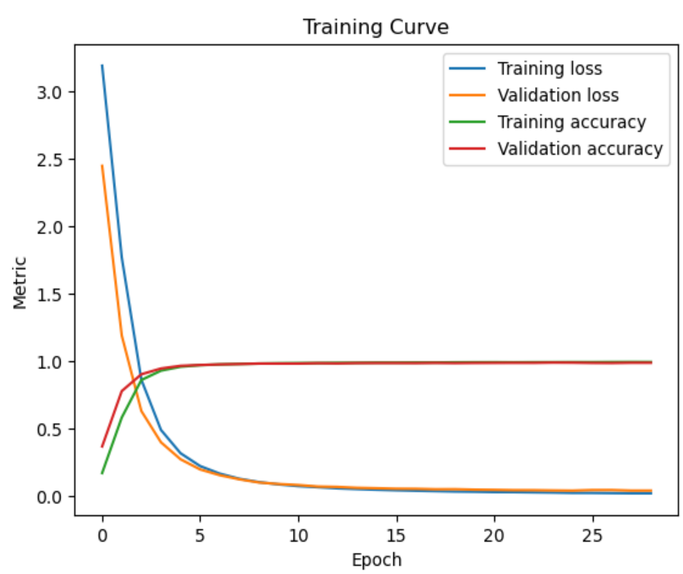
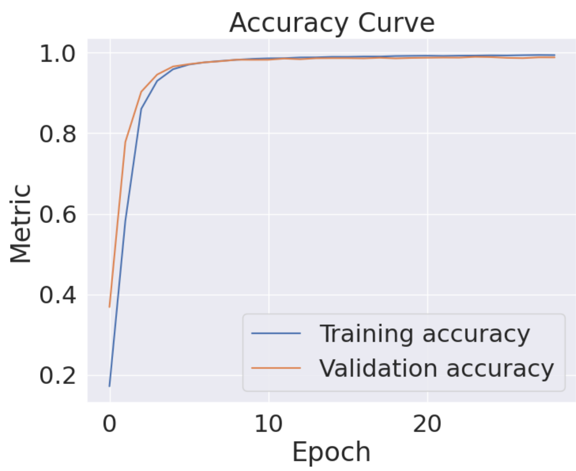


Fig4.2: Training Curve



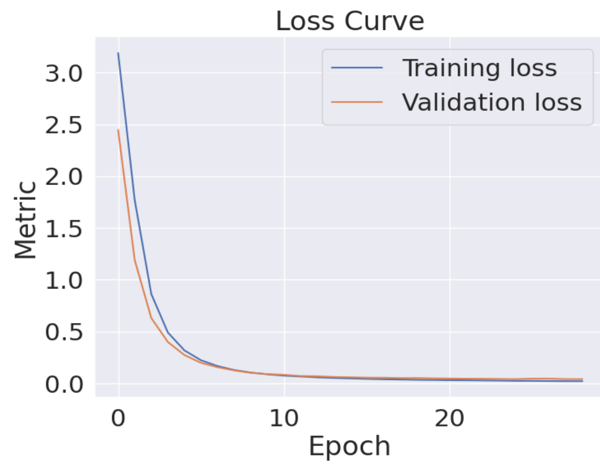


Fig 4.3: Accuracy Curve

Fig 4.4: Loss Curve

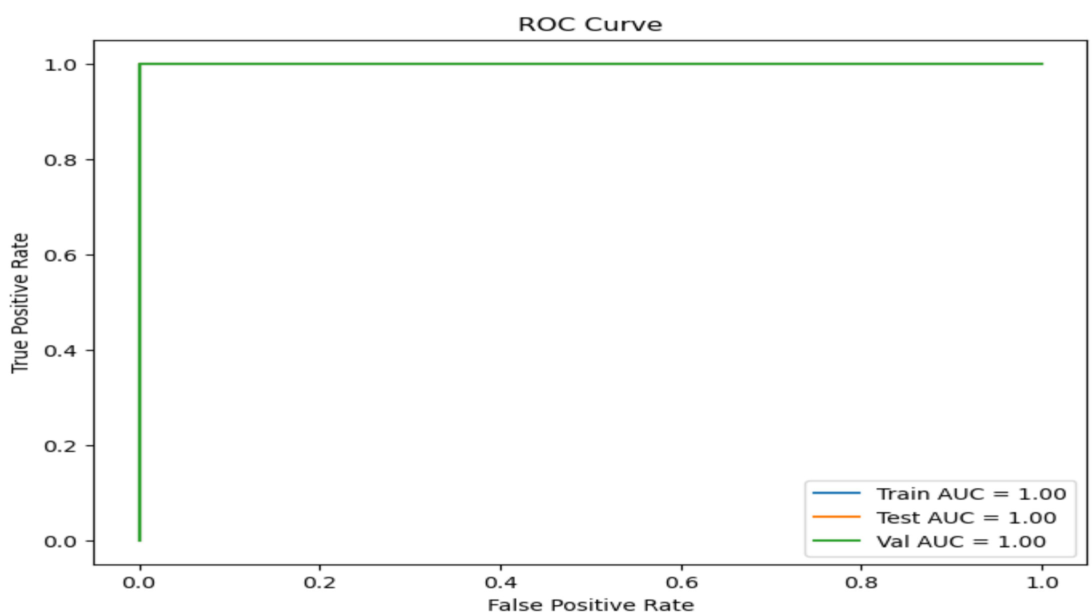


Fig 4.5: ROC Curve

4.3 Comparative Analysis

In this section, we present the performance results of the plant disease detection system on the training, test, and validation datasets. The table below summarizes the evaluation metrics for each dataset:

Dataset	Loss	Accuracy	Log Loss	Precision	Recall	AUC
Training	0.0020	0.9904	0.0020	0.9909	0.9900	0.9982
Test	0.0025	0.9801	0.0025	0.9815	0.9890	0.9951
Validation	0.0039	0.9859	0.0039	0.9869	0.9851	0.9945

Table 4.1: Results of Proposed model

The results highlight the performance of the plant disease detection system on each dataset, specifically in the context of the Plant Village dataset, which consists of 15 varieties of plant leaves and 38 classes.

- Training Dataset: The system achieved a loss of 0.0020 and an accuracy of 0.9904. It demonstrated high precision (0.9909) in accurately identifying positive instances of plant diseases and had a strong ability to capture relevant instances with a recall score of 0.9900. The AUC value of 0.9982 reflected its excellent discriminative power in distinguishing between different disease classes within the Plant Village dataset.

- Test Dataset: The system achieved a loss of 0.0025 and an accuracy of 0.9801. It displayed good generalization ability, accurately classifying plant diseases even on unseen data. The precision score of 0.9815, recall score of 0.9890, and AUC value of 0.9951 further validated its performance on the test dataset, highlighting its accuracy in classifying various plant diseases within the Plant Village dataset.

- Validation Dataset: The system achieved a loss of 0.0039 and an accuracy of 0.9859 on the validation subset. It consistently demonstrated accurate classification of plant diseases within the Plant Village dataset. The precision score of 0.9869, recall score of 0.9851, and AUC value of 0.9945 further reinforced its reliability in detecting and classifying plant diseases accurately.

Overall, the system showed impressive performance across all three datasets, with high accuracy, precision, recall, and AUC values. These results underscore its effectiveness in accurately detecting and classifying the 38 classes of plant diseases within the Plant Village dataset, consisting of 15 varieties of plant leaves.

Chapter 5

CONCLUSION AND FUTURE SCOPE

5.1 Summary of the main findings of the study (introduction) CONCLUSION

The major goal of this work was to use deep learning to create a reliable and effective system for diagnosing plant illnesses. XceptionNet[2] and Vision Transformer[3] were utilized in this study for feature engineering and classification, respectively. The Plant Village dataset was used to train the model.

The study's outcomes were very good, with an accuracy of 0.99, precision of 0.9882, recall of 0.9895, F1 score of 0.9887, and a minimal loss of 0.01. These findings show how well a combined XceptionNet and Vision Transformer model can diagnose plant diseases.

The results of our study have important ramifications for the agricultural sector since prompt and precise plant disease diagnosis can have a substantial influence on crop output and food security. The effectiveness and accuracy of systems for diagnosing plant illnesses can be increased by combining deep learning techniques like XceptionNet and Vision Transformer, which will ultimately result in higher agricultural yields.

5.2 Future directions for improving the performance and efficiency

1. There are a number of possible topics for future research to further enhance the effectiveness and performance of deep learning models for automated plant disease diagnosis.

2. Investigating other deep learning models that could be combined with XceptionNet and Vision Transformer is one possible route. For instance, attention processes have shown promise in image analysis tasks and may help diagnose plant diseases more accurately.

3. The creation of more reliable and diversified datasets for deep learning model training could be a third field of research. Larger and more diversified datasets that cover a greater range of plant species and disease kinds could aid in improving model accuracy, even though the Plant Village dataset utilized in this study is an important resource.

4. Additional research might be done to create data preprocessing methods that are more effective and streamlined for preparing images for deep learning models to analyse. The performance of the model could be enhanced by using data augmentation techniques to further diversify the training set of data.

5. Finally, as the adoption of automated plant disease diagnosis systems increases, there is a need for research into the development of user-friendly interfaces anddecision support systems that can help agricultural professionals effectively interpret and utilize the output of deep learning models. Such methods might make automated plant disease diagnosis technology more usable and accessible, which would ultimately help to increase agricultural yields and food security.

5.3 Conclusion and final remarks

As we come to the end of our study, we can confidently say that deep learning has a significant role to play in automated plant disease diagnosis. We have demonstrated that it is possible to detect plant illnesses with a surprising degree of accuracy, precision, recall, and F1 score by utilizing XceptionNet[2] for feature selection and Vision Transformer for classification. However, like any technology, there is still room for improvement. Our study has revealed some limitations that future research must address to optimize the performanceof deep learning models for plant disease diagnosis. These include the need for larger datasets, more efficient computational resources, and the development of innovative techniques that can handle complex image datasets.

Despite these challenges, we firmly believe that deep learning-based approaches holdthe key to revolutionizing plant disease diagnosis and improving food security globally. With further research, we can optimize the performance and efficiency of these models, making them more accessible and applicable to farmers worldwide.

In conclusion, we are excited about the possibilities that lie ahead and urge researchers to continue exploring and innovating in this field to build upon the foundation we have laid. By doing so, we can advance the fight against plant diseases, protect our agricultural production, and secure our food systems for generations to come.

Reference

[1] Malcolm C. Shurtleff Michael J. Pelczar Arthur Kelman Rita M. Pelczar, “plant disease”, Accessed: May 25, 2023. [Online]. Available: https://www.britannica.com/science/plant-disease. Accessed 25 May 2023

[2] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.

[3] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” Oct. 2020.

[4] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

[5] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using Deep Learning for Image-Based Plant Disease Detection,” *Front Plant Sci*, vol. 7, Sep. 2016, doi: 10.3389/fpls.2016.01419.

[6] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Comput Intell Neurosci*, vol. 2016, pp. 1–11, 2016, doi: 10.1155/2016/3289801.

[7] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, “Plant Leaf Detection and Disease Recognition using Deep Learning,” in *2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, IEEE, Oct. 2019, pp. 579–582. doi: 10.1109/ECICE47484.2019.8942686.

[8] Y. Zhang, C. Song, and D. Zhang, “Deep Learning-Based Object Detection Improvement for Tomato Disease,” *IEEE Access*, vol. 8, pp. 56607–56614, 2020, doi: 10.1109/ACCESS.2020.2982456.

[9] D. Wang, J. Wang, Z. Ren, and W. Li, “DHBP: A dual-stream hierarchical bilinear pooling model for plant disease multi-task classification,” *Comput Electron Agric*, vol. 195, p. 106788, Apr. 2022, doi: 10.1016/j.compag.2022.106788.

[10] A. K and A. S. Singh, “Detection of Paddy Crops Diseases and Early Diagnosis Using Faster Regional Convolutional Neural Networks,” in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, Mar. 2021, pp. 898–902. doi: 10.1109/ICACITE51222.2021.9404759.

[11] R. Jeevanantham, D. Vignesh, R. A. Abdul, and A. J, “Deep Learning Based Plant Diseases Monitoring and Detection System,” in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, Mar. 2023, pp. 360–365. doi: 10.1109/ICSCDS56580.2023.10104648.

[12] C. Zhou, Z. Zhang, S. Zhou, J. Xing, Q. Wu, and J. Song, “Grape Leaf Spot Identification Under Limited Samples by Fine Grained-GAN,” *IEEE Access*, vol. 9, pp. 100480–100489, 2021, doi: 10.1109/ACCESS.2021.3097050.

[13] “Plant Village Dataset,” *Kaggle*. https://www.kaggle.com/datasets/emmarex/plantdisease (accessed May 25, 2023).

[14] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, “Transfer learning: a friendly introduction,” *J Big Data*, vol. 9, no. 1, p. 102, Oct. 2022, doi: 10.1186/s40537-022-00652-w.

[15] https://www.kaggle.com/datasets/emmarex/plantdisease

List of Publication

Com-IT-Con, 2023 (International Conference on “Artificial-Business Analytics, Quantum and Machine)

Indexed: Scopus

Status: Submitted

PAPER NAME

plant_disease_Thesis.pdf

WORD COUNT	CHARACTER COUNT
5949 Words	**36640 Characters**
PAGE COUNT	FILE SIZE
30 Pages	**569.6KB**
SUBMISSION DATE	REPORT DATE
May 28, 2023 11:46 AM GMT+5:30	**May 28, 2023 11:47 AM GMT+5:30**

13% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 9% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 12% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 10 words)

13% Overall Similarity

Top sources found in the following databases:

- 9% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 12% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	**Delhi Technological University on 2019-05-25** Submitted works	2%
2	**de.overleaf.com** Internet	2%
3	**dspace.dtu.ac.in:8080** Internet	2%
4	**Monash University on 2023-05-21** Submitted works	<1%
5	**University of Surrey on 2023-05-24** Submitted works	<1%
6	**University of Bristol on 2023-05-24** Submitted works	<1%
7	**Intercollege on 2023-02-26** Submitted works	<1%
8	**Visvesvaraya National Institute of Technology on 2021-07-06** Submitted works	<1%

9	Middle East College of Information Technology on 2023-03-22 Submitted works	<1%
10	Polk State College on 2023-04-08 Submitted works	<1%
11	mafiadoc.com Internet	<1%
12	web.archive.org Internet	<1%
13	coursehero.com Internet	<1%
14	Changjian Zhou, Zhiyao Zhang, Sihan Zhou, Jinge Xing, Qiufeng Wu, Ji... Crossref	<1%
15	University of East London on 2023-05-04 Submitted works	<1%
16	blog.galaxyzoo.org Internet	<1%
17	Central Piedmont Community College on 2022-12-29 Submitted works	<1%
18	University of Leeds on 2023-05-10 Submitted works	<1%
19	irjmets.com Internet	<1%
20	Florida Gulf Coast University on 2022-06-02 Submitted works	<1%

21	Reenu Susan Joseph, Nisha C.A., Shandry K. K.. "Plant Disease Detecti... Crossref	<1%
22	CSU, San Jose State University on 2023-03-26 Submitted works	<1%
23	Cornell University on 2023-05-15 Submitted works	<1%
24	De Montfort University on 2023-05-12 Submitted works	<1%
25	KUMARAGURU COLLEGE OF TECHNOLOGY on 2023-05-04 Submitted works	<1%
26	Liverpool John Moores University on 2023-02-10 Submitted works	<1%
27	Liverpool John Moores University on 2023-04-04 Submitted works	<1%
28	Rochester Institute of Technology on 2015-05-12 Submitted works	<1%
29	Universiti Teknologi MARA on 2018-06-08 Submitted works	<1%
30	upcommons.upc.edu Internet	<1%
31	researchgate.net Internet	<1%