AI-Based Forest Health Monitoring and Its Regulatory Compliance Challenges in Cyber Security

By

Vijay Kumar

(00000361559)

Department of Cyber Security

Pakistan Navy Engineering College (PNEC)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2023)

AI-Based Forest Health Monitoring and Its Regulatory Compliance Challenges in Cyber Security

By

Vijay Kumar

(Registration No: 00000361559)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in

Cyber Security

Supervisor: Cdre. Dr. Nadeem Kureshi

Pakistan Navy Engineering College (PNEC)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2023)

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at Department of Cyber Security at Pakistan Navy Engineering College (PNEC) or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at Pakistan Navy Engineering College (PNEC) or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Name of Student: Vijay Kumar

Date: 26-JAN-2024

FORM TH-4

National University of Sciences and Technology

MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by Vijay Kumar (00000361559) Titled: AI-Based Forest Health Monitoring and it's Regulatory Compliance Challenges in Cyber Security be accepted in partial fulfillment of the requirements for the award of Master's degree.

EXAMINATION COMMITTEE MEMBERS

1. Name: Lt. Cdr Dr. Gul Shehzad PN   Signature: ____________

2. Name: Dr. Muhammad Usama   Signature: ____________

Supervisor's name: Cdre Dr Nadeem Kureshi   Signature: ____________ DR NADEEM KUR Commodore DEAN MIS PNS JAUHAR

Date: 26-JAN-2024

AALIYA ALI
Lt Cdr Pakistan Navy
HOD CySD

Head of Department

26-JAN-2024
Date

COUNTERSIGNED

Date: 26-JAN-2024

DR NADEEM KURESHI
Commodore
DEAN MIS
PNS JAUHAR

Dean / Principal

CERTIFICATE FOR PLAGIARISM

1. It is certified that PhD / M.Phil / **MS** Thesis Titled "AI-Based Forest Health Monitoring and it's Regulatory Compliance Challenges in Cyber Security." by **Vijay Kumar (2021-NUST-MS Cyber Security (CyS Fall 21)** has been examined by us. We undertake the follows:

a. Thesis has significant new work / knowledge as compared already published or is under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.

c. There is no fabrication of data or results which have been compiled / analyzed.

d. There is no falsification by manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC Plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

DR NADEEM KURESHI
Commodore
DEAN MIS
PNS JAUHAR

Thesis Acceptance Certificate

Certified that final copy of MS/MPhil thesis entitled "AI-Based Forest Health Monitoring and it's Regulatory Compliance Challenges in Cyber Security"written by Vijay Kumar, (Registration No 00000361559), of Pakistan Navy Engineering College (PNEC) has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ____________________

Name of Advisor: **Cdre. Dr. Nadeem Kureshi**

Date: 26-JAN-2024

AALIYA ALI
Lt Cdr Pakistan Navy
HOD CySD

Signature (HOD): ____________________

Date: 26-JAN-2024

DR NADEEM KURESHI
Commodore
DEAN MIS
PNS JAUHAR

Signature (Dean): ____________________

Date: 26-JAN-2024

CERTIFICATE OF APPROVAL

It is certified that the contents and form of the thesis entitled "AI-Based Forest Health Monitoring and it's Regulatory Compliance Challenges in Cyber Security." submitted by Shahid have been found satisfactory for the requirement of the degree.

Advisor: **Cdre. Dr. Nadeem Kureshi**

Signature: ____________________

Date: 26-JAN-2024

NADEEM KURESHI
Commodore
PN MIS
PNS JAUHAR

Committee Member 1: **Asst. Prof. Dr. M Usama**

Signature: ____________________

Date: 26-JAN-2024

Committee Member 2: **Lt. Cdr Dr. Gul Shehzad PN**

Signature: ____________________

Date: 26-JAN-2024

ACKNOWLEDGMENT

All honor and praise belong to Allah, I am thankful and admire Him for His greatness, righteousness, splendor, and favors which have made it possible to finish this thesis. I am truly grateful to Allah for His unwavering presence, direction, and support during the difficult times, which inspired and committed me to finish this task.

I am grateful for Allah's help, which has come to me repeatedly in many forms. I am sincerely thankful to Cdr. Dr. Nadeem Kureshi, my supervisor, Dr. Muhammad Usama, and Lt. Cdr. Dr. Gul Shehzad PN, my co-supervisor, for their encouragement, support, and direction, which were essential to the completion of this thesis. It has been a source of honor for us to work with such distinguished mentors, and we pray that Allah would continue to bless them abundantly and provide them success in life.

In the future, we hope to work with similarly powerful and encouraging people in our endeavors, which will help us advance our knowledge and development.

Contents

List of Figures

Abstract

Agriculture plays an important role in Pakistan due to its food demand and population growth. The problem can be tackled with robust machine learning and deep learning techniques which can reduce the monitoring time can help in data-driven decision making.

We commenced the training of models over the dataset of plant village dataset having 20642 images, with 15 different classes of 3 species, including pepper, potato, and tomato we used ResNet50 with a span of 10 epochs, and we were able to obtain the utmost training accuracy of 98.09%. Nevertheless, the pinnacle of the validation (testing) accuracy culminated at 91.5%.

we used the Fine-Tuning method with the ResNet50 pre-trained model in Keras, we selectively designated the last 75 layers of the ResNet50 model as trainable, enabling it to adapt and identify patterns in the current dataset, The model demonstrated notable improvements during this fine-tuning process, with training accuracy of 98.09 percent and testing accuracy of 95.23 percent, These results outperform the original ResNet50 model, demonstrating how the fine-tuning approach may be used to enhance the model's performance for the given job.

Additionally, in this study, two distinct instructional methodologies were examined for the Vision Transformer (VIT). The model underwent rigorous training over the course of 10 epochs However, despite these endeavors, the testing accuracy fell short of expectations, remaining at approximately 60%, Subsequently, an alternative approach was implemented whereby a pre-trained Vision Transformer (VIT) was utilized Through this utilization of pre-trained models, the aforementioned methodology exhibited a marked improvement in performance, ultimately attaining an accuracy level of 98%. The study also focuses on the regulatory compliance challenges for data.

1. Chapter 1: INTRODUCTION

1.1 Overview

In today's rapidly advancing world, where Agriculture stands as a cornerstone of our economy, it is crucial to address the challenges that hinder its growth and productivity. In today's modern world, nobody is interested in farming due to the many issues that farmers face. The Issue of keeping plants safe from diseases is connected to practical adaptation to climate change [1]. Research shows how climate change can impact the different stages and rates at which plant diseases develop, also the ability of the plants to resist these diseases may change [2]. There are methods for automating the detection of plant diseases using artificial intelligence, deep learning, and machine learning [3], these methods can effectively identify and detect plant diseases without any human interventions. Computer Vision has proven the best results over the course of time which is being used for feature detection, extraction, and enhancement for classification and identification of different applications including Vehicle detection, medical diagnosis, robotics, and processing of agricultural items [4]-[7], for feature extraction, segmentation, and classification the image processing was used before deep learning. Deep learning is state of art Machine learning technology, the more advanced methods of deep learning such as transfer learning have proven their worth in terms of performance for identification and classification-related tasks [8]. Let's now understand the few more concepts.

1.2 Convolutional neural networks

Convolutional neural networks belong to a class of neural networks specifically crafted for image recognition and classification, yielding outstanding outcomes. Initially, CNN has found widespread application in addressing object recognition challenges. However, its utilization has expanded into novel domains, including text recognition, object tracking, action detection, visual saliency detection, and scene labeling. Presently, CNN is extensively employed for the detection of diverse plant leaf diseases. Grasping the advancements in CNN architecture relies on a thorough understanding of the diverse components of CNN and their respective applications. Figure 2 illustrates multiple elements of CNN [20]. For the input images pixels are the foundational components of a computer image, representing visual data in binary form. Arranged in a matrix-like layout within the digital image, these pixels are sequentially organized from 0 to 255. The pixel value uniquely determines the brightness and hue of each pixel, In the training process of CNN layers, the emphasis is on recognizing fundamental patterns like lines and curves initially, followed by more sophisticated patterns such as faces and objects. This progression suggests that the adoption of CNNs could empower computers with visual perception [21-22].

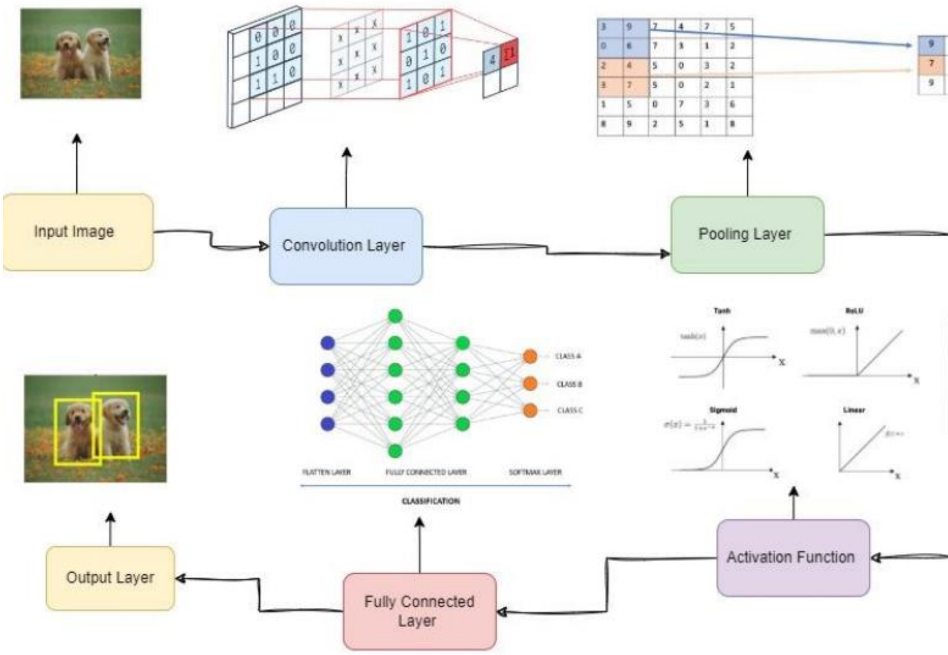

Figure 1: Components of CNN

1.3 The Convolutional Neural Network Generalization

When aiming to establish robust generalizations for CNN models, the central hurdle is addressing over-fitting. Over-fitting arises when a model excels on training data but falters when applied to test data, representing data it hasn't encountered previously, as detailed in the following section. In situations where the model doesn't capture enough information from the training data, it is deemed under-fitted. A model is deemed "appropriately fitted" when it yields satisfactory outcomes on both the training and testing datasets as shown in Fig 2 [23-24].

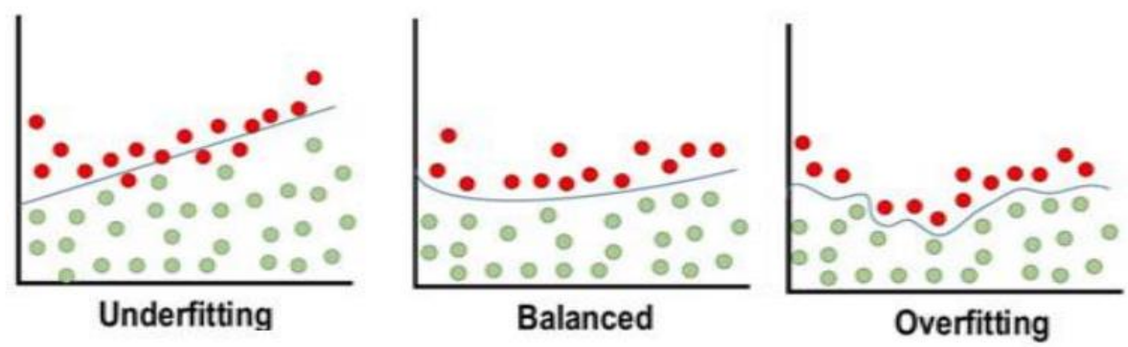

Figure 2: Convolutional Neural Network Generalization

1.4 Popular Convolutional Neural Network Models

R-CNN [36–38] stood out as a groundbreaking model, introducing convolutional neural networks (CNN) into the scene. The model, tasked with classifying images, generates 2000 region proposals, resizing them to 227 × 227. Its distinctive features include a region-of-interest (RoI) classifier based on a deep convolutional neural network (DCN) for specific region classification within input images. The process involves utilizing a convolutional neural network (CNN) for feature extraction and model training, followed by object categorization using a support vector machine (SVM) classifier. Despite its effectiveness, the model operates at a notably slow pace. In response, Fast R-CNN emerged in 2015 to address both accuracy and speed concerns [37,38], emphasizing RoI extraction from feature maps in SPPNet and Fast R-CNN, an advanced iteration of R-CNN that outperformed the conventional framework.

Faster R-CNN advances this trajectory by introducing region proposal networks for feature extraction, aiming to eliminate storage costs [37,38]. This enhanced version of Fast R-CNN achieves comprehensive end-to-end training using RPN-based fully contained region proposal networks. Regression-based region-of-interest (RoI) networks (RPNs) play a crucial role in generating RoIs.

Compared to its predecessors, Faster R-CNN excels in accuracy and speed; however, a misalignment between ground truth and predicted bounding boxes persists. To address inaccuracies in the region of interest (RoI) pooling layer stemming from the quantization process, the authors propose the incorporation of Mask R-CNN.

Mask R-CNN takes a step further, building upon Faster R-CNN by adding a mask prediction branch, enabling simultaneous object detection and mask prediction. R-FCN, meanwhile, opts for replacing fully connected layers with position-aware score maps, resulting in improved object detection capabilities.

1.5 RESNET50

Often used in transfer learning, the AlexNet, AlexNetOWTBn, GooLeNet, Overfeat, and VGG models have a stacked configuration of convolutional layers. Deep CNN networks do, however, face difficulties such degradation problems, the vanishing gradient problem, and network optimization. To tackle these issues, the Residual network (ResNet) presents a revolutionary method that improves detection accuracy and provides solutions for complex jobs. ResNet focuses on deep CNN training challenges, including as saturation and accuracy degradation. This work uses the ResNet50 architecture, which is represented in Fig. 3 and consists of 50 layers of residual networks.

In Fig. 3, several groups of similar layers that make up the ResNet50 structure are indicated by different colors. The curving lines represent identity blocks, which indicate that later levels make use of earlier layers. This feature, which mitigates problems like vanishing or expanding gradients and the degradation issue during the training of deep networks, is crucial to ResNet50. The first layer in Figure 3 consists of 64 filters with a 7×7 kernel size and is succeeded by a max-pooling layer that is 3×3. Different sets of layers, each denoted by a different hue, are made up of different numbers of similar blocks. Blue curves represent the connections between layers of varying sizes. For the classification challenge, a total of 38 fully connected layers are created after these blocks. Nevertheless, these completely connected layers are not used in our suggested model.

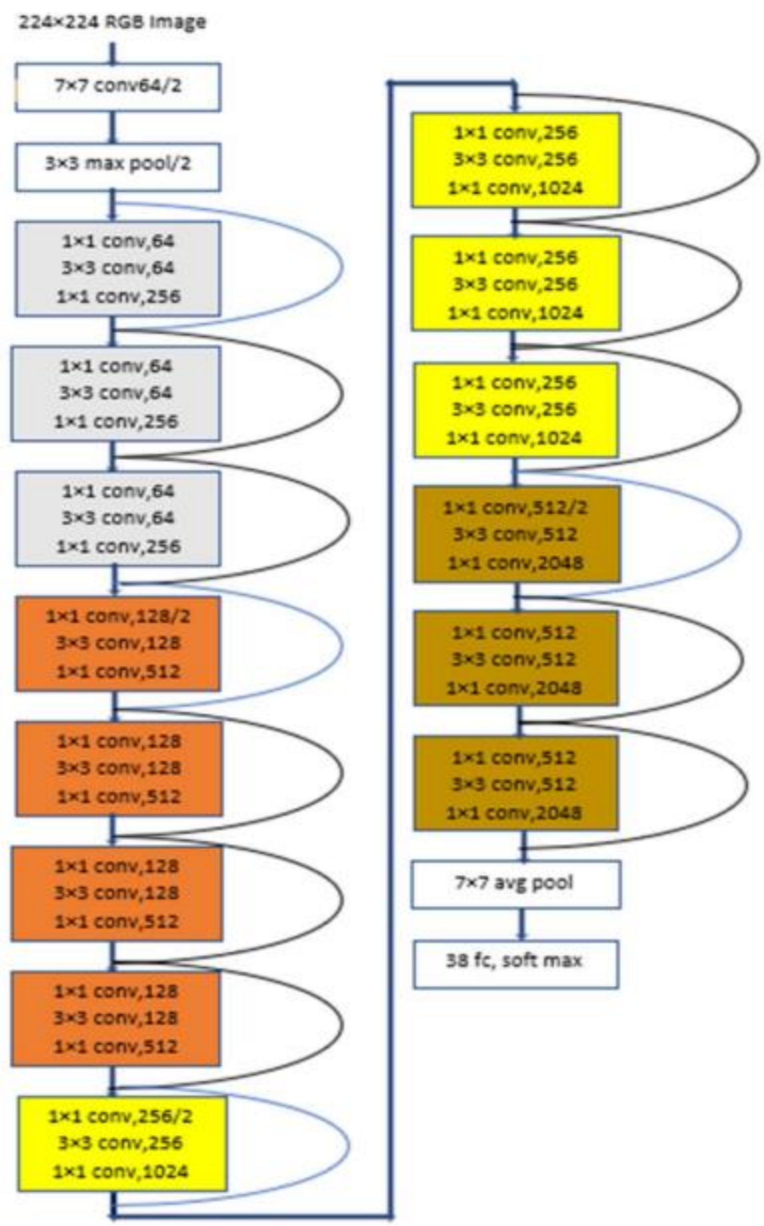

Figure 3: ResNet50 Architecture

1.6 Vision Transformer

ViTs are inspired by the well-known Transformer approach that is commonly used in natural language processing [26]. Because of their self-attention mechanisms, which maintain the dependency of words in the sequence representation, transformers—which are specifically made for processing word sequences—are excellent at maintaining long-distance dependencies [26]. In the field of computer vision, Ref. [27] offered a pure Transformer technique after several approaches for hybrid architectures fusing convolutional processes and self-attention mechanisms. Compared to state-of-the-art CNNs, the first version of ViT uses a transformer applied directly to picture patch sequences, and it achieves similar results on a variety of image classification datasets with a much less training computational cost. An image is divided into fixed-size patches that are all linearly embedded during the ViT process. After adding position embedding, the vector sequence is put into a typical Transformer encoder. ViTs are distinguished by their relaxation of the CNN-associated locally constrained receptive field and the translation invariance constraint. This suggests that ViTs are more appropriate for encapsulating an image's global object organization [28].

New self-supervised learning (SSL) techniques specifically designed for ViTs highlight the already noted performance advantage of ViTs over CNNs. The main goal of an SSL approach in computer vision is to force a network to predict any concealed portion of an image from an unhidden portion. The concealed portion could entail cropping and zooming, as in DINO (self-distillation with no labels) [29], or partial masking of image patches, as in Masked AutoEncoders (MAE) [30], or the Bidirectional Encoder representation from Image Transformer (BeiT) SSL method [31], which was particularly selected and assessed for this study. The model can be more effectively fine-tuned in a traditional supervised fashion on datasets having picture labels, like ImageNet, after being pre-trained using the SSL technique on a large collection of unlabeled photos. This rigorous procedure, as demonstrated in this paper, produces a new, better-pretrained model that may then be improved for a variety of downstream tasks. The second part of this study consists of private data protection challenges, and regulatory compliance challenges in cybersecurity, with increasing automation technologies, big data, and Artificial Intelligence, making human lives easier but also bringing new kinds of cyber threats and challenges [9]. Recital 1 of GDPR which is the "*Data protection as the fundamental rights*" declares that one of the fundamental rights is the protection of natural persons with regard to the handling of their personal data. 2 Everyone has the right to the protection of personal information pertaining to them, as stated in Article 8(1) of the Charter of Fundamental Rights of the European Union (the "Charter") and Article 16(1) of the Treaty on the Functioning of the European Union (TFEU) [34]. On the other hand, Pakistan lacks comprehensive data protection laws that strictly govern issues pertaining to the handling of personal data. Currently, the main piece of legislation that offers a legal framework for many types of electronic crimes and also covers illegal access to personal data is the Prevention of Electronic Crimes Act, 2016 (commonly known as "PECA"). To amend the Personal Data Protection Act, 2023 (also known as "the Act"), MOITT has introduced the Personal Data Protection Bill 2023 (also known as "the Bill"). The Act has not yet been signed into law. Once the Bill is passed into law, it will serve as the primary piece of legislation governing data controllers and processors in Pakistan. It will apply to anyone who handles, oversees, or approves the processing of any personal data, as long as the data subject, data controller, or data processor (local or foreign) is based in Pakistan [35]. In addition to all above, the core objectives of cyber security regulations are to prevent data breaches and protect private data ensuring the confidentiality and integrity of private data, also regulatory compliances have many further challenges for example they must be continuously updated to address the emerging and advanced threats to protect the private data [10].

2. Chapter 2: LITERATURE REVIEW

This section of our study is about the discussion of recent related work on Plant health monitoring and disease detection based on deep learning models. Digital photography is a very useful tool for early-stage symptom recognition when it comes to plant disease detection and identification. Several scientists have developed agricultural applications, such as ones for fruit and leaf disease detection [11]. Even though conventional techniques are accurate, there is a rising need for automated disease detection to reduce the need for human involvement in the analysis and diagnosis of illnesses in plant leaf patterns [12]. Using their plant-specific datasets, Sladojevic and colleagues [13] applied the CaffeNet model for the identification of illnesses in various plants. To improve the training of their models, they included a data augmentation procedure, especially to deal with issues related to tiny datasets, after 100,000 iterations, the top-5 success rate was 99.99 percent and the top-1 success rate was 96.3 percent. These findings show how well the deep convolutional network model recognizes plant illnesses from leaf photos, with a high degree of efficiency and accuracy. In a different study, Mohanty et al. [14] used both transfer learning and the learning-from-scratch method for model training. They used AlexNet and GoogleNet to detect 26 illnesses of 14 crops in the plant village dataset. With GoogleNet in place, a phenomenal peak accuracy of 99.34 percent was achieved. Ferentinos [15] identified 58 plant leaf diseases with success using five different pre-trained deep learning models: VGG, AlexNet, AlexNetOWTBn, Overfeat, and GoogleNet. In order to detect plant illnesses, Geetharamani and Pandian [16] used a nine-layer deep CNN, achieving a 96.46 percent accuracy rate. By replacing the fully linked layer of AlexNet with the inception layer, Liu et al. [17] created a model that was inspired by the topologies of GoogleNet and AlexNet. This model had a 97.62 percent accuracy rate in correctly identifying four different apple leaf diseases. For plant leaf disease identification, Transfer learning of deep convolutional neural networks was used by Junde et al. [18]; the authors pre-trained VGGNet using ImageNet and selected the Inception modules for the approach. The recommended approach outperforms other state-of-the-art methods and achieves a validation accuracy of at least 91.83 percent on the publicly available dataset. In this paper [19], the author created a deep ensemble neural network that incorporates transfer learning as part of an efficient model to identify plant leaf disease within limited computational resources, hence reducing false positives and false negatives. The architecture of this proposed methodology is illustrated in Fig. 4. Additionally, the authors in their study used an ensemble neural network and DenseNet. DenseNets serve to streamline the connectivity patterns between layers by establishing direct connections among them, addressing the constraint on maximum information flow. In comparison to traditional CNN, DenseNets exhibit a reduced number of feature maps.

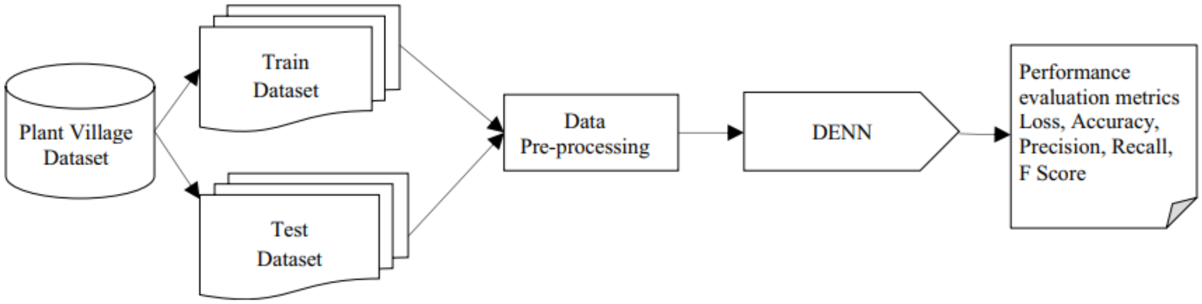

Figure 4: Architecture of the proposed study.

Deep learning models, specifically ResNet-50 and VGG-19, for identifying and classifying diseases in banana plants based on leaf images the Convolutional Neural Networks (CNN) to process the image data and classify it into different disease categories such as health, Sigatoka, Cordana, and Pestaloptiosis the use of ResNet-50 and VGG-19 models to achieve accurate disease identification and classification [32] several research papers and articles that have utilized deep learning models for various image processing tasks, including disease diagnosis, image captioning, and object classification but the architecture and features of ResNet-50 and VGG-19, highlighting their capabilities in handling image data and extracting meaningful patterns for classification purposes, explains the study methods in further detail, covering model evaluation, model architecture design, and dataset processing. In an effort to determine the best method for identifying diseases in banana plants, it also compares the outcomes of the ResNet-50 and VGG-19 models. In general, the work focuses on employing deep learning models—ResNet-50 and VGG-19 in particular—to handle the significant challenge of using leaf pictures to identify and categorize illnesses in banana plants [32].

The paper's authors [33] address the shortcomings of Pakistan's cyber security framework and offer suggestions for how to make it better. They also emphasize the country's evolving cyber laws and the Electronic Transactions Ordinance, 2002's narrow scope. They also talk about Pakistan's difficulties in controlling the digital world, particularly internet users' lack of information technology literacy. To protect financial assets and respond to cyberattacks, the study suggests establishing a thorough cybersecurity policy, collaborating with other countries, and setting up a National Cyber Coordination Center. Furthermore, recommended by the authors are running awareness programs and taking part in global initiatives to establish cyber norms.

3. Chapter 3: METHODOLOGY

3.1 Material and Methods

The primary aim of this study is to facilitate the seamless integration of computer vision technologies in order to address common challenges in the classification of symptoms associated with plant diseases. Beyond plant disease diagnosis, the suggested approaches find use in a variety of industrial settings where the quick development of machine learning algorithms is required. When a large enough training dataset is available, prior studies have demonstrated the remarkable performance of sophisticated convolutional neural networks (CNNs) in these kinds of tasks. But it's frequently not possible to get a lot of photos for particular crop-disease combos. By using data augmentation approaches, which allow CNNs to learn unique aspects of illness classes, this lack of data and class imbalance are efficiently addressed.

Furthermore, inaccuracies in data collection from natural settings could lead to misclassification due to issues including dim lighting, blurry images, and unclear symptoms. Due to their high computational costs, large model sizes, and lengthy running times, classical CNNs are not as successful when employed in mobile devices with limited resources. This document offers a thorough explanation of the experimental configuration created to examine the effects of these circumstances and suggests fixes inside our process.

We suggest a unique method that use a CNN without a dense layer to address model size concerns. As a result, there are far fewer add-multiplication processes, which makes the model more efficient and compact. Furthermore, we recognize the potential of vision transformers to bring about transformation and systematically incorporate this architecture into our methodology. Since vision transformers use attention techniques to extract long-range relationships from the data, they present a viable option, particularly in situations when computational resources are scarce. The datasets and their difficult qualities are thoroughly discussed in the portions of the study that follow. Comprehensive justifications of the methods utilised to improve the overall functionality of our suggested system—which now includes vision transformers—are also provided.

3.2 Dataset

Many well-known datasets, such as the Plant Pathology Dataset, Fruit Disease Dataset, Tomato Leaf Dataset, and others, include a large number of photos intended for the purpose of identifying plant diseases. A large variety of crops and illnesses are covered by these datasets. For example, the Plant Pathology dataset available on Kaggle displays leaves impacted by rust, scab, and several leaf spot kinds. The Fruit Disease Dataset, which is also accessible on Kaggle, is an important tool for researchers studying fruit-related disorders. Conversely, the Tomato Leaf dataset addresses particular tomato-related illnesses. Additionally, there are crop-specific datasets for wheat, rice, and soybeans, which contribute to a comprehensive comprehension of plant diseases across various crops. In the quest for an advanced computer vision algorithm for detecting crop diseases, the PlantVillage Dataset was selected as a fundamental learning resource. The approach involved transfer learning using a pretrained model from the ImageNet dataset, followed by fine-tuning it with the extensive PlantVillage dataset. This method took advantage of the pretrained model's proficiency in identifying objects and features within images.

3.2.1. The PlantVillage Dataset

The PlantVillage dataset is a large and comprehensive collection of plant photos that have been carefully selected to further the development of computer vision algorithms for agricultural disease detection. The National Science Foundation and the Bill and Melinda Gates Foundation have generously provided financial support to the Penn State PlantVillage team, led by Dr. David Hughes. This team has worked to create this dataset. The PlantVillage initiative has demonstrated a noteworthy dedication to advancing agricultural research by providing the public with unrestricted and open access to this dataset. The extensive PlantVillage dataset contains more than 54,000 high-quality photos that illustrate 26 major agricultural diseases affecting 14 different plant species. A total of 38 class labels are added to the dataset by include photos of healthy plants. This repository was assembled by a collaborative effort involving contributions from professional photographers, citizen scientists, and farmers. In Figure 5, a representative snapshot of leaf images sourced from the PlantVillage dataset provides a visual insight into the diversity and scope of this invaluable resource.

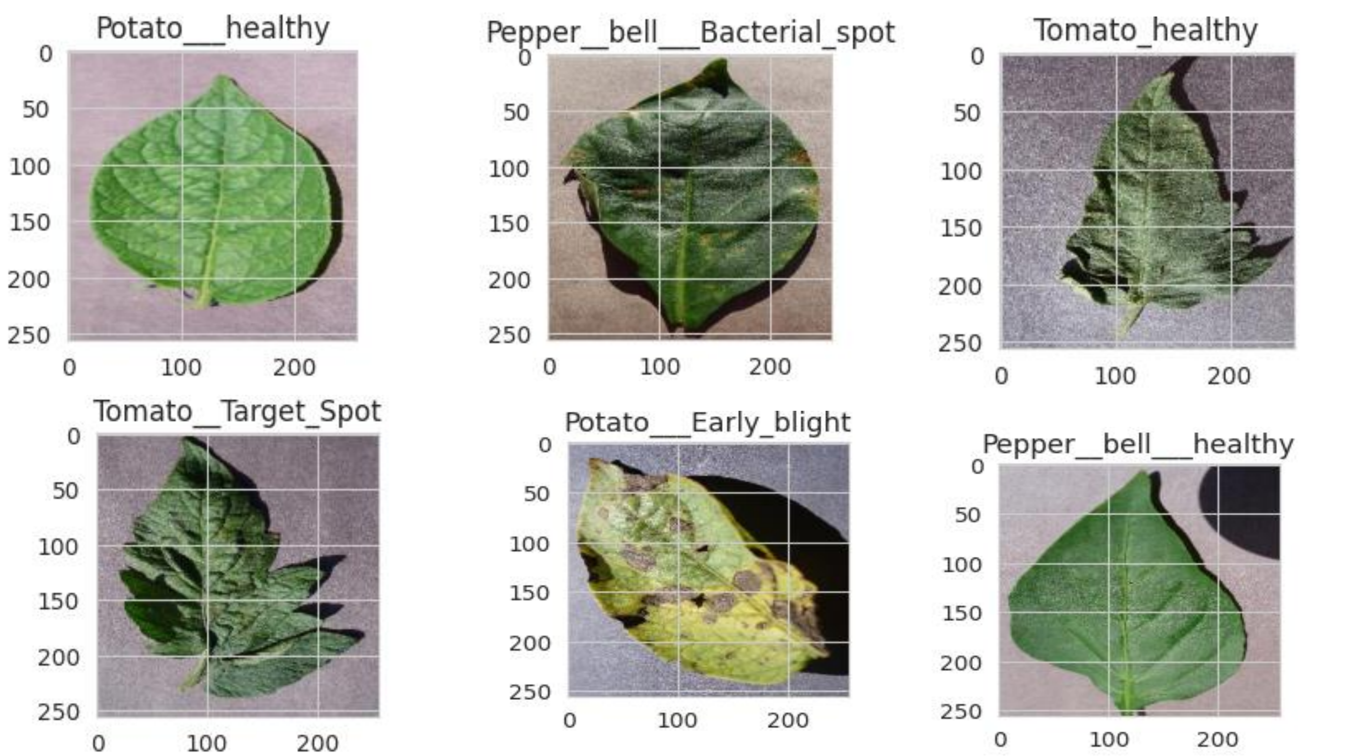

Figure 5: The Plant Village Dataset

It is crucial to underline that the PlantVillage dataset incorporates images not only of leaves but also fruits, showcasing the comprehensive nature of the dataset. The spectrum of plant species featured in this collection is extensive and diverse, encompassing apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry, and tomato. This deliberate inclusion ensures that the dataset mirrors the agricultural landscape with a rich variety of plants susceptible to diseases.

Plant diseases such as early blight, bacterial spot, common rust, late blight, leaf curl, mosaic virus, powdery mildew, Septoria leaf spot, spider mites, target spot, tomato yellow leaf curl virus, and two-spotted spider mites are all included in the PlantVillage dataset. The inclusion of these varied illnesses highlights the importance of the dataset in tackling the complex problems caused by plant diseases in a range of crops. In addition, Figure 6 below shows the distribution of each class in the dataset.

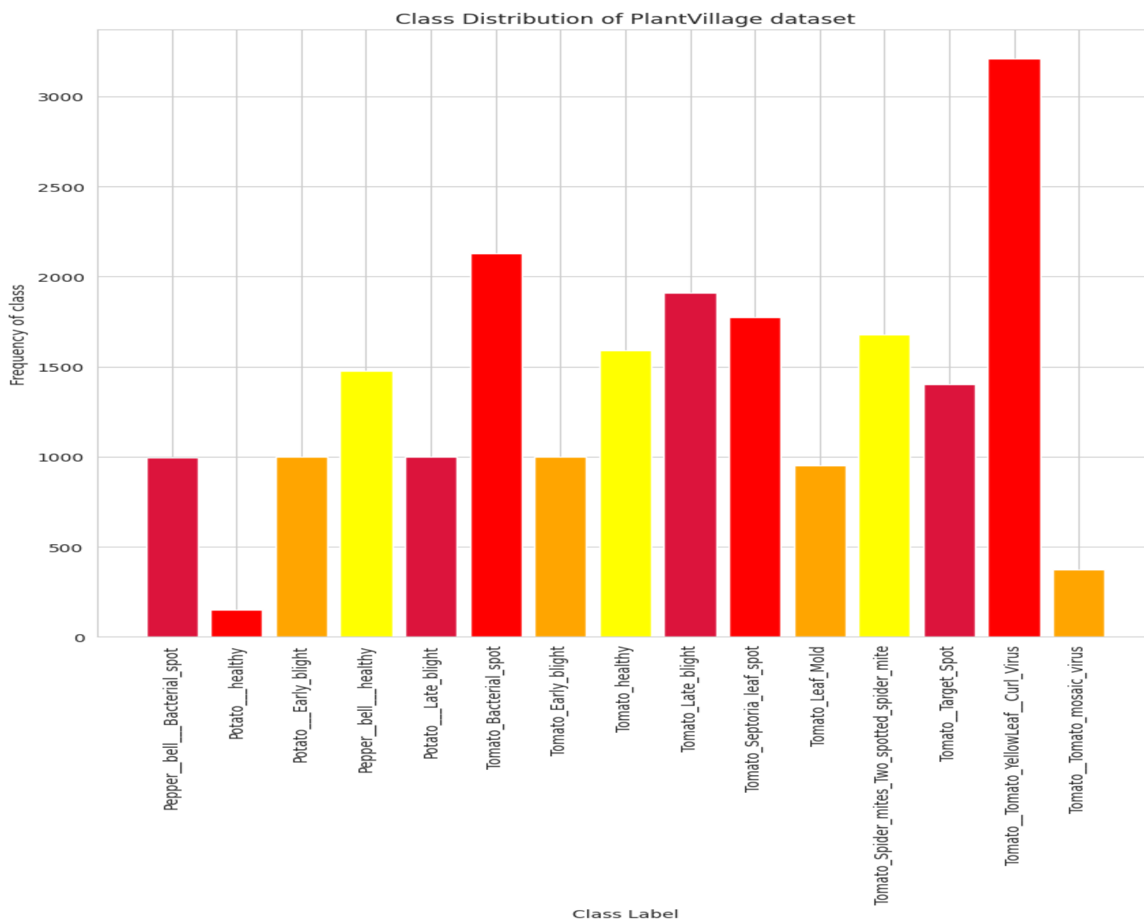

Figure 6: Class Distribution of PlantVillage dataset

As mentioned above in fig 6, our PlantVillage dataset contains 15 different classes, the classes explicitly are:

- Pepper__bell___Bacterial_spot
- Pepper__bell___healthy
- Potato___Early_blight
- Potato___healthy
- Potato___Late_blight
- Tomato__Target_Spot
- Tomato__Tomato_mosaic_virus
- Tomato__Tomato_YellowLeaf__Curl_Virus
- Tomato_Bacterial_spot
- Tomato_Early_blight
- Tomato_healthy
- Tomato_Late_blight
- Tomato_Leaf_Mold
- Tomato_Septoria_leaf_spot
- Tomato_Spider_mites_Two_spotted_spider_mite

3.2.2. The ImageNet Dataset

The research is greatly aided by the use of the ImageNet dataset [25], which produces a number of benefits. It's one of the biggest datasets out there right now, which makes it a great tool for building machine learning models. Using the WordNet hierarchy to classify photographs improves the precision and effectiveness of object recognition models. Machine-learning algorithms have benefited greatly from the use of the ImageNet dataset.

This research's favored data source includes over 50,000 color images of crop leaves, including both healthy and injured plants. This dataset's significant sample size, accurate labeling, and capacity to detect agricultural illnesses are among its noteworthy qualities.

14,197,122 annotated photos arranged according to the WordNet hierarchy make up the ImageNet dataset. The richness and diversity of the dataset are further enhanced by the fact that there are 1000 separate object categories, and each category has an average of 1000 photos.

3.3 Data Pre-Processing

The process of analyzing data involves extracting pertinent information from data sources, and in this investigation, the utilization of Tensor Flow played a crucial role in guaranteeing meticulous findings. Various Python libraries were utilised to examine the information contained in the PlantVillage databases. In particular, Matplotlib was used to visualise the libraries used in the study, and Tensor Flow, NumPy, and Keras were used to build the neural network architecture. These libraries are essential to the research and provide different functions. Among these, TensorFlow stands out as an essential library for machine learning model development and implementation, offering a feature-rich and flexible neural network construction environment. One important feature of this study is that it is particularly well-suited for managing large datasets due to its effectiveness and scalability. Unquestionably, data analysis ensures the accuracy and dependability of information. TensorFlow, NumPy, Keras, and Matplotlib are just a few of the Python-based tools that make it easier to analyse PlantVillage datasets. These tools work together to ensure that the data is analysed thoroughly and effectively, producing reliable results.

3.4 Models

3.4.1 CNN

One popular deep learning framework that is often used for image and video analysis is the Convolutional Neural Network (CNN). Its architecture is made up of several layers that are intended to extract information from input movies or photos. Low-level characteristics like edges and corners are identified by the first layer, sometimes referred to as the convolutional layer. In order to extract higher-level features like forms and objects, the ensuing layers expand upon these low-level properties. Pooling layers are used after the convolutional layer to down-sample the feature maps and reduce spatial dimensions without losing important information. For tasks involving classification or regression, the output from the convolutional and pooling layers is subsequently flattened and fed through fully connected layers. CNNs are able to process images of different sizes and orientations because of their special capacity to learn the spatial hierarchies of features through convolution and pooling processes. We especially used the well-known to be successful ResNet50 CNN architectures and ResNet50 with transfer learning for our investigation. In the framework of this work, we first designed a data generator that efficiently handles input data, from which we were able to construct a strong model for image classification. Then, we leveraged the ResNet50 architecture that we had acquired from Keras Applications to apply transfer learning. The ResNet50 model's layers were all made non-trainable in order to improve the model's flexibility. Additionally, we introduced a specialized final classification layer to accurately categorize 15 distinct classes. Following this configuration, the model was compiled, integrating a carefully selected optimizer, loss function, and evaluation metric.

Furthermore, the model was systematically trained on the PlantVillage dataset, adjusting its parameters to optimize performance for the specific classification task at hand. The subsequent phase involved fine-tuning, where a pretrained model was used, but only a subset of the layers was designated as trainable in order to identify patterns within the current dataset. To further improve the robustness of the model, regularization was implemented through the strategic inclusion of dropout layers.

During the fine-tuning process of the ResNet50 model, we loaded the model from Keras Applications and specifically designated the last 75 layers of the ResNet50 model as trainable. A final layer was then added to facilitate classification across the 15 distinct classes. Following this fine-tuning configuration, the model was compiled, incorporating a suitable optimizer, loss function, and evaluation metric. The concluding step involved training the model on the available dataset, and refining its parameters to optimize performance that is specifically tailored to the classification task at hand.

3.4.2 Vision Transformer

To construct a Vision Transformer architecture, the procedure encompasses multiple stages. Initially, an image is divided into patches, and these patches are subsequently flattened into a singular dimension, resulting in feature maps of dimensions (196, 768). The resultant output is then converted into the desired format of $N \times (P^2 \cdot C)$, where P represents the patch size, and C signifies the channel dimension.

a. Normalization layer

In order to facilitate the Transformer's encoder, essential layers are incorporated, commencing with Layer Normalization (LN or LayerNorm) utilizing torch.nn.LayerNorm(). This process of normalization, employed over the final dimension, aids in enhancing training efficiency and the model's generalization ability to unseen data.

b. Multi-head Attention

The Multi-Head Self Attention (MSA) layer is implemented by utilizing torch.nn.MultiheadAttention(), with parameters such as the embedding dimension (D), the number of attention heads, and the option to apply dropout for regularization. This particular layer is of utmost importance in capturing intricate relationships within the data.

c. MLP block

Additionally, an MLP Block is seamlessly integrated into the architecture, and the complete Transformer Encoder block is assembled by amalgamating the custom-made layers mentioned above. This comprehensive block plays a pivotal role in processing and extracting meaningful features from the input data.

Once these components are in place, the subsequent step entails creating the entirety of the Vision Transformer. Now the architecture has been established shown in Fig 7, the model is trained using the PlantVillage dataset. The training process encompasses optimizing the model parameters to enhance its proficiency in performing image classification or other pertinent tasks.

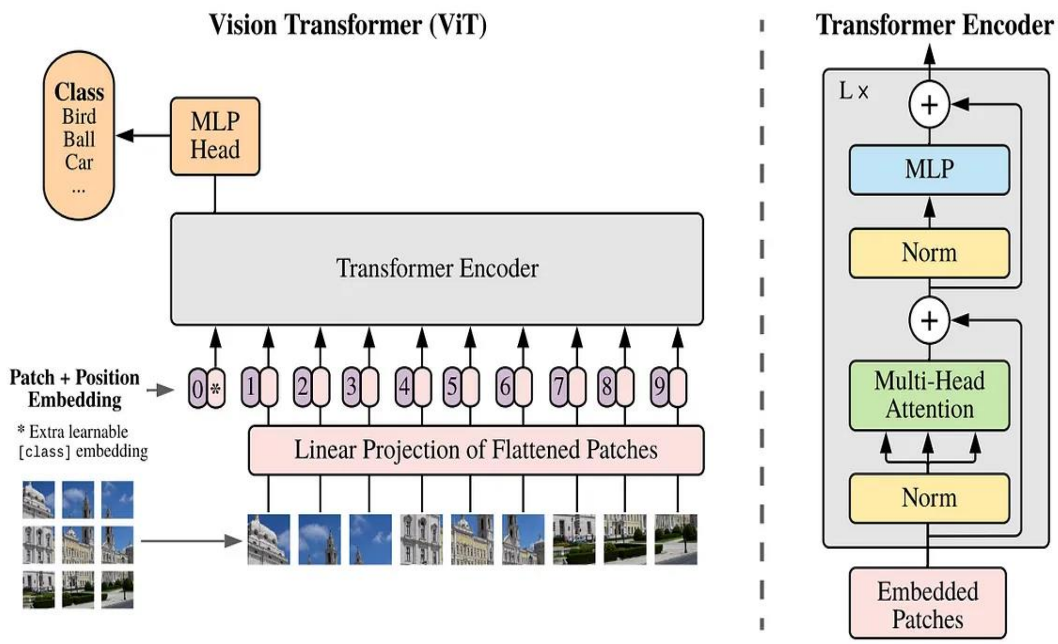

Figure 7: Architecture of Vision Transformer

During the initial phase of the training process, the effort was directed towards training the Vision Transformer (VIT) from the ground up. The dataset was divided into separate sets for training and testing purposes. In order to make efficient use of the computational resources, a batch size of 32 was utilized, and the available CPU count was taken into account by setting NUM_WORKERS appropriately. It is worth noting that the image size was standardized to 224, and for optimal feature extraction, a patch size of 16 was employed. The input image, which adhered to the dimensions torch Size ([1, 3, 224, 224]), underwent a transformation that resulted in an output patch embedding with dimensions torch Size ([1, 196, 768]). Throughout the training process, which spanned 10 epochs, the model was refined using the Cross Entropy Loss technique to optimize its parameters and improve its capabilities in image classification and other relevant tasks. This rigorous training process aimed to equip the Vision Transformer with the ability to discern complex patterns and relationships within the data, ultimately leading to enhanced performance across various image-based tasks.

4. Chapter 4: Results

4.1 Results and Discussion

We commenced the training of ResNet50 with a span of 10 epochs, wherein we were able to attain an utmost training accuracy of 98.09%. Nevertheless, the pinnacle of the validation (testing) accuracy culminated at 91.5%. Training accuracy and validation accuracy graph is shown in fig 8.

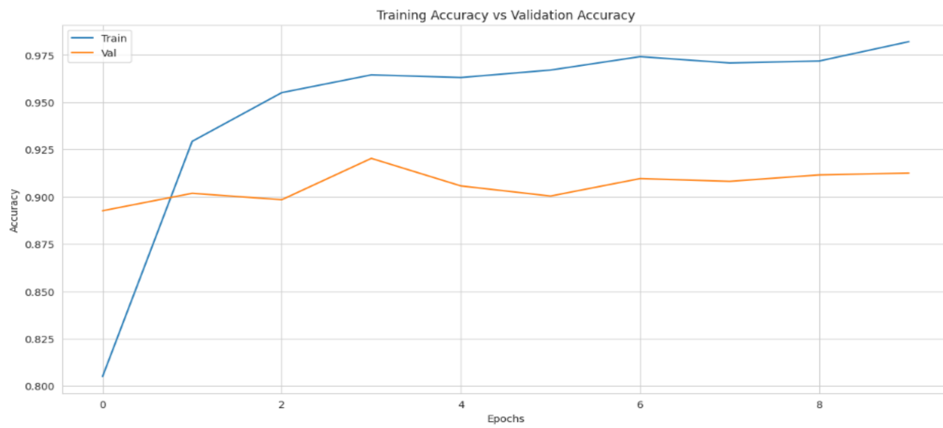

Figure 8: Resnet50

In order to improve even more, we used the Fine-Tuning method with the ResNet50 pre-trained model in Keras. The ImageNet dataset was used to train ResNet50 initially. In this approach, we selectively designated the last 75 layers of the ResNet50 model as trainable, enabling it to adapt and identify patterns in the current dataset. Additionally, regularization methods were integrated by incorporating dropout layers. The model demonstrated notable improvements during this fine-tuning process, with training accuracy of 98.09 percent and testing accuracy of 95.23 percent, as indicated in table 1. These results outperform the original ResNet50 model, demonstrating how the fine-tuning approach may be used to enhance the model's performance for the given job. The training accuracy and validation accuracy are shown graphically in figure 9.

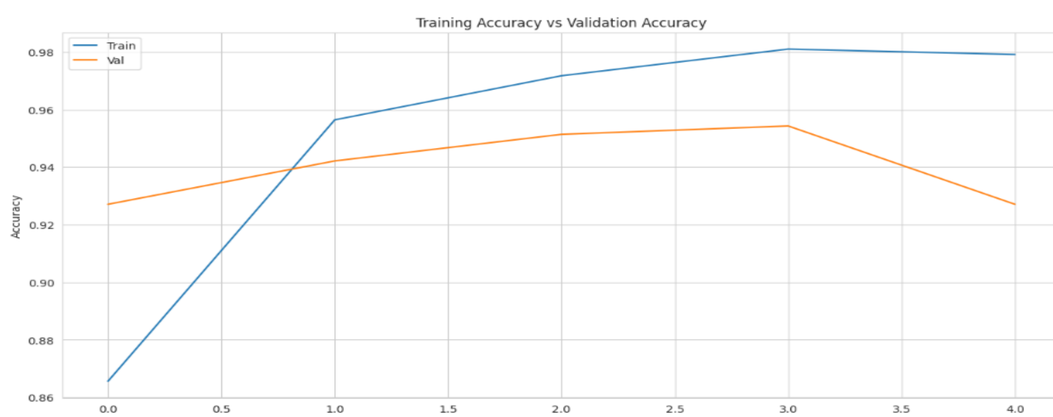

Figure 9: Resnet50 with transfer learning

Furthermore, in this study two distinct instructional methodologies were examined for the Vision Transformer (VIT). At first, the focus was on carefully dividing the dataset into training and testing sets in order to train the VIT architecture from the start. A batch size of 32 was utilized, and computational resources were fine-tuned by setting NUM_WORKERS to the available CPU count. The standardized image size was established at 224, employing a patch size of 16 to facilitate efficient feature extraction. Input images underwent transformation, leading to an output patch embedding.

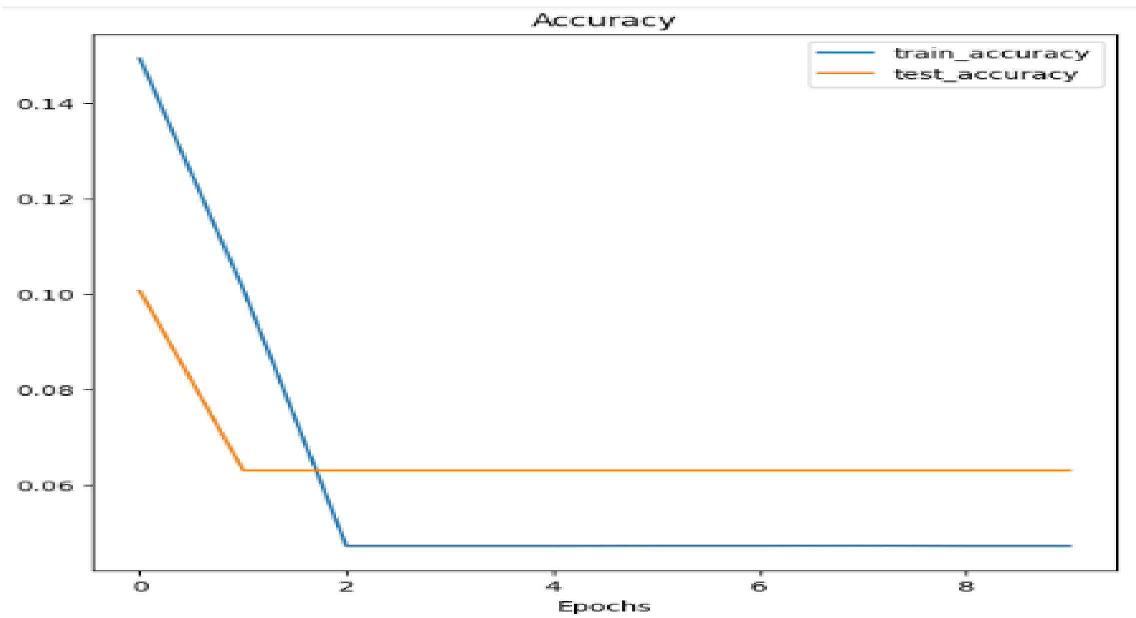

Figure 10: VIT Scratch

The model underwent rigorous training over the course of 10 epochs, shown in fig 10 utilizing Cross Entropy Loss for parameter optimization in the realm of image classification and related tasks. However, despite these endeavors, the testing accuracy fell short of expectations, remaining at approximately 60% given in table 1.

Subsequently, an alternative approach was implemented whereby a pretrained Vision Transformer (VIT) was utilized. Through this utilization of pretrained models, the aforementioned methodology exhibited a marked improvement in performance, ultimately attaining an accuracy level of 98% given in table 1. This observation serves to underscore the efficacy of harnessing pre-existing knowledge in the guise of pretrained models, thereby augmenting the overall performance and accuracy of the Vision Transformer. The figure 11 depicts the graphical representation of the training accuracy and validation accuracy.

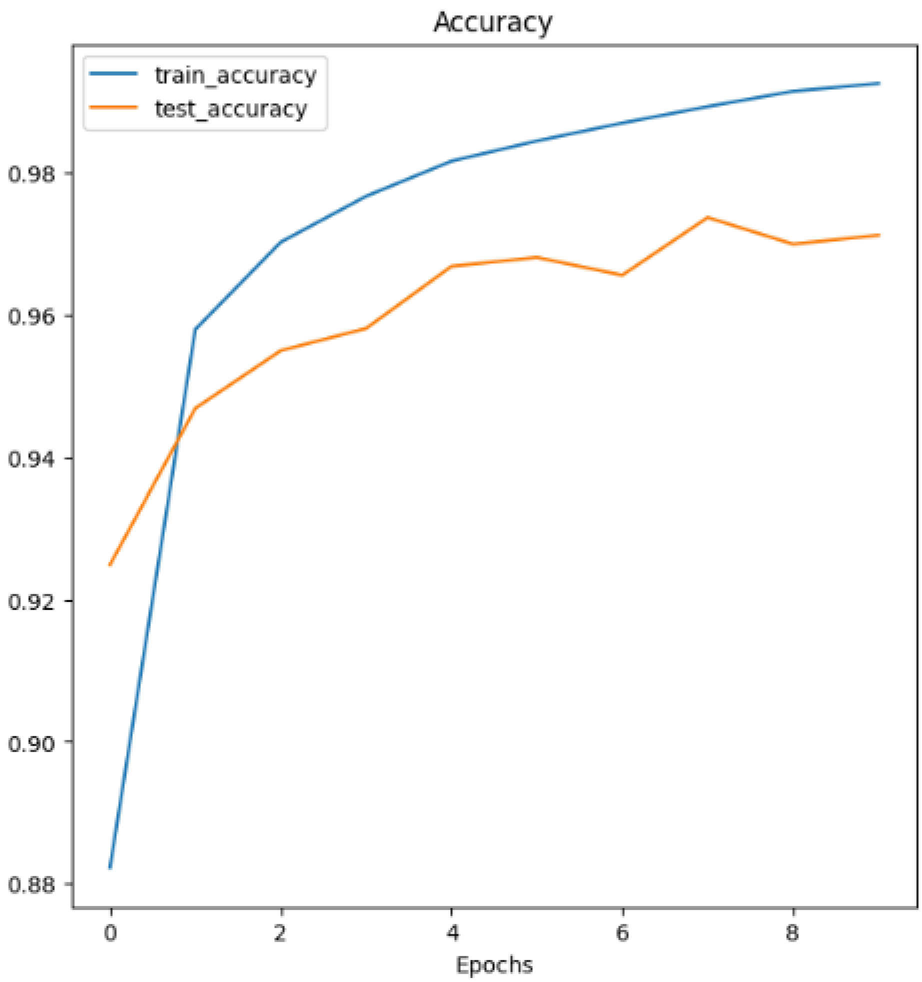

Figure 11: VIT pretrained

The evaluations encompassed various performance metrics, such as:

1. Training Accuracy: The percentage of successfully predicted occurrences relative to the total number of training examples is measured by the training accuracy statistics. It provides insightful information on how well the model learned from the given training set.

2. Test Accuracy: The ratio of accurately predicted instances to the total number of test instances is known as test accuracy. It functions as a gauge of the model's generalization ability by assessing how well it performed on unknown data during training.

3. F1 Score: The F1 score represents the harmonic mean of precision and recall. This score strikes a balance between precision, which denotes the ratio of true positive predictions to the total predicted positives, and recall, which signifies the ratio of true positive predictions to the total actual positives. The F1 score proves particularly valuable when confronted with an imbalanced class distribution.

4. Recall: Recall, also referred to as sensitivity or true positive rate, quantifies the ratio of true positive predictions to the total actual positives. It serves as a measure of the model's aptitude in accurately identifying instances belonging to a specific class.

5. Precision: Precision, on the other hand, denotes the ratio of true positive predictions to the total predicted positives. It quantifies the accuracy of positive predictions and is particularly critical when the occurrence of false positives bears significant consequences.

6. Confusion matrix: A tabular representation called a confusion matrix compares the actual and predicted class labels to assess how well a classification model performs. There are four types of predictions in it: False Positives (positives that are not accurately anticipated), False Negatives (negatives that are correctly predicted), and True Positives (positives that are correctly predicted) (negatives incorrectly predicted). The heatmap provides a visual representation of these measures; the cell colours represent the frequency of predictions shown in Figures 7 and 8, respectively, for resnet50 and resnet50 with transfer learning.

S.no	Model	Epochs	Training accuracy	Testing accuracy
1.	Resnet50	10	97.50%	91.474
2.	Resnet50 with transfer learning	10	98.09%	95.23%
3.	VIT	10	48	64%
4.	Pretrained VIT	10	99	97%

Table 1: Comparison of Models

	Resnet50			Resnet50 with Transfer Learning		
Classes	Precision	Recall	F1-score	Precision	Recall	F1-score
Pepper bell Bacterial spot	0.88	0.93	0.90	0.96	1.00	0.98
Pepper bell healthy	0.90	0.99	0.95	0.99	0.98	0.99
Potato Early blight	0.96	0.93	0.94	1.00	0.99	0.99
Potato Late blight	0.95	0.72	0.82	0.84	0.95	0.89
Potato healthy	0.89	0.50	0.64	0.74	0.88	0.80
Tomato Bacterial spot	0.87	0.96	0.91	0.99	0.99	0.99
Tomato Early blight	0.86	0.55	0.67	0.89	0.94	0.91
Tomato Late blight	0.82	0.88	0.85	0.97	0.90	0.94
Tomato Leaf Mold	0.84	0.93	0.88	0.99	0.97	0.98
Tomato Septoria leaf spot	0.82	0.95	0.88	0.98	0.96	0.97
Tomato Spider mites Two spotted spider mite	0.96	0.79	0.86	0.94	0.89	0.92
Tomato Target Spot	0.85	0.85	0.85	0.92	0.87	0.90
Tomato Yellow Leaf Curl Virus	0.96	0.96	0.96	0.98	0.98	0.98
Tomato mosaic virus	0.93	1.00	0.96	0.88	1.00	0.94
Tomato healthy	0.98	0.99	0.98	0.95	0.99	0.97

Table 2: Classes and Scores

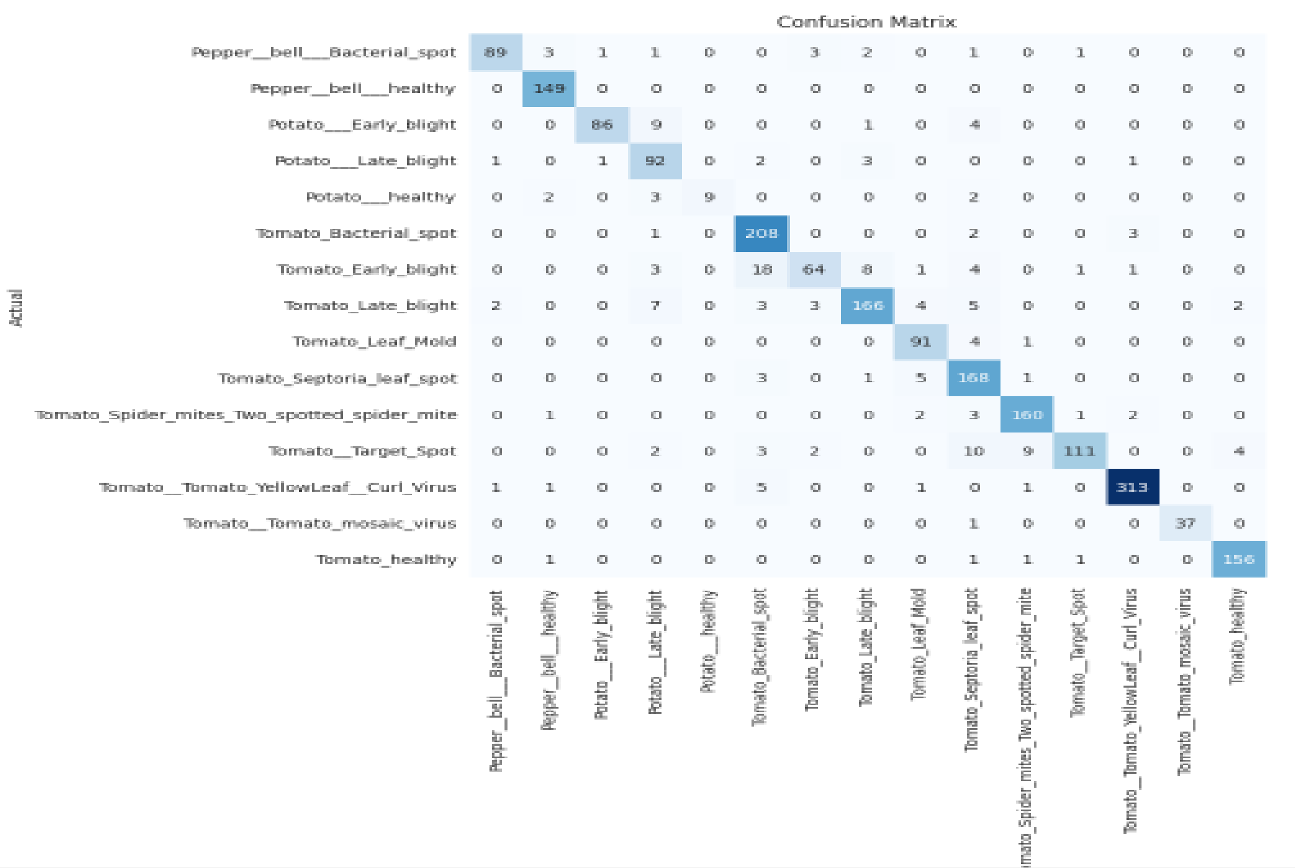

Figure 12: Resnet50

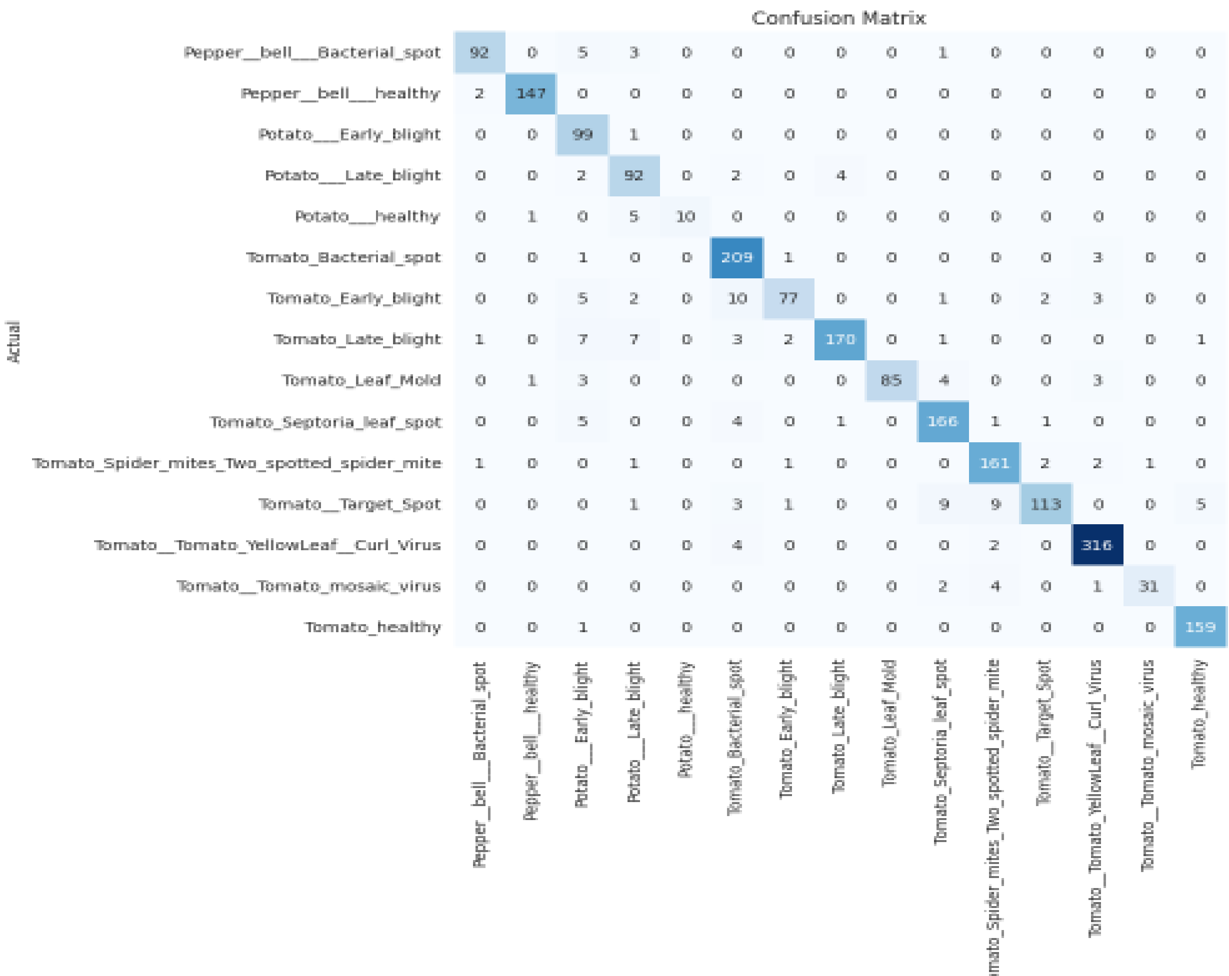

Figure 13: Confusion matrix for resnet50 with transfer learning

4.2 Data Protection and Challenges:

As an emerging country located in the Global South, Pakistan acquired internet connectivity in the early 1990s. Currently, according to the Pakistan Telecommunication Authority (PTA), broadband penetration is reported at 40.95%, with an impressive 87 million subscribers. Currently, 54% of the nation's populace enjoys mobile broadband access, and mobile internet penetration is documented at 26% (GSMA, 2020).

With a significant population relying on information and communication technologies, the cyber domain has emerged as a new frontier, bringing forth challenges associated with the regulation of cybersecurity. As per the 2018 Global Cyber Security Index Report (GCI), Pakistan secured the 94th global ranking (International Telecommunication Union, 2018).

Within the current context of shaping cybersecurity regulations in Pakistan, the foremost challenge revolves around effective implementation. As previously highlighted, the deficient institutional structure in Pakistan poses a major impediment to executing cybersecurity laws. This challenge is further compounded by other inherent issues, including the presence of adversarial intelligence networks and elements that are against the state.

There is a lack of sufficient technological expertise, especially in monitoring foreign spy agencies. Moreover, the nation is susceptible to malware, capable of installing additional malicious software and pilfering personal information from infected computer systems. Another vulnerability arises from Distributed Denial of Service (DDOS) attacks or the unauthorized transmission of data within a computer without the user's knowledge. A notable illustration is the banking sector in the country, which faces the risk of cyber-attacks, leading to a trust deficit between customers and banks.

Similar to any digital information, the security of plant health data is susceptible to cyber threats. The ongoing challenge involves implementing strong cybersecurity measures to shield against data breaches and cyber-attacks.

To tackle these difficulties, a comprehensive and cooperative strategy is necessary. This approach should engage government entities, research institutions, industry stakeholders, and the public to formulate regulatory frameworks for plant health data that are both effective and ethical.

5. Chapter 5: CONCLUSION

Conclusion

Our research significantly advances the field of smart agriculture by introducing innovative methodologies for advanced computer vision algorithms. The acknowledgment of data protection and cybersecurity challenges, particularly in regions like Pakistan, underscores the necessity for collaborative efforts in establishing ethical regulatory frameworks to protect forest health data. In summary, our investigation into plant disease classification reveals the profound impact of model architecture and training approaches. ResNet50, initially achieving a training accuracy of 97.50%, exemplifies the efficacy of transfer learning, resulting in a noteworthy testing accuracy of 95.23%. While Vision Transformers (VIT), a transformative shift occurred with a pretrained VIT model, achieving an exceptional training accuracy of 99% and an impressive testing accuracy of 97%. These findings underscore the crucial role of leveraging pre-existing knowledge, highlighting the broader applicability of transfer learning to optimize model performance. As we navigate the landscape of smart agriculture, these insights contribute significantly to the ongoing refinement of advanced computer vision algorithms, promoting precision and resilience in plant disease detection.

References:

[1] M. Loey, A. ElSawy, M. Afify, Deep learning in plant diseases detection for agricultural crops: a survey, Int. J. Serv. Sci. Manag. Eng. Technol. 11 (2020) 41–58.

[2] U. Shruthi, V. Nagaveni, B. Raghavendra, A review on machine learning classification techniques for plant disease detection, in: 2019 5th International Conference on Advanced Computing & Communication Systems, ICACCS), 2019, pp. 281–284.

[3] P.K. Sethy, N.K. Barpanda, A.K. Rath, S.K. Behera, Deep feature-based rice leaf disease identification using support vector machine, Comput. Electron. Agric. 175 (May) (2020) 105527, doi:10.1016/j.compag.2020.105527.

[4] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern Trends in Hyperspectral Image Analysis: A Review," IEEE Access, vol. 6, pp. 14118–14129, 2018.

[5] J. De Fauw et al., Clinically applicable deep learning for diagnosis and referral in retinal disease, vol. 24, no. 9. 2018.

[6] M. Bakator and D. Radosav, "Deep learning and medical diagnosis: A review of literature," Multimodal Technol. Interact., vol. 2, no. 3, 2018.

[7] A. Rastogi, R. Arora, and S. Sharma, "Leaf disease detection and grading using computer vision technology & fuzzy logic," 2nd Int. Conf. Signal Process. Integr. Networks, SPIN 2015, pp. 500–505, 2015.

[8] W. Dawei, D. Limiao, N. Jiangong, G. Jiyue, Z. Hongfei, and H. Zhongzhi, "Recognition pest by image-based transfer learning," J. Sci. Food Agric., vol. 99, no. 10, pp. 4524–4531, 2019.

[9] Demertzi, V., Demertzis, S. and Demertzis, K. (2023) An overview of cyber threats, attacks and countermeasures on the primary domains of Smart Cities, MDPI.

[10] S. S. Uddin, Governors State University Opus Open Portal to University Scholarship, https://opus.govst.edu/cgi/viewcontent.cgi?article=1137&context=theses.

[11] S. Nosratabadi, S. Ardabili, Z. Lakner, C. Mako, and A. Mosavi, "Prediction of food production using machine learning algorithms of multilayer perceptron and anfis," Agriculture, vol. 11, no.5, p.408, 2021.

[12] H. Guo, K. Jeong, J. Lim, J. Jo, Y. M. Kim, J.-p. Park, J. H. Kim, and K. H. Cho, "Prediction of effluent concentration in a wastewater treatment plant using machine learning models," Journal of Environmental Sciences, vol. 32, pp. 90–101, 2015.

[13] S Sladojevic, M Arsenovic, A Anderla, D Culibrk, and D Stefanovic "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image classification" Computational Intelligence and Neuroscience, Hindwai 2016.

[14] S.P. Mohanty, D.P.Hughes and M. Salathe "Using Deep Learning for Image-Based Plant Disease Detection" Journal Frontiers in Plant Science, Frontiers,2016.

[15] K. P. Ferentinos, ''Deep learning models for plant disease detection and diagnosis,'' Comput. Electron. Agricult., vol. 145, pp. 311–318, Feb. 2018.

[16] G. Geetharamani and P. J. Arun, ''Identification of plant leaf diseases using a nine-layer deep convolutional neural network,'' Comput. Electr. Eng., vol. 76, pp. 323–338, Jun. 2019.

[17] B. Liu, Y. Zhang, D. He, and Y. Li, ''Identification of apple leaf diseases based on deep convolutional neural networks,'' Symmetry, vol. 10, no. 1, p. 11, 2018.

[18] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," Computers and Electronics in Agriculture, vol. 173, p. 105393, 2020. doi:10.1016/j.compag.2020.105393.

[19] S. S. Anupam Bapat, "Plant Leaf disease detection using Deep Learning," International Journal of Advanced Science and Technology, http://sersc.org/journals/index.php/IJAST/article/view/14162.

[20] M. M. Taye, "Theoretical understanding of Convolutional Neural Network: Concepts, architectures, applications, Future Directions," Computation, vol. 11, no. 3, p. 52, 2023. doi:10.3390/computation11030052.

[21] J. Koushik, "Understanding convolutional neural networks," arXiv.org, http://arxiv.org/abs/1605.09081 (accessed Dec. 20, 2023).

[22] T. Bezdan and N. Bačanin Džakula, "Convolutional neural network layers and architectures," Proceedings of the International Scientific Conference - Sinteza 2019, 2019. doi:10.15308/sinteza-2019-445-451.

[23] P. K. Bhanu, R. Kannan, S. A. Alexander, and G. R. Kanagachidambaresan, "Advanced deep learning for engineers and scientists," EAI/Springer Innovations in Communication and Computing, 2021. doi:10.1007/978-3-030-66519-7.

[24] Zhang, S.; Wen, L.; Bian, X.; Lei, Z.; Li, S. Single-shot refinement neural network for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4203–4212.

[25] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[26] A. Vaswani et al., "Attention is all you need," arXiv.org, https://arxiv.org/abs/1706.03762 (accessed Jan. 7, 2024).

[27] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," arXiv.org, https://arxiv.org/abs/2010.11929 (accessed Jan. 7, 2024).

[28] N. Tumanyan, O. Bar-Tal, S. Bagon, and T. Dekel, "Splicing VIT features for semantic appearance transfer," arXiv.org, https://arxiv.org/abs/2201.00424 (accessed Jan. 7, 2024).

[29] M. Caron et al., "Emerging properties in self-supervised Vision Transformers," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021. doi:10.1109/iccv48922.2021.00951.

[30] K. He et al., "Masked autoencoders are scalable vision learners," arXiv.org, https://arxiv.org/abs/2111.06377 (accessed Jan. 7, 2024).

[31] H. Bao, L. Dong, S. Piao, and F. Wei, "Beit: Bert pre-training of Image transformers," arXiv.org, https://arxiv.org/abs/2106.08254 (accessed Jan. 7, 2024).

[32] I. R. Syihad, M. Rizal, Z. Sari, and Y. Azhar, "CNN Method to Identify the Banana Plant Diseases based on Banana Leaf Images by Giving Models of ResNet50 and VGG-19" *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 6, pp. 1311-1318, 2023. [Online]. Available: https://doi.org/10.29207/resti.v7i6.5000.

[33] U. P. Khan and M. W. Anwar, "Cybersecurity in Pakistan: Regulations, Gaps and a Way Forward," Cyberpolitik Journal, vol. 5, no. 10, Winter 2020.

[34] European Union, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation)," Official Journal of the European Union, vol. 59, no. 1, pp. 1–88, May 4, 2016. [Online]. Available:https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

[35] DataGuidance, "Pakistan Data Protection Overview," DataGuidance. [Online]. Available: https://www.dataguidance.com/notes/pakistan-data-protection-overview. [Accessed: Month Day, Year].