

**T.R.  
ONDOKUZ MAYIS UNIVERSITY  
INSTITUTE OF GRADUATE STUDIES  
FACULTY OF AGRICULTURE  
DEPARTMENT OF AGRICULTURAL MACHINERY AND TECHNOLOGIES  
ENGINEERING**



**APPLICATION OF POTENTIAL DISEASE DETECTION TECHNIQUES IN  
GREENHOUSES AND THEIR IMPACT ON SUSTAINABILITY**

**PH.D. THESIS**

**Ako Kakarash OMER**

**SUPERVISOR**

**Prof. Dr. Y. Benal ÖZTEKİN**

**SAMSUN  
2024**

## ACCEPTANCE AND APPROVAL OF THE THESIS

The study entitled “**APPLICATION OF POTENTIAL DISEASE DETECTION TECHNIQUES IN GREENHOUSES AND THEIR IMPACT ON SUSTAINABILITY**” prepared by **Ako Kakarash OMER**, and supervised by **Prof. Dr. Y. Benal ÖZTEKİN** was found successful and unanimously accepted by committee members as a Ph.D. thesis, following the examination on the date 24.07.2024.

**Mr.Ako Kakarash OMER**

Ondokuz Mayıs University

**Department/** Agricultural Machinery and  
Technologies Engineering

**Final  
Decision**

|                 |                                                                                                                                          |                                                                               |
|-----------------|------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| <b>Chairman</b> | Prof. Dr. Ufuk Türker<br>Ankara University<br>Department of Agricultural Machinery<br>and Technologies Engineering                       | <input checked="" type="checkbox"/> Accept<br><input type="checkbox"/> Reject |
| <b>Member</b>   | Prof. Dr. Yeşim Benal Öztekin<br>Ondokuz Mayıs University<br>Department of Agricultural Machinery<br>and Technologies Engineering        | <input checked="" type="checkbox"/> Accept<br><input type="checkbox"/> Reject |
| <b>Member</b>   | Prof. Dr. Gürkan Alp Kağan Gürdil<br>Ondokuz Mayıs University<br>Department of Agricultural Machinery<br>and Technologies Engineering    | <input checked="" type="checkbox"/> Accept<br><input type="checkbox"/> Reject |
| <b>Member</b>   | Assoc. Prof. Dr. Kemal Çağatay Selvi<br>Ondokuz Mayıs University<br>Department of Agricultural Machinery<br>and Technologies Engineering | <input checked="" type="checkbox"/> Accept<br><input type="checkbox"/> Reject |
| <b>Member</b>   | Assoc. Prof. Dr. Hüseyin Kürşat Çelik<br>Akdeniz University<br>Department of Agricultural Machinery<br>and Technologies Engineering      | <input checked="" type="checkbox"/> Accept<br><input type="checkbox"/> Reject |

This thesis has been approved by the committee members already stated above and determined by the Institute Executive Board.

Prof. Dr. Ahmet Tabak  
Head of Institute of Graduate Studies

## DECLARATION OF COMPLIANCE WITH SCIENTIFIC ETHIC

I hereby declare and undertake that I complied with scientific ethics and academic rules in all stages of my Ph. D. Thesis, that I have referred to each quotation that I use directly or indirectly in the study, and that the works I have used consist of those shown in the sources, that it was written in accordance with the institute writing guide and that the situations stated in the article 3, section 9 of the Regulation for TÜBİTAK Research and Publication Ethics Board were not violated.

Is the Ethics Committee Necessary?

Yes ☐ (If it is necessary, please add appendices.)

No ☒

Signature

24/07/2024

Ako Kakarash OMER

## DECLARATION OF THE THESIS STUDY ORIGINALITY REPORT

**Thesis Title:** APPLICATION OF POTENTIAL DISEASE DETECTION  
TECHNIQUES IN GREENHOUSES AND THEIR IMPACT ON  
SUSTAINABILITY

As a result of the originality report taken by me from the plagiarism detection program on 25 /06 /2024 for the title above;

Similarity ratio: % 30

Single resource rate: % 2 has been released.

Signature

24/07/ 2024

Prof. Dr. Y. Benal ÖZTEKİN

## ÖZET

### SERALARDA POTANSİYEL HASTALIK TESPİT TEKNİKLERİNİN UYGULANMASI VE SÜRDÜRÜLEBİLİRLİK ÜZERİNDEKİ ETKİLERİ

Ako Kakarash OMER

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Ziraat Fakültesi

Tarım Makinaları ve Teknolojileri Mühendisliği Ana Bilim Dalı

Doktora, Temmuz/2024

Prof. Dr. Y. Benal ÖZTEKİN

Tarımda yapay zeka kullanımı, ürün verimini artırmak, hastalık ve zararlılar nedeniyle %40'lara kadar ulaşan ürün kayıplarını en aza indirmek için zorunlu hale gelmiştir. Bu bağlamda, yapraktan bitki hastalığı tespiti odaklanılan önemli bir araştırma alanı haline gelmiştir. Farklı araştırmacılar, yapay zeka teknikleri kullanarak yaprak virüsü tanımlaması için aktif çözümler aramaktadırlar. Bu çalışmanın öncelikli amacı, CNN modelleri ile, hastalık seraya yayılmadan önce sınırlı veri mevcudiyeti ile domates bitkisinde bazı hastalık belirtilerini erken dönemde yapraktan tespit edebilen güvenilir bir yapay zeka modeli oluşturmaktır. Bu amaçla öncelikle, Altıayak-Antalya'da yer alan iki farklı seradan sağlıklı ve üç farklı tür hastalığa sahip domates bitkisi yaprak görüntüleri bir cep telefonu kamerası yardımıyla alınarak bu görüntülerden seçim yapılmış ve görüntüler etiketlenmiştir. Sınırlı sayıda elde edilen görüntülere veri artırma metodu uygulanmış ve bunlar en yaygın kullanılan beş farklı CNN mimarisi (VGG16, VGG19, InceptionV3, DenseNet201 ve MobileNet) ile önceden eğitilmiş model ağırlıklarından yararlanmak için transfer öğrenmesi kullanılarak eğitilmiş ve sonuçta eşit sayıda epok ve hiperparametre için sıfırdan yeni bir CNN modeli elde edilmiştir.

Modellerin sınıflandırma doğrulukları VGG16, VGG19, InceptionV3, MobileNet, DenseNet201 ve yeni elde edilen CNN modeli için sırasıyla %89.23, %83.5, %87.88, %94.28, %90.91 ve %83.0 olarak bulunmuştur. Elde edilen sonuçlar, sınırlı verilerle bile olsa sera ortamlarında hastalık tespiti için gelişmiş CNN mimarileri kullanımının uygulanabilirliğini ortaya koymuştur. Tüm modeller arasında MobileNet, %94.28'lik bir sınıflandırma doğruluğu ile diğer tüm modellere göre daha başarılı olmuştur. Bu sonuçlar MobileNet modelinin domates hastalıklarını doğru bir şekilde tespit etmedeki potansiyelini ortaya koymaktadır. Etkili mimarisi, özel düzenlemesi ve veriye uygunluğu göz önüne alındığında en yüksek doğruluk elde etme başarısına sahip mimarinin MobileNet olduğu söylenebilir. Sonuç olarak, kapsam ve uygulama alanının genişletilmesi için kamu kurumları, tarımsal araştırma merkezleri ve işletmelerin işbirliği yaparak daha fazla veri kullanımı ile daha fazla sınıfa yönelik gerçek zamanlı sera izleme sistemleri konusunda araştırmalar yapılması önerilmektedir.

**Anahtar sözcükler:** Sera, Yapay sinir ağları, Derin öğrenme, Evrişimli sinir ağları, Görüntü işleme, Hastalık tespiti

## ABSTRACT

### APPLICATION OF POTENTIAL DISEASE DETECTION TECHNIQUES IN GREENHOUSES AND THEIR IMPACT ON SUSTAINABILITY

Ako Kakarash OMER

Ondokuz Mayıs University

Institute of Graduate Studies

Faculty of Agriculture

Department of Agricultural Machinery and Technologies Engineering

Ph.D., July/2024

Supervisor: Prof. Dr. Y.Benal ÖZTEKİN

Artificial intelligence has become imperative in agriculture, improving production efficiency and minimizing losses, which reach up to 40% due to pests and diseases. Thus, early plant leaf disease detection has become a vital area of focus. Researchers have been actively searching for solutions for vegetable leaf virus identification. The primary objective of this study is to create a reliable artificial intelligence model, particularly for CNN models, that could identify tomato leaf infections early even with limited data available before the diseases spread across the greenhouse. Images of healthy and three types of diseased tomato leaves were captured, collected, selected and labeled from two greenhouses in Altıayak, Antalya-Türkiye with a mobile phone camera. Despite having limited data, employing data augmentation, five popular CNN architectures (VGG16, VGG19, InceptionV3, DenseNet201, and MobileNet) were trained using transfer learning to take advantage of the weights of pre-trained models and a new CNN model was trained from scratch for equal number of epochs and hyperparameters.

The classification accuracies of the models were 89.23%, 83.5%, 87.88%, 94.28%, 90.91% and 83.0% for VGG16, VGG19, InceptionV3, MobileNet, DenseNet201 and the new CNN model respectively. Therefore, the findings of this research demonstrated the feasibility of using advanced CNN architectures for disease detection in greenhouse environments even with limited data. Out of all the models, MobileNet outperformed all other models with a classification accuracy of 94.28%. These results highlighted the potential of the MobileNet models in accurately detecting tomato diseases. It can be concluded that its efficient architecture, specific regularizations, and suitability, as obtained from the particular data, have contributed to the overall success of the MobileNet CNN model. The study suggests that further research should be conducted in real-time greenhouse monitoring systems using a larger amount of data, for an increased number of classes with cooperation of government bodies, agricultural research centers and business enterprises to broaden its scope and implementation.

**Keywords:** Greenhouse, Artificial neural networks, Deep learning, Convolutional neural networks, Image processing, Disease detection

## ACKNOWLEDGEMENT

I want to convey my sincere gratitude to the most revered, illustrious, and devoted supervisor, Prof. Dr. Y. Benal ÖZTEKİN, for her understanding, counsel, and knowledge that she shared with me. I wish to extend my genuine appreciation to the members of their PhD thesis committee, Prof. Dr. Ufuk TÜRKER from Ankara University, Prof. Dr. Gürkan Alp Kağan GÜRDİL, and Assoc. Prof. Dr. Kemal Çağatay SELVI from Ondokuz Mayıs University and, Assoc. Prof. Dr. Hüseyin Kürşat ÇELİK from Akdeniz University for their insightful feedback and constructive criticism that greatly enhanced the thesis.

Besides, I would like to express my thankfulness, especially to PhD student Mr. Ayas HUSSEIN, who has been a valuable asset in this journey. I want to thank my fellow student Mr. Khalid D. IDREES for his counsel and recommendations. Also, I thank Mr. Cihat GEDİK, and Mr. Mahtem T. MENGSTU, Ph.D. students for their advice and moral support and I am thankful to all students of my department. I thank Türkiye Scholarships, Ondokuz Mayıs University, and Koya University and its president, particularly the engineering faculty and the manufacturing and Industrial department, for their unwavering support and faith. I do not want to forget to send huge thanks to the thesis committee members as they recommended me in terms of knowledge and support at the early stage of the journey. Moreover, it is a great idea to thank the head and department members for always asking and supporting my studies. Lastly, I would like to thank my family for their support in this long journey.

July, 2024

Ako Kakarash OMER

## CONTENTS

|                                                                  |           |
|------------------------------------------------------------------|-----------|
| ACCEPTANCE AND APPROVAL OF THE THESIS.....                       | i         |
| DECLARATION OF COMPLIANCE WITH SCIENTIFIC ETHIC .....            | ii        |
| DECLARATION OF THE THESIS STUDY ORIGINALITY REPORT .....         | ii        |
| ÖZET .....                                                       | iii       |
| ABSTRACT .....                                                   | iv        |
| ACKNOWLEDGEMENT.....                                             | v         |
| CONTENTS .....                                                   | vi        |
| SYMBOLS AND ABBREVIATIONS.....                                   | vii       |
| FIGURES LEGENDS.....                                             | viii      |
| TABLES LEGENDS.....                                              | ix        |
| <b>1. INTRODUCTION .....</b>                                     | <b>1</b>  |
| 1.1. Research Background .....                                   | 1         |
| 1.2. Agricultural Production.....                                | 2         |
| 1.2.1. Recent technology in greenhouse .....                     | 3         |
| 1.2.2. Greenhouse distribution in Türkiye and the globe .....    | 4         |
| 1.2.3. Permanent problem of pesticide.....                       | 6         |
| 1.2.4. Tomatoes production in the world .....                    | 6         |
| 1.2.5. Tomato production in Türkiye .....                        | 7         |
| 1.2.6. Cost of pesticide.....                                    | 8         |
| 1.2.7. Effective method of treating tomato early blight .....    | 8         |
| 1.2.8. Early blight(Alternaria).....                             | 8         |
| 1.2.9. Powdery mildew .....                                      | 9         |
| 1.2.10. Tomato yellow leaf curl virus .....                      | 10        |
| 1.3. Artificial Intelligence .....                               | 10        |
| 1.3.1. Significance of computer vision-based detection .....     | 13        |
| 1.3.2. Deep learning techniques in agricultural greenhouse ..... | 15        |
| 1.3.3. CNN Models .....                                          | 15        |
| 1.3.4. Artificial neural networks and greenhouses.....           | 16        |
| 1.3.5. Convolutional neural networks .....                       | 17        |
| 1.3.6. The importance of CNN.....                                | 18        |
| 1.3.7. Definition of CNN .....                                   | 19        |
| 1.3.8. Smart innovations and applications of greenhouse .....    | 20        |
| 1.4 Aim of the Study .....                                       | 25        |
| 1.4.1. Project scope .....                                       | 26        |
| 1.4.2. Unique values of this research .....                      | 27        |
| <b>2. LITERATURE REVIEW.....</b>                                 | <b>27</b> |
| 2.1. General Literature Review .....                             | 27        |
| 2.2. Transfer Learning on Plant Disease Detection .....          | 40        |
| <b>3. MATERIALS AND METHODS.....</b>                             | <b>44</b> |
| 3.1. Image Acquisition .....                                     | 44        |
| 3.2. Image Preprocessing.....                                    | 45        |
| 3.3. Transfer Learning Models .....                              | 46        |
| 3.3.1. VGG.....                                                  | 48        |
| 3.3.2. InceptionV3 .....                                         | 49        |
| 3.3.3. MobileNet .....                                           | 50        |
| 3.3.4. DenseNet201 .....                                         | 50        |
| 3.3.5. Proposed CNN model .....                                  | 51        |
| 3.4 Performance Metrics.....                                     | 52        |

|                                                |           |
|------------------------------------------------|-----------|
| <b>4. RESULTS AND DISCUSSIONS.....</b>         | <b>53</b> |
| 4.1. Results of the Training .....             | 53        |
| 4.1.1. VGG16.....                              | 53        |
| 4.1.2. VGG19.....                              | 55        |
| 4.1.3. InceptionV3 .....                       | 57        |
| 4.1.4. MobileNet.....                          | 59        |
| 4.1.5. DenseNet201.....                        | 61        |
| 4.1.6. Proposed CNN model .....                | 63        |
| 4.2. Discussion.....                           | 65        |
| <b>5. CONCLUSIONS AND RECOMMENDATIONS.....</b> | <b>68</b> |
| <b>REFERENCES.....</b>                         | <b>72</b> |
| <b>APPENDIX.....</b>                           | <b>83</b> |
| <b>CURRICULUM VITAE.....</b>                   | <b>86</b> |



## SYMBOLS AND ABBREVIATIONS

|        |                                     |
|--------|-------------------------------------|
| AI     | : Artificial Intelligence           |
| ANN    | : Artificial Neural Network         |
| CNN    | : Convolutional Neural Network      |
| DL     | : Deep Learning                     |
| EU     | : European Union                    |
| FAO    | : Food and Agriculture Organization |
| GAN    | : Generative Adversarial Networks   |
| IoT    | : Internet of Things                |
| IPM    | : Integrated Pest Management        |
| LFM    | : Leaf Foreground Module            |
| ML     | : Machine Learning                  |
| SF     | : Smart Farming                     |
| SVM    | : Support Vector Machines           |
| TÜİK   | : Turkish Statistical Institute     |
| USA    | : United States of America          |
| VGG-16 | : Visual Geometry Group-16          |
| VGG-19 | : Visual Geometry Group- 19         |

## FIGURE LEGENDS

|                                                                                               |    |
|-----------------------------------------------------------------------------------------------|----|
| Figure 1.1. Early blight.....                                                                 | 9  |
| Figure 1.2. Tomato powdery mildew.....                                                        | 9  |
| Figure 1.3. Tomato yellow leaf.....                                                           | 10 |
| Figure 1.4. The illustration manual disease detection.....                                    | 12 |
| Figure 1.5. Camera vision-based and CNN for diagnosing and clustering.....                    | 13 |
| Figure 1.6. Understanding of AI concept.....                                                  | 14 |
| Figure 1.7. Schematic of ANN (Maind, and Wankar, 2014).....                                   | 16 |
| Figure 1.8. Schematic of a CNN with two hidden layers (Dalai, 2020).....                      | 20 |
| Figure 1.9. The potential of smart agriculture (Bacco, et al., 2019).....                     | 21 |
| Figure 1.10. Demonstrates the parts of a plant that face diseases mostly.....                 | 22 |
| Figure 3.1. Sample images of the four classes.....                                            | 44 |
| Figure 3.2. Original and resized images.....                                                  | 45 |
| Figure 3.3. (a) Original image, (b) 15 <sup>0</sup> rotated and (c) horizontally flipped..... | 46 |
| Figure 3.4. Six categories of images removed from the dataset.....                            | 46 |
| Figure 3.5. Typical CNN architecture.....                                                     | 47 |
| Figure 3.6. VGG16 architecture.....                                                           | 48 |
| Figure 3.7. VGG19 architecture.....                                                           | 49 |
| Figure 3.8. InceptionV3 architecture.....                                                     | 49 |
| Figure 3.9. MobileNet architecture.....                                                       | 50 |
| Figure 3.10. DensNet201 architecture.....                                                     | 51 |
| Figure 3.11. The architecture of the proposed CNN model.....                                  | 52 |
| Figure 3.12. The overall process of tomato disease detection model.....                       | 52 |
| Figure 4.1. Training and validation accuracy curve of VGG16.....                              | 53 |
| Figure 4.2. Training and validation loss curve of VGG16.....                                  | 53 |
| Figure 4.3. Confusion matrix of VGG16.....                                                    | 54 |
| Figure 4.4. Training and validation accuracy Curve of VGG19.....                              | 55 |
| Figure 4.5. Training and validation loss of VGG19.....                                        | 55 |
| Figure 4.6. Confusion matrix of VGG19.....                                                    | 56 |
| Figure 4.7. Training and validation accuracy curve of inceptionV3.....                        | 57 |
| Figure 4.8. Training and validation loss curve of the inceptionV3 model.....                  | 57 |
| Figure 4.9. Confusion matrix of inceptionV3.....                                              | 58 |
| Figure 4.10. Training and validation accuracy curve of MobileNet model.....                   | 59 |
| Figure 4.11. Training and validation loss of mobileNet model.....                             | 59 |
| Figure 4.12. Confusion matrix of MobileNet.....                                               | 60 |
| Figure 4.13. Training and validation accuracy curve of the denseNet201 model.....             | 61 |
| Figure 4.14. Training and validation loss curve of the denseNet201 model.....                 | 61 |
| Figure 4.15. Confusion matrix of denseNet201.....                                             | 62 |
| Figure 4.16. Training and validation accuracy curve of the proposed model.....                | 63 |
| Figure 4.17. Training and validation loss curve of the proposed model.....                    | 63 |
| Figure 4.18. Confusion matrix of the proposed model.....                                      | 64 |

## TABLES LEGENDS

|                                                                                      |    |
|--------------------------------------------------------------------------------------|----|
| Table 1.1. Types of greenhouse structure.....                                        | 5  |
| Table 1.2. Top five world's largest tomato-producing countries in 2018.....          | 7  |
| Table 3.1. Proportion of training, validation and test datasets in four classes..... | 45 |
| Table 3.2. Training configurations use.....                                          | 53 |
| Table 4.1. Classification report for VGG16 accuracy 89.23%.....                      | 55 |
| Table 4.2. Classification report for VGG19, accuracy 87.88%.....                     | 57 |
| Table 4.3. Classification report for inceptionV3, accuracy 83.5%.....                | 59 |
| Table 4.4. Classification report for mobileNet, accuracy 94.28%.....                 | 61 |
| Table 4.5. Classification report for denseNet201 accuracy 90.91%.....                | 63 |
| Table 4.6. Classification report for the proposed model, accuracy 83.50%.....        | 65 |
| Table 4.7. Performance metrics of all models.....                                    | 66 |
| Table 4.8. Trending results of disease detection algorithms .....                    | 70 |



# **1. INTRODUCTION**

## **1.1. Research Background**

Technological progress has significantly changed many aspects of our daily lives in the past few decades. The obvious changes can be seen in business, education, agriculture and healthcare industries. These areas have all gained benefits from this advancement. Technology use has been fundamental because it is not imaginable to perform daily tasks effectively and reliably without it. The world has witnessed numerous technological innovations that have driven rapid changes. Artificial Intelligence (AI) stands out as a considerable advancement in the information technology sector (Gupta, 2021). This research seeks to examine the agriculture sector, which is dramatically influenced by AI innovation. This study will later explore several reasons that will be examined to explain why AI has become important in agriculture.

Over the past 23 years, the population has globally risen by 1.7 billion, and it is expected to grow by an additional 2 billion in the next 28 years based on the United Nations report (United Nations, 2022). This rapid population growth creates serious challenges in ensuring food security particularly, as addressed by Roser et al. (2019). To state these challenges, food production should increase by 60% over the next 30 years, a target highlighted by Ristaino et al. (2021). Improving crop production efficiency and minimizing losses, which can reach up to 40% because of pests and diseases. According to the Food and Agriculture Organization, these are crucial strategies (FAO, 2021). Therefore, the detection of crop diseases becomes a vital area of focus. The main approaches to detecting crop diseases are done manually through which labor-intensive is required, and exposed to errors (Thangaraj et al., 2022a).

The research community has been actively seeking automated solutions for plant disease detection, which has seen increasing scholarly attention. There have been notable contributions to a new deep-learning method for diagnosing rice diseases (Lu et al., 2017). Also, there has been a system developed by DeChant et al. (2017) for identifying northern leaf blight in maize. Although various studies demonstrate the high effectiveness of automated plant disease detection methods, particularly those using Machine Learning (ML) and Deep Learning (DL) techniques, challenges remain. Studies often report high accuracy rates for Convolutional Neural Networks (CNN) models, some even claiming 100% accuracy (Ayikpa et al., 2022). However, these

models are frequently trained and evaluated using unrealistic datasets comprising laboratory-generated images, limited sample sizes, or heavily processed images, which limits their practical applicability (Abade et al., 2021a; Ahmed et al., 2021; Thangaraj et al., 2022; Tugrul et al., 2022).

This research aims to address these gaps by building robust CNN architecture solutions that train methods in enhancing the accuracy of plant disease diagnosis and contribute to more sustainable agricultural practices to improve food security. Today with the new growth of the economy, the agriculture sector is transforming and progressing from the traditional to a modern cycle, and greenhouses, as crucial support for modern agriculture, are widely used. A greenhouse can be kept untouched and affected by the external environment and at the same time not be limited by the geographical and seasonal factors of each crop's cultivation, consequently showing the capabilities of meticulous environmental agricultural production (Xie, et al., 2017). According to Strange et al. (2005), global food security is extremely threatened by plant diseases. On the other hand, traditional methods for diagnosing these plant diseases rely on visual inspection laboratory tests which significantly require time and resources (Khakimov et al., 2022). In addition, these methods often fail to provide good results, notably in remote and poor regions. CNNs have accomplished considerable success in image recognition tasks. They are being highly applied in agricultural areas because they are capable of identifying and classifying plant diseases from digital images. This method offers many advantages over traditional methods due to improved accuracy and speed, cost reduction and scalable solutions for different geographic areas (Olaoye and Potter, 2024).

## **1.2. Agricultural Production**

Agriculture is not only essential for enhancing nutrition, but it is also a major source of cash for many nations. Increased food output is the key to reducing hunger and fostering social and economic advancement. The last century has significantly changed the world's crop production. The quantity of land we utilize for agriculture has increased, but the main driver has been a quick rise in crop yields, which have caused a significant increase in the amount of food we produce. In many nations around the world, there is a greater variety of diets. The majority of agricultural produce used to be made up of cereals, roots, and other staple crops.

Now, this includes foods such as nuts, seeds, fruits, vegetables, legumes and other nutrition. International trade has also significantly increased agricultural productivity. Everywhere in the past, food production was localized; farmers produced food for their households or local groups. Even as it grew more widely, domestic markets accounted for the majority of the food trade. Large-scale international food trade now exists; this has greatly increased the variety of foods around the world and can be a significant source of money for farmers.

#### **1.2.1. Recent technology in greenhouse**

In the year 2018, globally, the total area covered by greenhouse vegetables was approximately 5.6 million ha (Roble, 2019). Due to their extensive energy requirements, most greenhouses have substantial operational expenses. The considerable heat loss causes the high energy consumption in greenhouses due to the greenhouse outer material. The most recent technical advancements are employed in greenhouses throughout many nations to regulate the microclimate inside the greenhouse with a focus on passive methods. It has been discovered that PCM (Privacy-preserving model customization) was recently employed to offer cooling and heating for a Mediterranean climate. Additionally, a system based on the closed greenhouse concept for northern climates enhances energy demand reduction by 80%, with a potential payback of 6 years. Additionally, double-glazing envelopes are usually the most effective for improving greenhouse performance in the majority of countries (Harjunowibowo et al., 2018).

Greenhouse technology is employed to experience quick growth. According to a study by Briassoulis et al. (1997) the most basic greenhouse uses a plastic cover over wooden stands. Furthermore, traditional Chinese greenhouses are heated with the help of coal burners, natural gas, and the combustion of straw. After that, the modern greenhouse's cover is made of glass and steel (Carlini et al., 2012). These are designed to provide the best circumstances possible for plant growth (Pérez-Alonso et al., 2012) and pest prevention (Yang et al., 2014). Despite the main functions of these technologies, there is a gap in terms of pesticide control automatically farmers still struggling with the prevention of plant leaf diseases Therefore, with modern greenhouses and robotization cultivation operations plant protection also has

significance to plant growers in terms of knowing the disease and early protection of spread and control the diseases.

### **1.2.2. Greenhouse distribution in Türkiye and the globe**

In the world plastic films have changed the agricultural industry since they first entered it, greatly enhancing both the quality and quantity of crops. In recent years, agricultural films have attracted not only a lot of interest and attention but also a sizable industry that is expanding across all of Europe, America, and Asia, with a persistently good trend that is crisis-free. The term "plasticulture" refers to the field of plastic films used in agriculture. These films include photo-selective films, solarization films, low tunnels, greenhouse coverings, mulches, and impermeable films. It consists of raw materials (many different types of polymers), various film forms, applications, agronomical performances, and recycling issues. As a result, using future automation systems brings many benefits. An accurate definition of a greenhouse described by Von Zabeltitz (2011) is that for the plants growing within greenhouses, the climate must be ideal. They must shelter plants against extreme cold, wind, rain, hail, birds, and insects (Castellano et al., 2006). Shape, orientation, structure, cladding material, base, ventilation, and technological temperature control equipment are all crucial elements in a successful greenhouse design.

Specifically, in Türkiye, the economy heavily depends on the horticulture sector. Numerous fruits and vegetables can be grown in Türkiye due to their suitable environments and climatologic conditions. Besides, Türkiye provides various horticulture products because of the limitless climatic and soil conditions throughout a large territory (Yilmaz et al., 2005). At research facilities in Türkiye, greenhouse cultivation began hesitantly in the 1940s. In 1970, commercial greenhouses were built in Izmir and Antalya (Kutlar, 2019). The total agricultural land in Türkiye is 23.4 million hectares. The average size of agricultural land sections is 1.29 hectares, and there are 5.9 agricultural land sections per firm (TÜİK, 2019). One of Türkiye's major agricultural and tourist centers is Antalya. By population, it is the fifth-largest province. Its agricultural output is worth 9.5 billion Turkish lira. In terms of value, it represents 6.8% of Türkiye's vegetable production. Approximately 156,000 growers reside in the province.

The greenhouse business appears to be Türkiye's fastest-growing sector of agriculture at the moment, mostly due to the country's favorable climatic circumstances. However, in recent years, businesses have been compelled to take on a more competitive position in the market. Growers may not always receive any guarantees concerning sales volume, a fair price, or a positive financial outcome due to the competitive market environment for greenhouse products. Operators of greenhouses in Türkiye currently deal with issues like diminishing crop prices, price swings caused by an excess of supply, weak market structures and sales uncertainty, and a lack of farmer cooperatives. Operators of greenhouses now face unclear income and market concerns as a result of these issues. Other flaws in the Turkish greenhouse business include its heavy reliance on foreign inputs and overuse of chemicals. Greenhouses in the Türkiye area gradually expanded from 1003 ha in 1960 to 24 458 ha in 2002, and 650321 ha in 2018. The annual growth rate shows that up until 1990, there was a gradual decline in expansion, but recently, the rate of growth has increased. Türkiye produced 1.37 million tons of tomatoes in 2000 however now it has rapidly increased and recently in 2023 the production of tomatoes has dramatically increased to 13.5 million tons (TUIK, 2023).

Table 1.1. Types of greenhouse structure

| Greenhouses Types              | Structure Design                            |
|--------------------------------|---------------------------------------------|
| A. Based On Cost Investment    | 1. Low technology greenhouses               |
|                                | 2. Medium technology greenhouse             |
|                                | 3. High-level greenhouses                   |
|                                | 4. Lean-to type greenhouse                  |
| B. Based On Shape              | 1. Uneven span-type greenhouse              |
|                                | 2. Ridge and furrow-type greenhouse         |
|                                | 3. Quonset greenhouse                       |
| C. Based On Utility            | 1. Saw tooth-type greenhouse                |
|                                | 2. The active heating system of greenhouses |
|                                | 3. The active cooling system                |
| D. Based On Covering Materials | 1. Glass greenhouses                        |
|                                | 2. Plastic film greenhouses                 |
|                                | 3. Rigid panel greenhouses                  |
| E. Based On Construction       | 1. Wooden framed structures                 |
|                                | 2. Truss framed structures                  |

### **1.2.3. Permanent problem of pesticide**

Excessive use of fertilizers and pesticides, fertilizers, and other chemicals are the main inputs utilized in greenhouse cultivation. Some of these inputs might be harmful to human health and might pollute the environment. Recent research from the province of Antalya indicates that excessive amounts of pesticides and fertilizers are used in greenhouse agriculture. According to Yilmaz et al. (2005), just 7% of growers administered pesticides in proportion to the pest's economic threshold, while 69% of growers did so as soon as they noticed pests in their greenhouse. A third of those polled (32%), on average, said they use pesticides more frequently than is advised. According to Ozkan (2000), excessive pesticide use occurs during the production of greenhouse tomatoes, peppers, and cucumbers. These costs account for 25% to 55% of the overall production costs. Econometric studies reveal that pesticide use is incompatible with the technical optimum point in several greenhouse crops when excess use is seen from an economic perspective (Yilmaz et al., 2005). Future export potential will be significantly impacted by pesticide residue on the crops. Additionally, growers overuse fertilizer when producing in greenhouses, which causes soil salinity issues and nitrate poisoning of aquifers. Salinity issues arise in greenhouse soils, and a study in the region of Antalya revealed that 50% of wells had nitrate levels that are higher than what is permissible (Kaplan, 2000).

### **1.2.4. Tomatoes production in the world**

According to data compiled by FAO (2022), the total amount of tomatoes produced worldwide in 2021 for processing and fresh consumption was just over 189.1 million metric tons, up 2% from the 184.8 million metric tons grown in 2020 and 4% from the average (182.7 million metric tons) of the preceding three years (2018-2020),.

These numbers show that in several nations, annual tomato output either surpassed or beyond the one million metric ton mark. Three times less than Iran (3.4 million metric tons) and six times less than Italy (6.6 million metric tons), Cameroon and Indonesia combined generated about 1.1 million metric tons. Türkiye produced twice as much as Italy (13 million metric tons), but barely a fifth of China's crop (67 million metric tons), which is responsible for about 36% of the world harvest.

Record yields of fresh tomatoes were observed in Poland and the Netherlands, even though the growing conditions were substantially different from those used to process tomatoes. After Spain, Türkiye, and Brazil, the nations with the largest yields were Portugal, the United States, and Morocco. At 36.6 t/ha, the yield worldwide was nearly unchanged from 2020.

FAO data shows that there is a large difference in the growth rates of the two sectors, with fresh market tomato production having the advantage. Over the previous 30 years (1991–2021), the output of raw tomatoes has increased at an average annual rate of 3.5%, whereas the production of processed tomatoes has increased at an average annual rate of only 1.85%. For processed tomatoes, growth rates have been 2.6%, whereas fresh tomato growth rates have been above 3% during a shorter period, from 2001 to 2021. Over the previous ten years (2011–2021), fresh market production has increased at a rate of more than 2%, but the processing sector has only had marginally positive growth at an annual rate of 0.4%. Table 1.2. shows the top five tomato producers.

Table 1.2. Top five world's largest tomato-producing countries in 2021

| Country | Production (mT) |
|---------|-----------------|
| China   | 67,538,340      |
| India   | 21,181,000      |
| Türkiye | 13,095,258      |
| USA     | 10,475,265      |
| Italy   | 6,644,790       |

#### 1.2.5. Tomato production in Türkiye

One of the most popular vegetable food sources is the tomato. Türkiye has a per capita yearly fresh tomato consumption of approximately 80kg due to favorable ecological circumstances, tomatoes' high vitamin and mineral content, their significance in dietary habits, and their relatively low cost. According to Keskin et al. (2010), this amount is three to four times higher than in the EU and several other producer nations. China was the leading tomato producer in 2007 with a share of 26%, followed by the EU, the USA, and Türkiye with shares of 13%, 9%, and 8%, respectively (Anonymous, 2008). While tomato production is practiced in nearly every province of Türkiye, it is typically done in the Mediterranean region under greenhouse conditions. The production of industrial tomatoes differs regionally as well; typically, these productivities take place in the Aegean and The Ministry of Agriculture and Rural

Affairs Marmara areas (Keskin et al., 2010). The Ministry of Agriculture and Rural Affairs (MARA) has certified seeds for 60 different industrial tomato varieties and 400 different table tomato varieties. 90% of the tomatoes processed in Türkiye's tomato industry are used to make tomato paste, with the other 10% going to pickles, dried tomatoes, and other tomato-based goods. The EU and the USA are significant destinations for Türkiye's exports of organic tomatoes, which are produced for exportation together with processed goods. Following the global trend of tomato production Türkiye has provided a governmental strategy to increase productivity based on the customer's demand and having a good geographical region and climate and policies to support farmers therefore Ministry of Agriculture wants to pay attention to using technology and innovation to increase that productivity.

Due to its favorable temperature and usage of greenhouses for production, Turkey's Mediterranean region has the best productivity. Turkey produces roughly 40 tons of tomatoes per hectare on average, which is much less than the needed amount even though it is nearly 50% higher than the global average. Italy produces 56 tons of tomatoes per hectare on average, Greece produces 61 tons, Morocco produces 78 tons, Israel produces 78 tons, and Spain produces 82 tons per hectare. These are the countries with similar climates (Özaydin and Direk, 2022).

#### **1.2.6. Cost of pesticide**

The cost of pesticides in greenhouses for tomato plants can vary depending on several factors, such as the type and brand of pesticide being used, the size of the greenhouse, and the severity of the pest infestation. According to (Engindeniz and Cosar 2013) average pesticide and pesticide application costs of the processing and table tomatoes were determined to be \$445/ha and \$502/ha, respectively. The breakeven yields for the processing and table tomatoes were calculated to be 3,708 and 3,138 kg/ha, respectively. It is important to keep in mind that while pesticides can be effective in controlling pests in tomato plants, they also have the potential to harm beneficial insects and impact the environment. Integrated Pest Management (IPM) strategies which incorporate cultural, biological, and physical methods in conjunction with pesticide use can help to minimize the overall cost of pest management and reduce negative impacts.

### 1.2.7. Effective method of treating tomato early blight

Treatment of early blight on tomatoes needs to be treated right away to prevent the diseases from taking over the plants. Spray bonide tomato and vegetable or bonide liquid copper fungicide concentrate liberally all over the plant, including the bottoms of the leaves. These two remedies are both natural. Powdery mildew, downy mildew, black spots, peach leaf curl, rust, and other diseases are all controlled by liquid copper fungicide concentrate.

### 1.2.8. Early blight (*Alternaria*)

Early Blight is brought on by the fungus *Alternaria*, which also affects tomato plants. Lower leaves have dark-edged brown or black dots that resemble targets. Fruit stem tips may be harmed, displaying significant sunken, dark patches with concentric rings. Usually, this fungus appears after plants produce fruit. Lower tomato leaves are produced by tomatoes. Early blight is a fungus that causes yellow leaves with brown blotches. Perhaps the most prevalent tomato foliar disease is *Alternaria Solanica* reported by Nashwa and Abo-Elyousr, (2012).



Figure 1.1. Early blight

### 1.2.9. Powdery mildew

As its name suggests, powdery mildew spores on leaves make them look white on the upper or lower surfaces. You might also notice spots of dead tissue on the lower leaves, which may be surrounded by a yellow.



Figure 1.2. Powdery mildew

*Oidiopsis taurica* fungi do not infect stems and fruits. Spores are usually common in areas near farms and fields since the fungus commonly infects crops and weeds in the Solanaceae family. The wind carries these spores to tomato plants in home gardens. This disease is usually a problem late in the season and in humid conditions, with temperatures between 15-21°C. Several infections may be treated with fungicides, or the bio fungicide *Bacillus pumilus*, and some types of organic sprays, such as neem oil, can also be used.

#### **1.2.10. Tomato yellow leaf curl virus**

Tomato yellow leaf curl virus (TYLCV) can infect over 30 different kinds of plants, but it is mainly known to cause devastating losses of up to 100 percent in the yield of tomatoes. Both field and glasshouse-grown tomatoes are susceptible. TYLCV is vectored by silverleaf whitefly (*Bemisia tabaci*).



Figure 1.3. Tomato yellow leaf

### **1.3 . Artificial Intelligence**

According to a study conducted by Escamilla-García et al. (2020), AI is the study and development of intelligent software and devices which includes disciplines like genetic algorithms, particle swarm optimization, simulation, ANNs and hybrid models. AI involves mapping non-linear behavior between process inputs and outputs. With functions like perception, recognition, decision-making, and control combined with knowledge from the brain and related disciplines like cognitive science and psychology, AI consists of many useful tools that enable solving challenging problems and tasks that call for biological or human intelligence (Escamilla- García et al., 2020).

Many scholars and researchers state that non-destructive recognition and early identification of vegetable growth disorders are vital to developing precision and ecological agriculture. Primary detection and prevention can efficiently slow down the spread of plant growth disorders. Hence, it requests to adopt a proper procedure for precise detection (Wang and Liu, 2021).

Among various crops, the tomato is the first vegetable that can be eaten as fruit, and its nutritional rate is much higher than that of fruit (Manish, 2017). Its yield is high, and its planting area is increasing, exclusively in greenhouses where the planting area increases rapidly (Takayama and Ezura, 2015). However, tomatoes are susceptible to diseases and pests during their growth, which extremely affects their produce and quality and sources enormous economic losses to planters (Xu, et al., 2000). The occurrence is uncertain due to the variety of tomato diseases and pests, the types of chemical sprays are diverse and the cost of treatment is often higher than that of prevention which negatively brings notorious economic conditions (Fuentes et al., 2016).

Pesticides not only destroy the plants' products but also destroy the ecology and may damage the environment and human well-being. Therefore, any attempts such as early disease detection and adequate pesticide application to reduce pesticides will be accepted, and lower application of pesticides means lower expenditure.

There are two types of plant disease detection which are the traditional method and the systematic model. Firstly, traditional identification and early warning decisions of tomato diseases and pests are largely based on field surveys by plant protection specialists and experienced tomato producers according to the growth status of tomatoes

(Martinelli, et al., 2015). Another point is crucial only relying on human being decision resources to identify the evidence on plant growth disorders is labor-intensive, time-consuming, inaccurate, and slow, and deviations may occur, in this case, hindering the timely treatment of diseases and pests (Czosnek and Laterrot, 2020). Therefore, today's agricultural production needs a new system to release producers from inefficient, multipart conditions and pests' documentation processes.



Figure 1.3. Illustration of manual disease detection

The second method of plant disease detection is a technology and internet-related method. In recent years, due to the fast development of DL-related theories and the enhancement of computational ability, deep convolutional networks have attained great success in supercomputer vision. In terms of item detection, the precision of DL-based methods greatly exceeds that of traditional methods based on manual design features. Body or object detection is to draw a range of objects of interest in an image, then select the target box with a rectangle and label it with a type.

Based on DL, there are primarily two types of object detection: one uses convolutional network structures based on region generation, with the representative network being the faster R-CNN (Zhao et al., 2018); the other approaches object detection as a regression problem, processing the entire image using the CNN network structure while simultaneously predicting the type and location of the object.

### 1.3.1. Significance of computer vision-based detection

In the domains of computer vision, sophisticated image processing, and ML, these past two decades of theoretical and applied research on early object detection methods of crop diseases and pests in the natural environment have been significant research routes. Traditional methods cannot achieve real-time and accurate disease detection due to the complexity of the early images of tomato diseases and pests in the natural

environment and the fact that greenhouses are the controlled environment where AI algorithms, particularly CNN, can accurately serve disease detection.



Figure 1.4. Camera vision-based and CNN for diagnosing and clustering

AI enables the forecasting of the thermophysical behavior of biomass, and tools like ANN have established themselves as essential in the advancement of research into the prediction of biomass energy, which in turn might be used to control greenhouse microclimates. Although ANNs may adapt to non-physical and non-linear data, they need a sizable multidimensional data set to lower the danger of extrapolation. Furthermore, AI uses a variety of mathematical and algorithmic techniques, including operational research, limited programming, DL, and ML. DL enhances traditional ML by giving modeling greater depth. Its benefit is featuring learning, which allows rapid resolution of complicated issues and the automatic extraction of features from unprocessed data. DL consists of a variety of mechanisms, including convolutions, clustering layers, fully connected layers, gates, memory cells, and activation functions as shown in Figure 1.7. depending on the network architecture employed (Alhnaity et al., 2019).

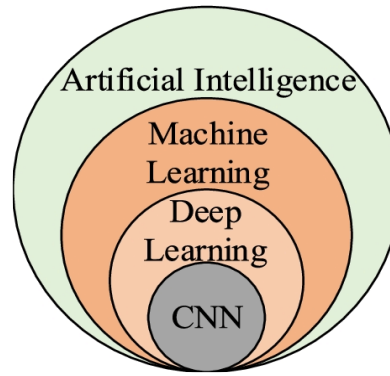


Figure 1.5. Understanding of AI concept

How much time is needed to control and decide the early detection? Based on the research that has been done by Jiang et al., (2019), experimental results show that the INAR-SSD model realizes a detection performance of 78.80% mAP on ALDD, with a high-detection speed of 23.13 FPS. The results demonstrate that the novel model provides a high-performance solution for the early diagnosis of apple leaf diseases that can perform real-time detection of these diseases with higher accuracy and faster detection speed than previous methods.

### 1.3.2. Deep learning techniques in agricultural greenhouse

DL has proven to be very capable in both pattern recognition and classification. By sequentially focusing attention on pertinent portions of the available data, this form of the network can actively learn to identify patterns (Schmidhuber, 2015). The advantage of DL over traditional neural networks is the ability to create simulated data sets to train the model, allowing it to solve real-world issues like greenhouse systems. The different uses of DL in agriculture are discussed in (Kamilaris and Prenafeta-Boldú, 2018), although its application in greenhouses is still limited.

### 1.3.3. CNN models

The CNN model is a development of the DL. They are made up of MLP networks, which include fully connected and various pools. Backpropagation is used to train and improve filters on each layer. The features that these trained and learned filters extract represent the input image. As this sort of model was able to outperform cutting-edge algorithms, it has taken the lead as the most sophisticated technique for a variety of data processing jobs. At the moment, CNN architectures are either created from scratch or are modified from pre-trained designs. Transfer learning can be used by using pre-trained architectures. Transfer learning is the process of applying models that have been

previously trained using substantial amounts of data from other systems, to issues or systems that are similar to those systems (Alzubaidi, et al., 2021).

The study conducted by Alhnaity et al. (2019) reported a new deep RNN and a long short-term memory neural network (LSTMNN) model to anticipate the stem diameter or tomato performance problems that used climatic factors such as CO<sub>2</sub>, humidity, radiation, external temperature, and indoor temperature, stands out among the studies that have been conducted on DL in greenhouses. This technique's biggest drawback—the substantial amount of data required for the training process—is revealed (Alhnaity et al., 2019).

The vast majority of DL algorithms are built on the BP network, which also enables models with many processing layers to represent data at different levels of abstraction. As a result, research into the BP network's use in agriculture has started. With picture recognition, it can be utilized for plant disease detection, weed management, and plant counting in the context of smart farming (SF). The most effective deep network types in this sector include CNNs, RNNs, and generative adversarial networks (GANs) (Zhu et al., 2018).

#### **1.3.4. Artificial neural networks and greenhouses**

ANNs have a significant impact on controlling different parameters in modern greenhouses for organizing microclimate energy distribution and plant growth stages therefore it's important to use resend artificial neural network applications to control plant protection systems and many researchers have been done to show what is necessary for the future of greenhouse management and increasing productions yield and a good quality.

The use of mathematical models to forecast changes and the adaptability of the greenhouse has been thoroughly studied. The complexity of these models is governed by the complexity of the greenhouse itself, and the most useful model to use physically or those that evaluate the system's inputs and outputs largely depends on the type of crop and greenhouse you have available. The use of ANNs as a response to these demands has been established for almost 40 years. 35 papers were examined, and it was discovered that 74% of them focus on defining and forecasting the microclimate, 9% on energy optimization, and 17% on diverse applications involving greenhouse networks (Escamilla-García et al., 2020).

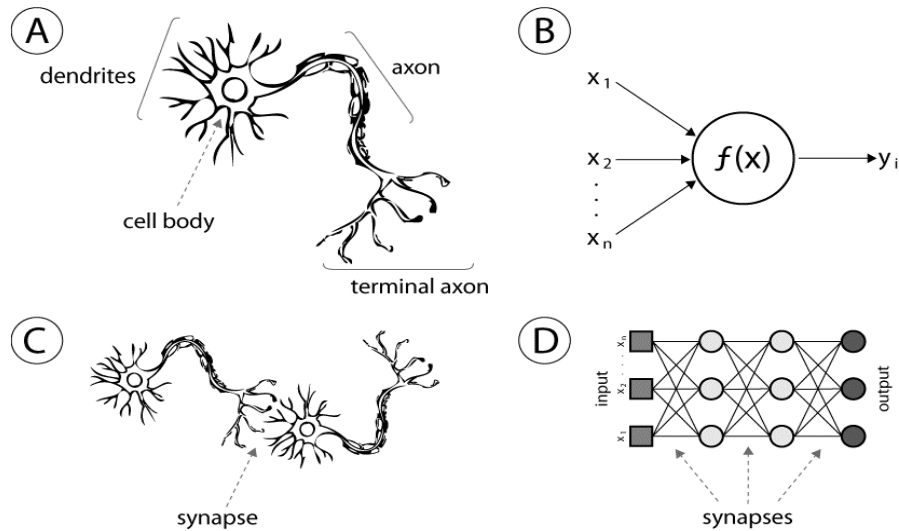


Figure 1.6. Schematic of ANN (Maind and Wankar, 2014)

### 1.3.5. Convolutional neural networks

Many scholars and researchers have reported using a variety of methods for monitoring and detecting plant diseases. Presenting the Gabor wavelet transform which is a method for removing tomato leaf characteristics. Mokhtar et al., (2015) applied SVM to find leaf ailments genuine sample photos for investigations of tomato leaves have been thought about, and two varieties of the disease, such as Early blight and Powdery mildew, in tomato leaves. Evidence of mildew phase pictures is used in preprocessing to 512\*512 resolutions to determine the computation period. The method of background subtraction has been used to take the image's background out. The SVM was trained for classification using a kernel function and evaluated (Sardogan et al, 2018).

It was suggested using fuzzy logic to computer vision segmentation techniques for early disease detection on plant leaves. Segmenting an image is additionally utilized to remove the damaged portion of the plant leaf from the source picture. Additionally, color space segmentation is used to determine the fruit's color or the presence of a disease area (Sehgal and Mathur, 2019).

The characteristics of color and texture are the images are edited. The suggested system is composed of two sections. Extracting disease characteristics is one, while the other falls under the same classification. The previous focus on the presence of a characteristic depending on a co-occurrence matrix is dependent on texture and gray-level equations. Later, the fuzzy neural network is deployed. Using a network, which is

essentially an unsupervised approach, sort various disease categories. The five classifications that are taken into consideration in this study are healthy plants free of disease, rust, scab, downy, and mildew. This range of photos of damaged grape leaves allows classification using the suggested system. When performance is measured in terms of accuracy, the desired result of 90% is attained (Sabrol and Satish, 2016).

The programmed determination framework based on plant malady highlights diminishes the reliance on specialists within the region concerned. The characteristic highlight of plant infection extricated from the infected locale of the disease-affected pictures utilizing picture preparing procedures. The disease-infected plant is recognized by design acknowledgment strategies such as “neural networks”, and “support vector machines”. In general classification precision is gotten as 97.3% is gotten with these six sorts of tomato images. Another study has detailed a smartphone-based diagnosis application for plant disease discovery. Color normalization has been utilized for the same (Petrellis, 2019).

Multiclass plant disease location utilizing SVM is actualized Wang et al., (2012) created a strategy to realize image recognition of plant maladies, four sorts of neural networks including backpropagation systems, spiral basis function neural systems, generalized regression networks and probabilistic neural networks were utilized to recognize wheat stripe rust from wheat leaf rust and to recognize grape wool buildup from grape fine buildup based on color highlights, shape features and surface highlights extricated from the disease images (Wang et al., 2012). K.K. Singh has proposed a cloud-based collaborative stage for plant malady distinguishing proof. So that it can be forecasted to ranchers for way better prevention purposes (Singh, 2018).

Different strategies for the segmentation method are examined assist which can be connected to the detection of different injury infections. To recognize various ailments in plant leaves a few calculations were utilized such as K-means cluster, thresholding, etc. For the distinctive inputs, k Figures delineate the first pictures taken after by the yield of thresholding and K-means cluster for portioned images. Although the calculation of k implies clustering needs prior information of detail around a few cluster centers. The modestly viable result for different data sets is given by this strategy in differentiation with the thresh-holding method (Li et al., 2012).

### **1.3.6. The importance of CNN**

DL constitutes a modern technique for image processing, with large potential. Having been successfully applied in various areas, it has recently also entered the domain of agriculture. CNN, a class of ANN that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including disease detection in agriculture. This chapter includes a survey of previous work that has been done to focus on the roles and application of CNN in disease detection and fruit productivity. For the picture like big dimensional and generally enormous information, there's a specialized arrangement to engineer called Convolutional Neural Systems (CNNs). CNNs are comprised of convolutional layers, pooling layers, enactment work layers, dropout layers, and fully-connected layers. Convolutional layers hold the comes about of the convolution of channels or bits with the past layers. These channels or parts comprise weights and predispositions to be learned. The purpose of the optimization work is to create these bits that speak to the information without mistakes.

Another factor why it is crucial to use leaves for disease identification because numerous earlier studies have taken this issue into account by analyzing photos of plant leaves and creating a specific classifier to divide the samples into images that are healthy and those that are unhealthy. The choice of plant leaf photos for analysis was made since most plant diseases usually manifest their symptoms first on the plant's leaves (Barbedo, 2019).

### **1.3.7. Definition of CNN**

The term convolutional neural systems alludes to a lesson of neural systems with a particular arranged design in Figure 1.7. , where each so-called covered-up layer ordinarily has two particular layers: the primary organize is the result of a neighborhood convolution of the last layer (the part has trainable weights), the second arrange may be a max-pooling organize, where the number of units is essentially reduced by keeping as it were the greatest reaction of several units of the primary stage. After a few covered-up layers, the ultimate layer is ordinarily completely associated. It encompasses a unit for each lesson that the organization predicts, and each of those units gets input from all units of the past layer.

In recent years, most works centered on picture classification of plant diseases and bother based on artificial neural network programs such as DL. There are a few considerations on question discovery of trim infections and pests based on profound learning (Arsenovic et al., 2019) The existing object detection strategies of tomato disease and bugs based on DL and CNN algorithms generally utilize the calculation based on region suggestion for location (Fuentes, 2022). The exactness has greatly made strides compared with conventional methods, but the question discovery handle takes a long time and it is difficult to distinguish and find tomato maladies and pests in actual-time beneath common conditions.

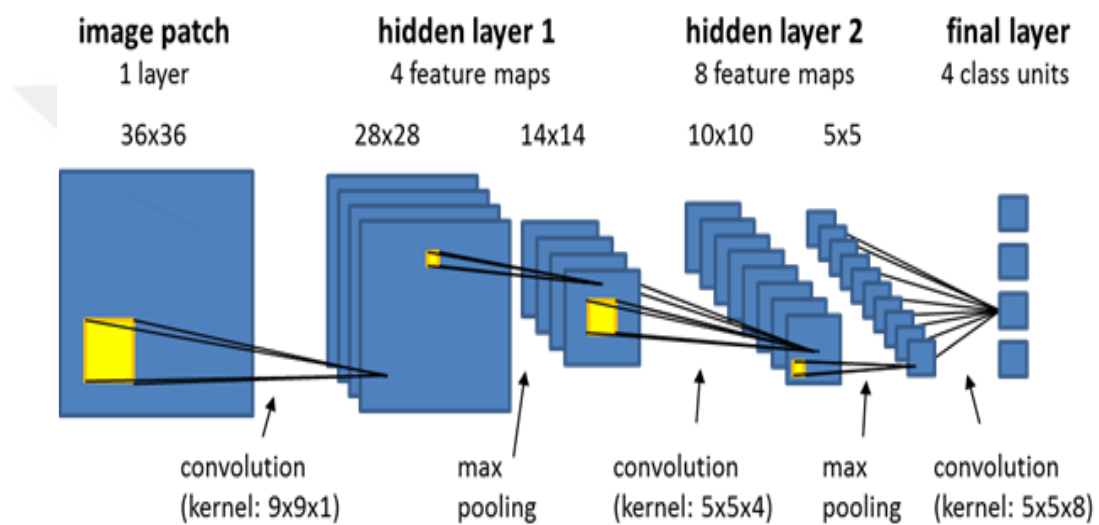


Figure 1.7. Schematic of a CNN with two hidden layers (Dalai, 2020)

### 1.3.8. Smart innovations and applications of greenhouse

The term SF describes the use of data collection and analysis to promote efficient production processes, which inspire researchers, practitioners, commercial enterprises, and government organizations to work toward the creation and promotion of cutting-edge technology to assist farmers locally. The utilization of hardware and software technologies, such as the deployment of sensor nodes, control systems, robotics, satellites for imagery and positioning, data storage and analysis, advisory systems, and terrestrial and aerial drones, is a crucial aspect of this novel approach to farm management. However, the goal of SF should not just be to industrialize agriculture but to improve it overall by making it more effective, sustainable, and high-quality while still taking into account the requirements of farmers.

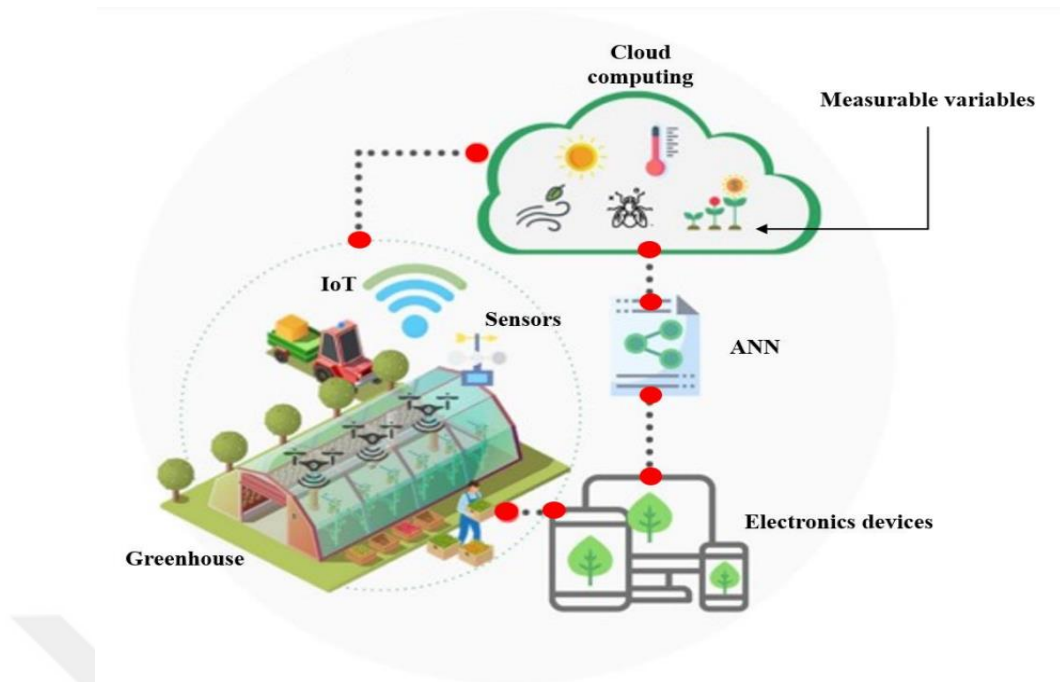


Figure 1.8. The potential of smart agriculture (Bacco, et al., 2019)

For a long time, greenhouse innovation in agribusiness has been well mechanized, and data innovation heading with the IoT (Internet of Things) innovation quick improvement and wide application. This framework is made up of front-end information securing, information preparation, information transmission, and information gathering. The surrounding temperature is handled by the temperature sensor of the information terminal hub. Prepared information is sent to the middle hub through a remote network (Dan et al., 2015).

The common plant disease that faces all parts of the plant, Figure 10. According to the Encyclopedia of Microbiology, the definition of plant disease is “the state of local or systemic abnormal physiological functioning of a plant, resulting from the continuous, prolonged ‘irritation’ caused by phytopathogenic organisms (infectious or biotic disease agents) and (Tampakaki et al., 2010).

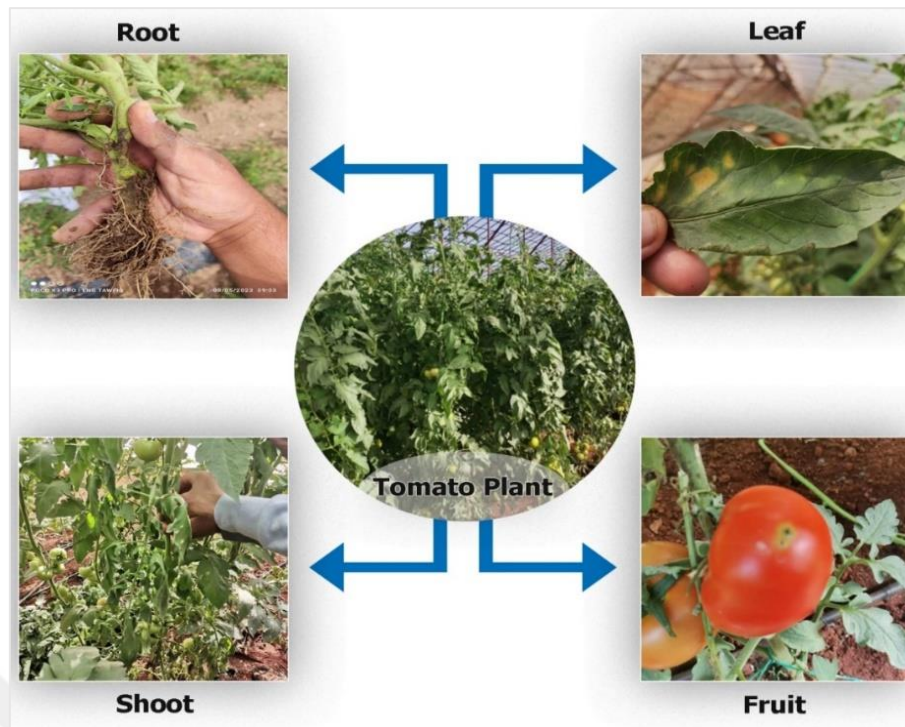


Figure 1.9. Demonstrates the parts of a plant that face diseases mostly

Crop disease detection and control are necessary to reduce plant diseases as plant diseases cause great losses to the production, economy, quality and quantity of agricultural products. Therefore, it is necessary to monitor the disease of crops from the first stage of the crop life cycle to before harvest. The traditional method is a visual observation, which is not only time-consuming and labor-intensive but also requires supervisors to have a lot of professional knowledge. To solve these problems, researchers have proposed a series of automated and intelligent disease detection methods. Here are some typical cases of crop pest detection using AI technology in recent years.

The main implementation methods of disease recognition are gradually moving from traditional DL to few-shot learning. This is primarily because the algorithms based on DL typically rely heavily on large amounts of data. DL driven by big data faces the challenges of the high cost of data acquisition, high cost of high-end hardware and high consumption of power resources (Li and Chao, 2021), which is not conducive to the sustainable development of agriculture. To conform to the sustainable development of agriculture, researchers should focus on the trade-off between data quality and quantity. For the data quality in the agricultural field (Li et al., 2022) believed that limited good data can defeat a large number of bad data. For the problem of data quantity (Yang et

al., 2022) examined the application of few-shot learning in smart agriculture, and (Li et al., 2020) investigated sustainable computing in intelligent agriculture. The results show that using small samples in some agricultural tasks can achieve a better model algorithm with less sample data.

A study conducted by Brahimi et al. (2017) proposed a deep-learning approach to build a classifier for disease classification. They introduced the Convolutional Neural Network (CNN) as a learning algorithm to train the classifier. Their dataset contained 14,828 images of tomato leaves infected with nine diseases. The obtained results were encouraging, reaching 99.18% accuracy, which outperforms dramatically shallow models, and they can be used as a practical tool for farmers to protect tomatoes against disease.

Similarly, Sahu et al. (2018) designed a system, that included a digital camera of a smartphone and applied CNN and DL to detect insects in a farm field and population estimation of insects in a farm field. The overall results of the study had 90.9% accuracy in which farmers can apply the results to control pests in a farm field and a healthy crop yielding for increased rice production.

Li and Yang, (2020) used two datasets to verify the effectiveness and feasibility of the few-shot model, one is the National Bureau of Agricultural Insect Resources (NBAIR), and the other is a dataset with the natural scenes. They adopted the Convolutional Neural Network (CNN) to extract feature vectors of images. The testing accuracy for the two datasets is 95.4% and 96.2% respectively, showing the generalization ability of the proposed few-shot model. Moreover, their work can also be regarded as a positive attempt to combine the software and hardware for the landing of intelligent algorithms in agricultural applications.

Li et al. (2020) proposed two methods namely SCNN-KSVM (Shallow CNN with Kernel SVM) and SCNN-RF (Shallow CNN with Random Forest) to detect the plant diseases in maize. The comparison experiments with other DL models were carried out on three different datasets. The results of testing accuracy for the datasets are 94% and 94% respectively showing that the CNN-SVM and CNN-RF outperform other pre-trained deep models on the indicators of precision, recall, and F1-score, with fewer parameters. The combination of shallow CNN and classic ML classification algorithms is a positive attempt to deal with plant disease identification simply.

In Chen et al. (2020) research, they used AI and image recognition technologies combined with environmental sensors and the Internet of Things (IoT) for pest identification. They also combined the current mature IoT technology and DL and applied it to smart agriculture. They used DL, and YOLOv3 for image recognition to obtain the location of *Tessaratoma papillosa* and analyze the environmental information from weather stations through Long Short-Term Memory to predict the occurrence of pests. The experimental results showed that the pest identification accuracy reached 90%.

In a recent paper conducted by Li and Chao (2021), they proposed a semi-supervised few-shot learning approach to solve plant leaf disease recognition. The results showed that the average improvement by the single semi-supervised method is 2.8% and that by the iterative semi-supervised method is 4.6% with DL accuracy of 92.6% so they concluded that the proposed methods can outperform other related works with fewer labeled training data.

The results of research published by Zeng et al. (2021) proposed a data generation model based on the cycle Generative Adversarial Network model which introduced a leaf foreground module (LFM) block. Experiments showed that the model can generate high-quality grape leaf disease images with DL accuracy of 90.91% and 92.44% correspondingly, which can improve the accuracy of grape disease recognition task in a Few-Shot Grape Leaf Diseases Classification task.

Li and Yang (2021) conducted a study for the first time on task-driven meta-learning few-shot classification in the ground of agriculture to the best of their knowledge. Specifically, they collected samples from publicly available resources to assemble a comprehensive dataset for the few-shot classification, covering both pests and plants to analyze the single domain or cross-domain. Then, they performed 36 groups of comparison experiments to establish the baselines of testing accuracy with 6000 images, and it was found that the percentage of CNN's Few-shot learning accuracy was 90.4%.

Mukhtar et al. (2021) proposed a wheat disease recognition network based on one-shot learning which not only needs a small number of images for training but also can accommodate new categories because it can be trained even on a few images of a new type. They have used the MobileNetv3 network as a feature extractor which is extremely fast and accurate classification. Their Mobilenetv3 model has achieved

around 98% training and 96% validation, while the whole one-shot network has achieved more than 92% accuracy, 84% precision, and 85 recall.

The research work conducted by Yasmeeen et al. (2021) proposes an automated framework using DL and best feature selection for citrus disease classification. The experimental process is conducted on three different datasets-Citrus Hybrid, Citrus Leaf, and Citrus Fruits. On these datasets, the best-achieved accuracy is 99.5%, 94%, and 97.7%, respectively.

#### **1.4. Aim of the Study**

Pests and diseases account for major losses in agriculture. Massive crops, vegetables and fruit products are lost due to pests and disease without considering other contributes to decreased productivity. Plant protection is one of the major disciplines of agricultural science, which principally goals to protect plants from getting affected. secondarily it takes measures to treat infected plants to overcome infestation and eventual product loss. Plant protection involves, addressing potential causes of diseases, assessing the occurrence and degree of severity of infection, forecasting the future infestation based on models considering various factors, examining and monitoring undergoing treatments and describing the efficacy and mechanisms of protective actions. In the process of plant protection, physical, chemical and biological elements are engaged as per the requirement of the action undertaken.

Fundamentally, in the phase of early disease detection, a physical diagnosis is required before the necessary chemical or biological treatment is administered. It is obvious that physical diagnosis and evaluation is tedious work and may span for a very long time in a large agricultural business, and of progression, requires highly skilled personnel. Concerning physical diagnosis, when it comes to determining the degree of severity the task is highly subjective and requires specialized skill and expertise. During protective treatments applying chemical pesticides assuming a higher degree of severity is one of the reasons associated with high operational costs and adverse ecological effects. Therefore, increasing the scale of diagnosis by adopting smart agriculture-based approaches has been in action in the area of plant disease protection. These methods are adopted both in the field and in greenhouses, for various crops, vegetables and fruit types. One of the characteristics of AI-based tools is that a certain model is highly specific and customized. For instance, a model developed to detect tomato disease such

as early blight in a field condition may not apply to the same case situated in a greenhouse. The efficacy of a model is influenced by the input data used to develop the particular model. Moreover, numerous reasons why the data quality, type, geographic location and environmental condition significantly affect the adaptability of a certain model. Hence, as long as a general-purpose model is not available custom-designed model based on the local and suitable data to which the model is going to be applied is unquestionably necessary. Thus, with the above statement and insight this study aims to develop an AI-based tomato disease classification model for greenhouses. Having this, the general aims and objectives of this study are listed below:

1. To collect images of healthy and diseased tomato leaves from a greenhouse located in Antalya- Türkiye.
2. To create a labeled database of tomato diseases which serves as a data bank for further study and analysis.
3. To develop a disease classification model using the custom image dataset.
4. To assess the performance of transfer learning adopting popular CNN architectures.
5. To compare transfer learning-based models and custom CNN models.

#### **1.4.1. Project scope**

The research focuses on collecting data from greenhouses in Antalya-Türkiye. The image datasets include relatively small image datasets, which confines the scope of the project to assess AI models with smaller datasets, and it consists of four classes namely healthy Early Blight, Powdery mildew and Yellow leaf disease. In the development and training of models, the scope of this study includes training and evaluating only CNN algorithms, both transfer learning which makes use of the power of popular CNN architectures and custom CNN training which involves training a model from scratch. In the performance assessment, accuracy, precision and recall metrics are applied. Having these elements, being confined to a single geographic location and data collection setup are the limitations of this study.

#### **1.4.2. Unique values of this research**

First, the creation of a labeled custom dataset specifically from the Antalya region has the potential of capturing unique characteristics of tomato disease prevalent in that particular area and making the dataset public data, it has contributed to enhancing

collaboration in agricultural research. Second, as it is a greenhouse-focused application, it can provide insight and understanding tailored to the requirements of geographic location. Third, while there are tomato disease detection studies, the study aims to show the capacity of adopting transfer learning models using limited data.



## **2. LITERATURE REVIEW**

### **2.1. General Literature Review**

This chapter covers related and relevant research works, regarding the application of AI in the area of plant disease detection. As AI is a vast subject, the scope of the review is confined to the techniques, features and models used in disease detection of different plants. The first section covers a general overview of models and algorithms used to recognize plant disease. Moreover, the second section covered only transfer learning-based research works related to plant disease recognition, which only addressed state-of-the-art vegetable virus detection.

In the literature section, related terms such as leaf diseases and plant virus detection, tomato, CNN, machine learning and transfer learning, and deep learning keywords were used as search queries in popular databases such as Springer, MDPI, Google Scholar, IEEE, ScienceDirect, and others. After manual filtering of the research articles based on recent periods and articles related to a few plants namely tomato, maize, potato, apple, rice and wheat, the review mainly focused on around 70 research articles and academic papers. Moreover, the review addressed the following questions- which plants were the subject of the study? How many diseases (class of diseases) were detected? Which models were used? What was the detection or classification rate of the successful models to summarize the research articles on the topic of this thesis?

Early information on crop health and disease detection may be obtained by implementing the proper management measures, such as fungicide applications, disease-specific chemical applications, and vector control through pesticide applications. This might help with disease management and boost output. A study conducted by Sankaran et al. (2010) informed the need for creating a quick, affordable, and dependable health monitoring sensor that supports agricultural developments. To build ground-based sensor systems to aid in monitoring plant health and diseases in the field, they discussed the currently employed technologies, including spectroscopic, imaging-based, and volatile profiling-based plant disease detection approaches.

It was decided to use image processing disease recognition approach among other commonly used approaches for plant disease diagnostics, such as double-stranded Ribonucleic Acid (RNA) analysis, nucleic acid probes, and microscopy, after analysis of their work and analysis presented by the authors of (Sladojevic et al., 2016). There

are numerous methods currently in use to identify plant diseases using computer vision. One of them is disease detection using color feature extraction. As a result of the employment of the color models in this investigation, disease spots were successfully identified and were unaffected by noise from other sources, such as camera flash.

The method of extracting shape data could also be used to detect plant diseases. In their final tests, Patil and Bodhe (2011) employed this technique to detect disease in sugarcane leaves. They used threshold segmentation to calculate the leaf area and a triangle threshold to determine the lesioning area. The ability to extract textural features may also be utilized to identify plant diseases. By generating the gray-level cooccurrence matrix on an image, (Patil et al., 2012) suggested a model for plant disease identification using texture properties such as inertia, homogeneity, and correlation. They experimented with disease detection on maize leaves in conjunction with color extraction.

Combining all of these features yields a rich feature set that may be used to enhance images and improve classification. The authors (Reed and Dubuf, 1993) have provided an overview of popular traditional feature extraction techniques. The application of these methodologies and strategies is the main emphasis of this paper's work because AI science is moving along so quickly. To advise corrective actions for pest or disease management in crops, they created a software model.

Another method suggested by Revathi and Hemalatha (2014) combines the features obtained from Particle Swarm Optimization (Zhou and Liao, 2013) with a forward neural network to identify the injured cotton leaf spot and improve the system's accuracy, which is ultimately 95%. Furthermore, Support Vector Machine techniques can be used to identify and classify plant diseases. Where the classification accuracy ranged from 65% to 90%, depending on the kind and stage of the disease.

For the detection of plant diseases, some techniques combine texture extraction and Neural Network Ensemble (NNE). NNE provides a higher generalization of learning ability by training a certain number of neural networks and merging their results later. Only tea leaf diseases were detected using this method, and the final testing accuracy was 91% (Karmokar et al., 2015).

K-means, a clustering strategy suggested by Al-Hiary et al. (2011), was combined with another method based on leaf pictures and using ANNs as a methodology for the

automatic detection and categorization of plant diseases. There were 10 secret layers in the ANN. The total number of outputs was 6, with each class representing one healthy leaf and five different diseases. This method's categorization accuracy was 94.67% on average.

Tarek et al. (2022), completed a crucial new project in Egypt (2022) The idea of "smart agriculture" is put into practice by suggesting a platform that can identify tomato leaf disease. The detection and diagnosis of tomato leaf diseases were examined using a variety of deep CNN approaches. The PlantVillage dataset was used to train the CNN models ResNet50, InceptionV3, AlexNet, and the three iterations of MobileNet. Several optimization methods were used to train each model. MobileNetV3 Large's greatest accuracy, as measured by the Adagrad optimizer, was 99.81% with a loss value of 0.0088. It is important to note that they used the data set's images of leaf disease rather than actual images while achieving a satisfactory accuracy ratio.

Because plant leaves come in a variety of sizes and shapes, diseases that affect them can also vary depending on how many plants' leaves are involved in the enhanced output. According to Chakraborty et al. (2022), several researchers have recently concentrated on developing deep-learning models for the categorization of diseases. Early disease detection is preferred to control the spread of pathogens and manage the sickness. Many techniques that enable the precise detection of objects have recently been developed as a result of recent advancements in the fields of AI and ML. By drawing inspiration from them, the results are rather encouraging since the improved model has an accuracy of 97.211% in identifying plants with symptomatic leaf curls.

Several researchers in the fields of biology, bioinformatics, biomedicine, robotics, and 3D technologies have introduced DL techniques for solving the most difficult problems. The work is an attempt to use DL to recognize tomato leaf disease, and it is motivated by DL techniques and their practical use. There is no proof that researchers investigated a DL approach for plant disease recognition from the leaf photos despite a thorough search of the state-of-the-art literature. In the materials and techniques, the deep CNN recognition technique is provided.

Also, another investigation states that right now, the agri-food division takes advantage of present-day apparatus, and apparatuses, and developing data and communication innovations that consider the Internet of Things (IoT) capabilities.

These executions have given way to an unused period of agri-food generation called 'Agri-Food 4.0', where computerization, network, digitalization, the utilization of renewable energies, and the effective utilization of assets are transcendent in this segment (Miranda et al., 2019).

Besides, the significance of later advances and applications is additionally, this paper states the specialized and logical state of the craftsmanship of remote sensor advances and standards for remote communications within the Agri-Food division. These advances are up-and-coming in a few areas such as natural observing, accuracy farming, cold chain control, and traceability (Ruiz-Garcia et al., 2009).

As it is known, agricultural activities are carried out in a closed area (with materials such as glass, plastic, etc.) and under controlled conditions. In other words, it is also called greenhouse vegetable and fruit cultivation. Under controlled conditions, controlling the independent variables such as temperature and humidity is one of the primary activities of greenhouse cultivation. However, the fact that the agricultural activities carried out in the greenhouse are in a closed environment, causes a rapid increase in plant pests in any part of the greenhouse. This negativity is a situation that the farmers dealing with greenhouse activities do not want. The biggest problem for the farmers is the inadequacy of the cost analysis (for pesticide applications) and the high cost of pesticides. In addition, early diagnosis of the disease is insufficient.

Kerkech et al. (2018) used a CNN model to detect Esca disease in grapevines using UAV RGB images. A CNN model was trained with different combinations of patch sizes  $16 \times 16$ ,  $32 \times 32$  and  $64 \times 64$  with different color spaces. All the different color spaces and vegetation indices used in this study can be calculated from the RGB images. It was pointed out that the lower accuracy of the models trained which combines all color information into a single channel and is less relevant for the network to learn the best color features.

Barbedo, (2018) investigated how dataset size and diversity affect the performance of DL techniques applied to plant diseases. An image dataset of leaves with a small number of samples for the CNN model was used to investigate the behavior of GoogLeNet under different conditions. Data augmentation and transfer learning methods were used to train the CNN. The results showed that even with transfer learning and augmentation techniques, CNN requires a large number of images to extract useful

features from the data. Even though a large number of images can be easily acquired with the new technology, labeling the dataset is time-consuming. One option proposed in this paper was to share the dataset, but as mentioned earlier, the two environments are not the same in agriculture. The effect of removing the background of the images on the accuracy of the model was also investigated. The model has trained again with the images without background. The model has trained again with the images without background. The results show different behaviors concerning the accuracy of the model, including no significant effect, a significant improvement in accuracy in some cases, a significant decrease in accuracy, and mixed results (improved accuracy in some diseases while the error rate increased in others). From the significant decrease in accuracy for some plants, it was inferred that the CNN model sometimes uses the background of the model to classify the images. An attempt that can be made here is to train the model with both datasets (with and without background) and investigate the accuracy of the model in this case.

Jiang et al. (2019) implemented the CNN models to detect five common apple leaf diseases using images from the field. By applying data augmentation, such as rotational transformations, horizontal and vertical flips, intensity disturbances, and Gaussian noise, 26,377 disease images were generated., where two GoogLeNet inception layers replaced two convolutional layers of the VGG model. Moreover, the structure of feature extraction and fusion is designed by applying the Rainbow concatenation method instead of the pyramid to improve the performance of feature fusion. The results showed that the data augmentation improved the accuracy of the model by 6.91% compared to the original dataset. To investigate the effect of using a deep model as a backbone, ResNet-101 was used as a feature extractor for SSD. The results show that ResNet-101 does not lead to any improvement. In popular terms of speed, Faster RNN was the slower model with more accuracy. The proposed model with 78.80% mAP and 23.13 Frames Per Second.

Karthik et al. (2020) proposed a Residual Attention Network for disease detection in tomato leaves. The main idea of attention is to focus on the relevant parts of the input to produce outputs. The dataset was collected from the open-source website Plant Village and contained one healthy class and three diseases. They implemented three models. One was a traditional CNN model, the second used the three residual layers introduced into the CNN model as part of the ResNet architecture, and the last used an

attention mechanism on top of the residual CNN for effective feature learning. The traditional CNN-based methods focus on ordered feature learning, starting from basic image-level features such as edges, color, etc., to complex texture-based differences. In the deeper layer, some important features extracted in the first layer may be lost. The residual layers are designed to avoid this problem. They concatenate the extracted features from the earlier layers with the deeper layers. In addition, the attention mechanism is used to extract the relevant parts of the feature maps. The Residual Attention Network CNN performed better with an overall accuracy of 98% than the Vanilla Residual Network with an accuracy of 95% and the traditional CNN model with an accuracy of 84%.

Kerkech et al. (2020) used the SegNet model to segment and detect grapevine diseases using images with RGB and infrared ranges. The dataset was acquired using a UAV device with two MAPIR Survey2 camera sensors, including a visible light (RGB) sensor set for automatic illumination and an infrared sensor. The dataset was labeled using a semi-automatic method, a sliding window to identify potentially diseased areas. Then, each block was classified by a LeNet5 network for pre-labeling. In the end, the labeled images were manually corrected. The model trained with RGB images (AC = 85.13) outperformed the model trained with infrared images (78.72). The fusion AND had the best performance, and the fusion OR had the worst accuracy. The runtime of SegNet on a UAV image was reported to be 140 s for visible and infrared images. The fusion between the two segmented images takes less than 2 seconds.

Lu et al. (2017) implemented the model Cascade Inception to detect four common apple leaf diseases in images captured in the field. The Cascade Inception was a modified AlexNet model with inception layers from GoogLeNet. Various data enhancement methods such as image rotation, mirror symmetry, brightness adjustment, and PCA jittering were applied to the training images. Moreover, the fully connected layers were replaced by convolutional layers, which results in fewer parameters and avoids overfitting. The transfer learning method was used to train VGGNet-16 and achieved 96% accuracy. Standard AlexNet, GoogLeNet, and ResNet-20 were trained from scratch using SGD optimization and achieved a maximum accuracy of 95.69%. The SVM and BP, which achieved an accuracy of less than 60%, show that the traditional approaches rely heavily on expert-developed classification features to improve the recognition accuracy. They also investigated the effect of data

augmentation methods and optimization algorithms on accuracy. The model with the SGD optimizer achieved 93.32% accuracy, while the model with NAG achieved 97.62% accuracy.

Picon et al. (2019) used the DL model to detect seventeen diseases of five crops (wheat, barley, corn, rice, and rapeseed) in images captured in the field. The dataset contained several challenges, such as multiple diseases on the same plant, similar visual symptoms among diseases, images of early and early-stage diseases, and diseases of leaves, and stems. The results show that by adding crop ID, the model can still benefit from more images for training, while crop ID information helps the network to discriminate between similar diseases.

Chen et al. (2021) used pre-trained MobileNet V2 networks to identify twelve rice plant diseases. The dataset was collected from online sources and real agricultural fields. An attention mechanism was added to the model, and transfer learning was performed twice during model training to achieve better performance. The MobileNet-V2 achieved 94.12% accuracy on the PlantVillage dataset, while the MobileNet-V2 with attention and transfer learning achieved 98.26%. The five CNNs such as MobileNetv1, MobileNet-V2, Mobile, EfficientNet-B0, and DenseNet121 were selected for comparison with the proposed models. In addition, two other networks named MobileNetv2-2stage and MobileAtten-alpha were trained. In the MobileNetv2-2stage model, transfer learning was performed twice to identify the images of plant diseases.

Similarly, in MobileAtten-alpha, the attention method was used instead of transfer learning twice. The proposed model (using attention and transfer learning twice) achieved the second-best accuracy of 98.84%. The DenseNet121 with an accuracy of 98.93% outperformed other models. MobileNetv2-2stage and MobileAtten-alpha achieved an accuracy of 98.68% and 96.80%, respectively. The proposed model was trained with images from the field and achieved an accuracy of 89.78%.

The overall findings of a study performed by Narvekar and Rao (2020), on the classification of flowers is a difficult task because of the huge number of flowering plant species, which are similar in shape, color and appearance. They have demonstrated and analyzed the recent developments in DL methods such as CNN and transfer learning in CNN. Prototype CNN model architecture proposed and transfer learning approach as

well examined on VGG16, MobileNet2 and Resnet50 architecture for flower classification on publicly available flower dataset.

Latif et al. (2022), proposed a deep convolutional neural network transfer learning-based approach for the accurate detection and classification of rice leaf disease. The modified proposed approach includes a modified VGG19-based transfer learning method. The proposed modified system can accurately detect and diagnose six distinct classes: healthy, narrow brown spot, leaf scald, leaf blast, brown spot, and bacterial leaf blight. The highest average accuracy is 96.08% using the non-normalized augmented dataset. The corresponding precision, recall, specificity, and F1-score were 0.9620, 0.9617, 0.9921, and 0.9616, respectively

In another research experiment by Liang et al. (2019), the authors proposed a deep convolutional neural network (CNN) for the recognition of rice blast disease. They used a dataset of 5812 divided equally between infected and non-infected rice plants, which are publicly available. Their method uses CNN for feature extraction and SVM for classification, and they reported an average accuracy of 95.83% for binary classification.

The results of Li et al. (2020), show that the SCNN-KSVM and SCNN-RF outperform other pre-trained deep models on the indicators of precision, recall, and F1-score, with fewer parameters. The combination of shallow CNN and classic ML classification algorithms is a positive attempt to deal with plant disease identification simply.

In the studies presented by Niedbała et al. (2020), ANN was adopted to examine the impact of the variety and weather conditions on the concentration of ferulic acid, deoxynivalenol, and nivalenol in winter wheat grain. As a result of the studies, three independent neural models were developed, one for each of the parameters considered. These models were based on the multilayer perceptron topology with two hidden layers and 14 inputs. Based on sensitivity analyses conducted for the models, the authors pointed out that the experiment variant and wheat variety were two of the most important features determining the concentration of ferulic acid, deoxynivalenol, and nivalenol in winter wheat grain.

The authors of the studies discussed in the article Li and Chao (2020) proposed an ANN-based continual classification method that may be used in solving many potential

classification tasks in the field of agriculture. The method is based on memory storage and retrieval, and it combines a convolutional neural network (CNN) and a generative adversarial network (GAN). The evaluation of this method was carried out for two classification problems: identification of crop pests and plant leaves. The method was also compared with the regular CNN approach. The results show that the regular CNNs classify the categories clearly for single-task problems, whereas in the continual tasks, they are prone to forgetting problems and cannot balance new and old tasks. The method proposed by the authors can accumulate knowledge and alleviate forgetting. Therefore, it can better distinguish all categories from both old and new tasks.

In the studies described by Aji et al. (2020), the authors attempted to apply ANN to model the dynamic responses of plant growth as affected by a change in root zone temperature in hydroponic chili pepper plants. They used a non-linear autoregressive model with exogenous input for this purpose. The performance of the developed model was evaluated by comparing the estimated dynamic response of plant growth values with the observed values. The authors assumed that the applied type of neural network was useful in identifying the complex process analyzed.

The authors of the studies presented in García-Martínez et al. (2020) designed neural network models that were used to estimate the corn grain yield. The authors also analyzed the relative importance of the predictor variables. Crop information was acquired using an unmanned aerial vehicle. The results presented in the paper show that the GA-BPNN model was more accurate in evaluating CLQ than the other two models.

In the studies presented by the article of Elavarasan et al. (2020), a hybrid feature extraction procedure to address the feature extraction problem in ML models was proposed. They also pointed out that the method may also be used in combination with artificial neural networks for agricultural applications.

In another review article conducted by Niedbała et al., (2020), all applications and potential applications of non-linear ML algorithms in classical and in vitro-based plant breeding methods have been discussed. Classification of genetic diversity, yield component analysis, yield stability and genotype-environment interaction ( $G \times E$ ) analysis, biotic and abiotic stress assessment, and predicting the outcome of mating designs and hybrid breeding programs through ML algorithms have been presented in

the first part of the review. In the third part of the review, the authors presented the application of ML in high-throughput phenotyping and precision agriculture.

The overall findings of Kamilaris and Prenafeta-Boldú (2018) indicate that CNN constitutes a promising technique with high performance in terms of precision and classification accuracy, outperforming existing commonly used image-processing techniques. However, the success of each CNN model is highly dependent on the quality of the data set used. In another study conducted by Sarma et al. (2022) on twelve types of vegetation for the study, out of which five disease classes of the tomato leaves are categorized using a Convolutional Neural Network (CNN). They have proved that CNN is an effective method in providing automated decision support for classifying the plant leaf disease types through a cloud server that can be accessed using the application farmland and regulation of watering mechanism of crops.

The results of research conducted by Abraham et al. (2020) indicated that ANN is the best approach to predicting soybean harvest area and production, while the classical linear function remains more effective in predicting soybean yield. Furthermore, ANN presents a reliable model to predict time series and can help the stakeholders anticipate the world soybean offer. Finally, the study shows that ANN can be effective in predicting a commodities index, even using a short time series, which in this case was 50 years

The findings of a study written by Wang et al. (2020) proposed a weakly supervised paddy rice mapping approach based on a long short-term memory network and dynamic time-warping distance. Standard temporal Synthetic Aperture Radar (SAR) backscatter profiles for each land cover type were constructed. Weak samples were after that labeled based on their DTW distances to the standard temporal profiles. A time-series feature set was at that point created that combined multi-spectral Sentinel-2 bands and Sentinel vertical received band. Experiments showed that weakly supervised learning outperformed supervised learning in paddy rice identification when field samples were insufficient. With only 10% of field samples, weakly supervised learning achieved better results in producer's accuracy (0.981 to 0.904) and user's accuracy (0.961 to 0.917) for paddy rice. Training with 50% of field samples also presented minor improvement.

In another article by Pentoś et al. (2021), the relationship between soil electrical parameters (apparent soil electrical conductivity and magnetic susceptibility measured at two depths, 0.5 and 1 m) and soil compaction was investigated with the use of ANN. The authors highlighted that both electrical parameters, and should be used in the process of the delimitation of management zones. This approach can be employed in precision agriculture, e.g., for the reduction in fuel consumption by variable-depth tillage technology based on soil compaction mapping.

A research article by Pérez-Pérez et al. (2021) describes the challenges date palm growers face in manual processes such as picking, grading, and packing. The authors highlighted that the main challenges are misclassification of dates, maturity level, size, and texture. This article assesses the CNN architectures: VGG-16, VGG-19, ResNet-50, ResNet-101, ResNet-152, AlexNet, Inception V3, and CNN from scratch, considering the learning transfer technique for their training and five hyperparameters to reduce time and improve accuracy in the selection process. The data set they used contains 1002 images corresponding to mature and immature dates in trays. The main results showed that the model obtained from the VGG-19 architecture with a batch of 128 and the Adam optimizer with a learning rate of 0.01 presented the best performance with a precision of 99.32%. Therefore, the authors suggest the possibility of using the VGG-19 model to build computer vision systems that help producers improve their classification process.

Cieniawska and Pentos (2021) presented the results of a laboratory study about the average degree of coverage and the coverage unevenness coefficient. The research was conducted with a special spray track machine, which functioned as a self-propelled field sprayer. Three artificial plants as three replicates were placed on the machine's route. Water-sensitive papers were attached to artificial plants to form specific surfaces. Four types of nozzles were selected for testing: standard, air-induction single flat fan, anti-drift, and air-induction dual flat fan. The following parameters of nozzles' work were used for the research: pressure, the height of the boom, spray angle, perpendicular to the ground, and driving speed. The research was conducted with a constant dose of liquid. Based on the research results, it was found that the highest average coverage was obtained for single standard flat fan nozzles and dual anti-drift flat fan nozzles. At the same time, the highest values of unevenness were observed for these nozzles. Inverse relationships were obtained for air-induction nozzles.

In a study conducted by Yarak et al. (2021), an alternative method for oil palm tree management is proposed by applying high-resolution imagery, combined with faster, automatic detection and health classification of oil palm trees. This study used a total of 4172 bounding boxes of healthy and unhealthy palm trees, constructed from  $2000 \times 2000$ -pixel images. Of the total dataset, 90% was used for training, and 10% was prepared for testing using ResNet-50 and VGG-16. Three techniques were used to assess the models' performance: model training evaluation, evaluation using visual interpretation, and ground sampling inspections. The study identified three characteristics needed for detection and health classification: crown size, color, and density.

The research of Shah et al. (2021) presents an alternative to the use of pesticides in agriculture and horticulture for weed control. Weeds are a major yield-limiting factor in food production, and the use of pesticides levies heavy costs on human and environmental health. The proposed autonomous weeding robot uses ANN to detect weeds and plants, following it up with the desired action to restrict the growth of weeds. During the training phase, three types of plants were used to assess the efficiency of the robot in identification and detection tasks. The study reported the highest classification accuracy of 95% for plant specimens and 99% for weed specimens.

The advancement of computer vision in recent years has made it possible to accurately diagnose tomato diseases in a new method. In the realm of computer vision, object detection is a crucial topic. The primary goal of object detection is to precisely pinpoint the area of interest in the image and identify each object's distinct category. In this study, various deep CNNs are integrated with object identification models to create a system that can not only recognize the many tomato diseases but also find the infected areas so that the proper therapy may be applied.

This study provided a strategy for categorizing bean leaf disease and locating and describing a practical network architecture (using hyperparameters and optimization techniques). The research also compared the findings of each architecture's application with each other to determine the optimal architecture configuration for categorizing bean leaf diseases and their outcomes.

The study has evaluated and implemented MobileNet architectures on one public dataset including two unhealthy classes (angular leaf spot disease and bean rust disease)

and one healthy class, the algorithm was tested on 1296 images of beans, and the model was trained using MobileNetV2 architecture to satisfy the classification requirements. To check if we could get faster training times, higher accuracy, and easier retraining, the model was trained using MobileNet under some controlled conditions as MobileNet. The collected findings demonstrated that our MobileNet model provides high classification performance for bean leaf disease, with the proposed model's average classification accuracy above 97% on training datasets and 92% on test datasets for two unhealthy classes and one healthy class, respectively.

## **2.2. Transfer Learning on Plant Disease Detection**

Sagar and Dheeba (2020) used transfer learning to detect 38 plant diseases using VGG16, ResNet50, InceptionV3, InceptionResNet and DenseNet169 and achieved results using ResNet50 with an accuracy of 0.982, precision of 0.94, recall of 0.94 and F1 score of 0.94.

Mukti and Biswas (2019) proposed a CNN model and adopted VGG16, VGG19 and AlexNet to classify plant diseases. Comparing their developed and transfer learning models their report shows the newly developed model provided 99.8% accuracy. Furthermore, (Bir et al., 2020) reported their study on the application of transfer learning in farming activities to classify nine types of diseases, they employed pre-trained EfficientNetB0, MobileNetV2 and VGG19 architectures for which they achieved a promising potential of using in mobile devices.

Shrivastava et al. (2019) reported their work on rice plant disease classification using transfer learning, using a pre-trained deep neural network as a feature extractor and a Support Vector Machine (SVM) as a classifier to identify rice diseases belonging to four classes and reported that their results encouraging as it facilitates early detection of plant diseases.

Chen et al. (2020), investigated DL with the approach of transfer learning to solve the rice disease classification problem. They proposed using the merits of both, the DenseNet pre-trained on ImageNet and Inception. The method has shown a greater performance and achieved a classification accuracy of at least 94.07% in the public dataset. Similarly, (Chen et al., 2020) studied transfer learning of the deep CNN for the identification of plant leaf diseases and considered using the pre-trained model learned from the large datasets and adapted it to the specific task using their data. In their

approach, VGGNet and Inception models were adopted. The weights adopted from the pre-trained models were used instead of training new models from scratch. The study presented a significant performance improvement compared to other contemporary approaches achieving a minimum validation accuracy of 91.83%. The results indicated the robustness of the methods to detect plant diseases, and even in complex image background cases the classification accuracy reached 92%.

Jiang et al. (2020), applied public data on plant disease to train the ResNet model. In their study, a transfer learning approach was applied to ResNet to develop a plant disease identification model that has good adaptability and efficiency. The experimental study obtained resulted in 83.75% classification accuracy higher than that of the traditional ResNet-101 model. Their experiment indicated; that the transfer learning approach is highly practicable in plant disease detection. Additionally, their research to detect tomato disease (Ouhami et al., 2020) aimed to develop an efficient ML algorithm using color images. DL models DensNet, 161 and 121 layers and VGG16 were considered by adopting transfer learning. The training was based on images of infected plant leaves six pest and disease classes. The classification accuracies were, 95.65%, 94.93% and 90.58% for DensNet161, DensNet121 and VGG16 respectively. In their work to identify plant-leaf diseases using CNNs and transfer learning, (Hassan et al., 2021) developed models trained with a public dataset comprising 14 different plant species, and 38 categories of diseased and healthy plants.

The classification accuracies of the trained models were 98.42%, 99.11%, 97.02%, and 99.56% with InceptionV3, InceptionResNetV2, MobileNetV2, and EfficientNetB0, respectively. The accuracies outperformed traditional techniques. Overall, the study showed the potential of these kinds of models to enhance disease detection in the agricultural sector.

Tunio et al. (2021) proposed the application of hybrid deep CNN transfer learning for the identification and classification of different rice diseases. Having a secondary data source, and employing Transfer Learning trained a model which its test accuracy was. The proposed model is 90.8%. The study demonstrated the techniques were practical, and it has the potential to be used as a disease detection mechanism.

In their work to detect disease in bell peppers using images from the public dataset (Thakur et al., 2021) developed three popular CNN architectures, which are

InceptionV3, ResNet50 and VGG16 adopting the transfer learning method. Comparing their results, VGG16 performed with an accuracy of 99.72% accuracy, ResNet50 had 99.31% accuracy and InceptionV3 showed the least of three with 95.77% classification accuracy. Overall, the study showed considerable improvement in the area of plant disease detection.

Menon et al. (2021), examined a multi-class classification model of several plant diseases over two different datasets. A transfer learning by applying several pre-trained models such as VGG16, Xception, InceptionV3, ResNet152, and MobileNetV2. Out of the trained models, for both training and validation sets, VGG16 and MobileNetV2 architectures achieved the highest accuracy with an accuracy of 92% for each set. Likewise, (Moyazzoma et al., 2021) used a pre-trained feature extraction method that is MobileNetv2 which is useful for mobile devices, the model obtained a validation accuracy of 90.38%.

Moid and Chaurasia (2021), proposed a work to develop a model based on transfer learning for disease classification the pre-trained architecture they adopted was Xception which has the potential to identify and classify diseases. In addition to the Xception model, the research also focused on InceptionV3, related to the former model. Because of their low computational requirements, the models are advantageous. After training using over 70,000 images the accuracy was 97.5% for 38 diseased and healthy categories. (Jiang et al., 2021) aimed to leverage the potential of VGG16 for multi-task learning in the scope of transfer learning. After utilizing and training the power of the model pre-trained on ImageNET, the model achieved a classification accuracy of 97.22% and 98.75% for rice and wheat leaf diseases respectively. Comparing it with single-task approaches, it was proved that multi-tasking was much more effective.

In their comparative study to detect tomato leaf disease, (Peyal et al., 2021), following the transfer learning technique compared popular CNN models namely VGG16, VGG19 and InceptionV3. Employing these models for a dataset of 11000 images categorized into 9 groups, Inception was found to be effective and best in terms of accuracy. Also, a study by Thangaraj et al. (2021), proposed a transfer learning-based deep CNN model to identify tomato leaf disease. The model used real-time images and stored tomato plant images.

In a study to automatically detect tomato diseases (Khasawneh et al., 2022), used deep transfer learning in the classification of nine tomato diseases including the healthy state. The method in this study inputted leaf images, into the CNN models which require no preprocessing and feature extraction. Besides, the algorithms are based on transfer learning of popular architectures. The results were intensively assessed using numerous tactics of data splitting and several metrics. Furthermore, the experimentations were replicated 10 times to test randomness. And the average accuracy was 99.4%.

Paymode and Malode (2022), aimed to develop a model to detect grape and tomato leaf diseases. A transfer learning technique was adopted and a popular network in which VGG was fine-tuned using crop leaf images. Taking accuracy precision, recall and F1-score performance metrics the proposed techniques achieved 98.4% and 95.71% classification accuracy for grapes and tomatoes respectively. Recently (Saeed et al., 2023), identified tomato leaf diseases by classifying healthy and diseased tomato leaf images using transfer learning techniques and applying widespread architectures namely InceptionV3 and ResNetV2. The models were trained with a total of 5225 images including public and custom images. The models were examined with different dropout rates, The Inception ResNetV2 model with a 15% dropout rate, achieved the best performance with an accuracy of 99.22%. The high efficiency indicates the potential of applying CNN models for identifying tomato diseases both in laboratory and field settings.

### **3. MATERIALS AND METHODS**

#### **3.1. Image Acquisition**

In the process of capturing and collecting the images of diseased tomato leaves, deciding which types of disease to consider was a challenging phase. To get the background and information regarding prevalent tomato leave disease the Horticultural Department of Ondokuz Mayıs University was contacted. According to their guidance, precisely knowing the types of plant disease is a highly subjective case as it depends on many factors such as agroecological, climatic and soil characteristics.

Following this, farmers and vegetable growers in Antalya were contacted and interviewed. As per the experience of the farmers, although there are common tomato diseases they highlighted that the prevalence is seasonal and highly variable. Moreover, they added that the internal temperature and moisture of greenhouses highly affect the occurrence of diseases. In the Mediterranean region, well-known tomato diseases, include tomato leaf diseases tomato wilt, tomato yellow leaf curl, early blight, tomato torrado virus, tomato infectious yellow, tomato yellows, tomato mosaic, and tomato mosaic virus. Having the information from the farmers, the image acquisition focused on three common tomato leave diseases namely early blight, yellow leaf disease and Powdery mildew. Along with these three disease types, images of healthy leaves were also captured to create a reference class or label in developing the detection model as the goal of the research is to train a model that recognizes and classifies healthy and diseased leaves.

Depending on the time and resource limitations, the image collection focused on two greenhouses in Antalya. The length, width and height of the greenhouses are 48 m, 18 m and 5 m respectively. The plant-to-plant and row-to-row distances of the tomato plants are 50-60 cm and 85-90 cm respectively. The greenhouses are located in Altıayak, 07210 Kepez/Antalya-Türkiye with coordinates 36.9865° N, 30.7478° E on 29th of October 2023. All images were captured using a mobile phone HUAWEI P40 lite, Model JNY-LX1 RAM 6. GB. As the aim of the study is to detect diseased tomato leaves, four classes were considered out of which the three are diseased and the fourth one a healthy class. In Figure 3.1. samples of the four classes have been provided.

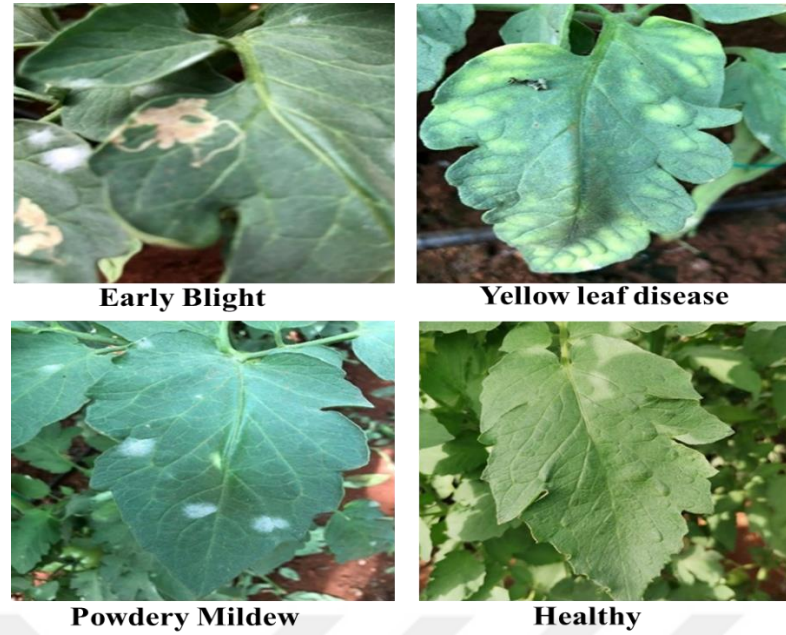


Figure 3.1. Sample images of the four classes

The total images were categorized into training, validation, and test datasets in the process of model training as presented in Table 3.1. The proportion of Training, validation and test datasets were 75%, 10% and 15% respectively.

Table 3.1. Proportion of training, validation and test datasets in four classes

| Class          | Training | Validation | Test |
|----------------|----------|------------|------|
| Early blight   | 229      | 32         | 68   |
| Yellow Disease | 393      | 56         | 112  |
| Powdery mildew | 362      | 50         | 101  |
| Health         | 378      | 57         | 113  |
| Total          | 1362     | 195        | 394  |

### 3.2. Image Pre-processing

Image preprocessing is an important step before an image-based model such as CNN is trained. The first basic preprocessing step is resizing images to a suitable size. It substantially helps to increase the efficiency of the model training process as it significantly reduces the data points a model has to iterate during its training. Besides, after a model is trained and implemented prediction of unseen data is fast for smaller-sized images. In this study, the original images captured in different sizes were resized to 250\*250 as presented in Figure 3.2.

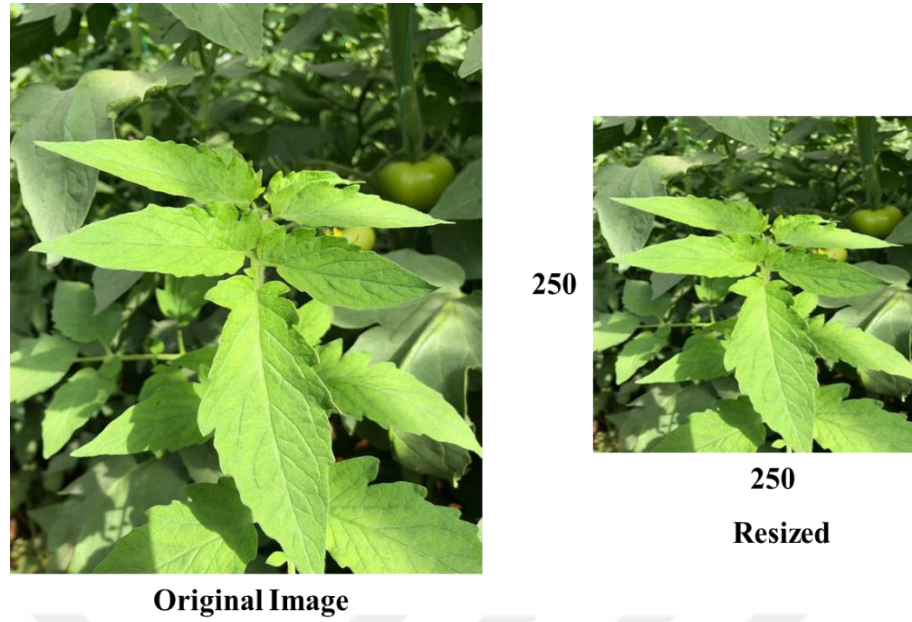


Figure 3.2. Original and resized images

Moreover, data augmentation was applied to increase the size of training data, because the number of images captured are considered as small dataset to train models. Thus, as a data augmentation technique, we adopted  $15^\circ$  random rotations and flipping horizontally as represented in the example given in Figure 3.3. Originally around 1800 images were captured, after manual selection and removal of low quality, shadowed images and much deviant images as shown in Figure 3.4 the number of images decreased to 1270 belonging to four classes. After data augmentation, which is horizontal flipping and rotation 679 images were added and the total number of images increased to 1951.

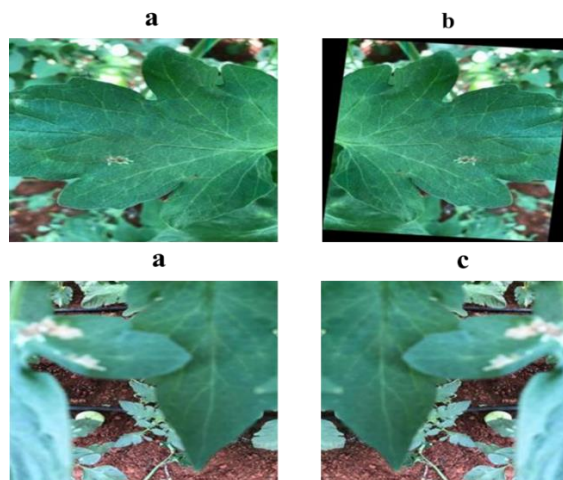


Figure 3.3. (a) Original image, (b)  $15^\circ$  rotated and (c) horizontally flipped



Figure 3.4 Six categories of images removed from the dataset

### 3.3. Transfer Learning Models

Transfer learning is associated with the utilization of CNNs, which are a type of ML model designed for image recognition and classification. These networks have gained widespread adoption due to their ability to directly learn powerful features from input images, eliminating the need for manual feature extraction. In agricultural contexts, such as identifying crop species and diseases, CNNs demonstrate superior performance compared to traditional methods that rely on feature extraction (Szegedy et al., 2015; Cetinic et al., 2018; Mohanty et al., 2016; Dyrmann et al., 2016). As it is presented in the Figure 3.5. , a typical CNN architecture mainly consists of convolution, pooling, and fully connected layers.

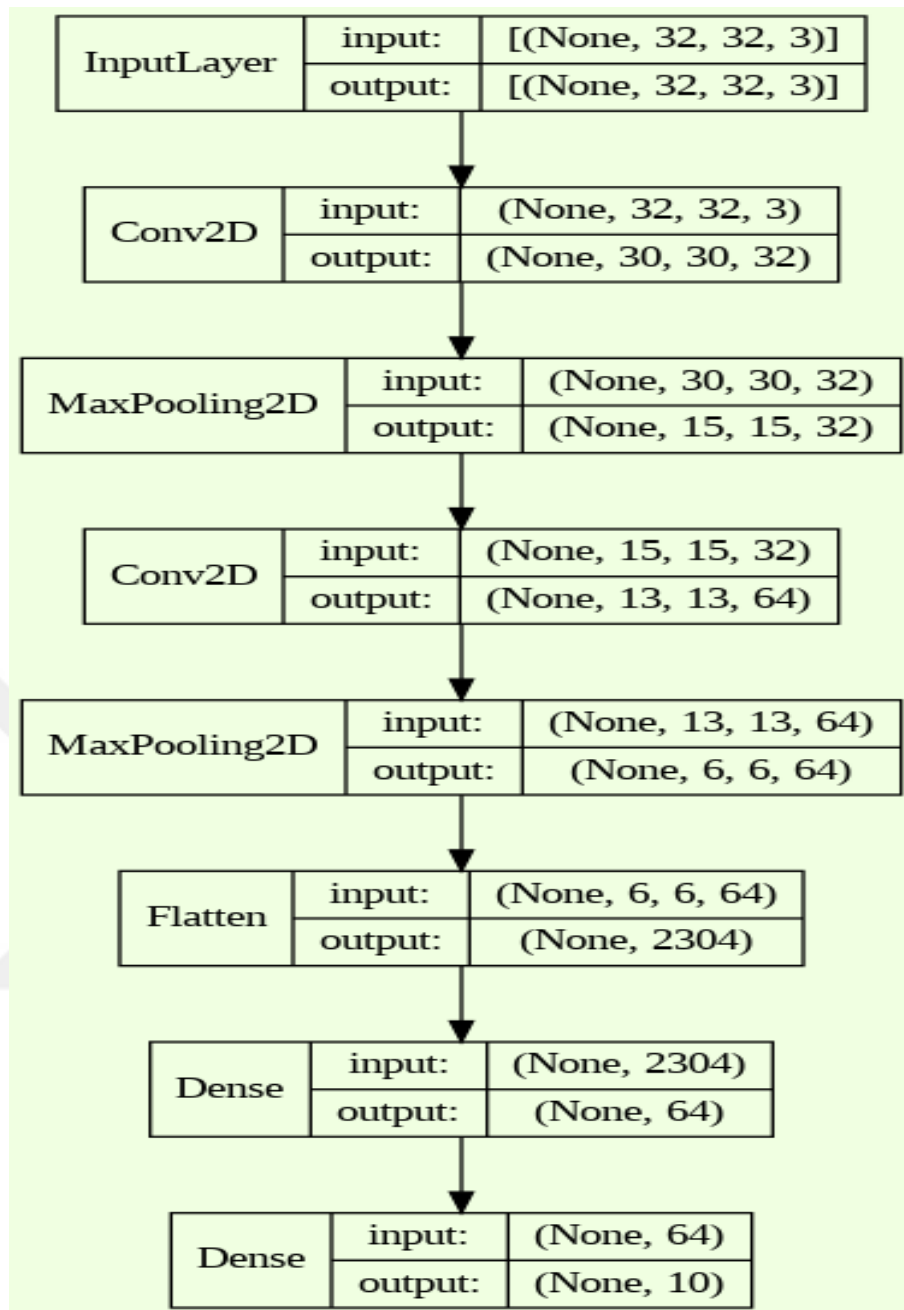


Figure 3.5. Typical CNN architecture

Generally, the training of CNNs is performed using large amounts of data. Their deeper and interconnected layers are the reason for their good performance. Nonetheless, CNNs are often trained on a small dataset, which can simply result in the overfitting of the networks, leading to poor and unconvincing performances. Likewise, training deep CNNs contributes to significant weaknesses, namely high processing costs and slow-running processes. Thus, to solve the aforementioned difficulties, transfer learning is widely adopted. Transfer learning is a DL approach that utilizes the information acquired from a well-established model to reinstate a new problem. CNNs

are trained using vast amounts of data, and their effectiveness originates from their deep and interconnected layers. However, CNNs are frequently trained on limited datasets, which inherently increases the risk of overfitting which results in poor generalization. Additionally, training deep CNNs incurs notable drawbacks, including high computational expenses and slow training speeds. To address these challenges, transfer learning is commonly employed. Transfer learning involves leveraging knowledge gained from a pre-trained model to address a new problem. Thus, in this study, the application of transfer learning, by selecting popular architectures has been investigated using data captured from greenhouses to detect tomato leaf diseases. As their background and structures are stated in the following sections (Figure 3.6 -Figure 3.11), the weights of the five pretrained architectures were adopted. Two dense layers containing 256 neurons each were added, and all training was implemented for an equal number of epochs as given in Table 3.2. To prevent overfitting, model checkpoints and precision-recall backs were implemented in which validation loss was monitored and the best model was saved.

### 3.3.1. VGG

VGG is a CNN model one of the most widespread models developed by Simonyan and Zisserman (2015). It is one of the most popular models trained by ImageNet for a classification task of 1000 classes. VGG includes two typical models known as the VGG16 and VGG19, which are available as the pre-trained models, with the 16 and 19 layers as shown in Figures 3.6 and 3.7, trained on the ImageNet (Russakovsky et al., 2015) dataset. Thus, in our experiment, both architectures were considered via transfer learning approach.

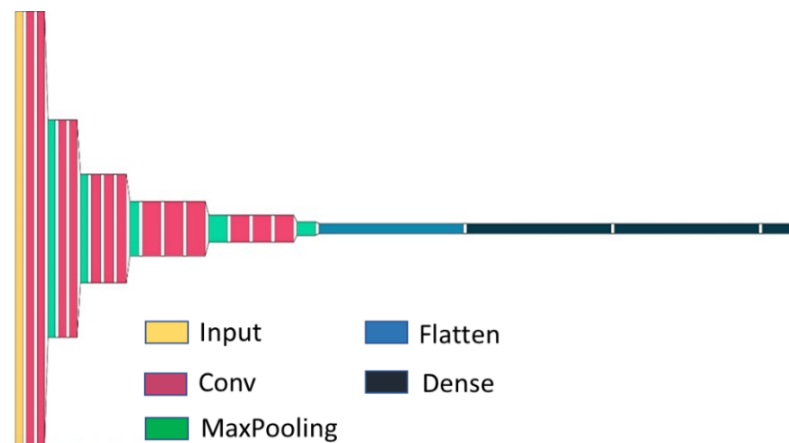


Figure 3.6. VGG16 Architecture

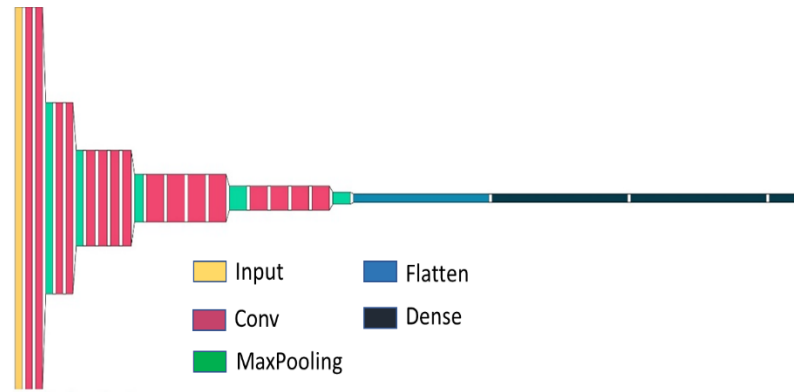


Figure 3.7. VGG19 architecture

### 3.3.2. InceptionV3

InceptionV3 is a CNN that is 48 layers deep belonging to the Inception family developed by Szegedy et al. (2016). The potential of the model which learned a wide range of feature representation for 1000 classes can be leveraged to use transfer learning technology. Due to its potential, InceptionV3 is widely adopted in agricultural applications such as disease detection and varieties classification (Xiaoling et al., 2017). The visualization of the different layers of InceptionV3 is provided in Figure 3.8.

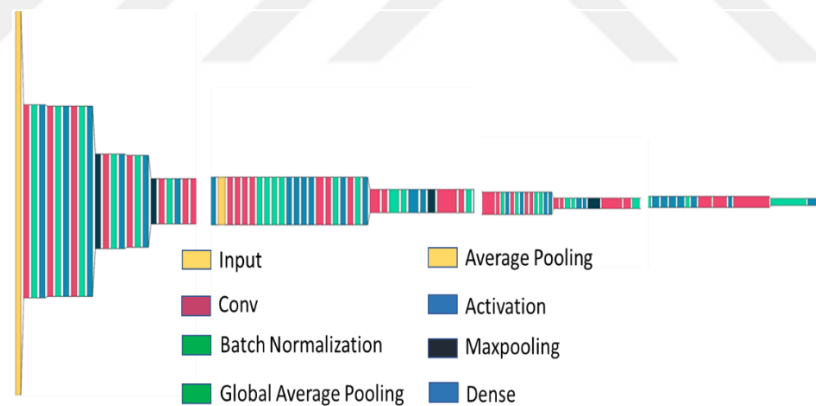


Figure 3.8. InceptionV3 architecture

### 3.3.3. MobileNet

MobileNet is characterized by depth-wise convolution. Unlike traditional convolutions, it is known for its computational efficiency (Howard et al., 2017). Trained by the ImageNet dataset, it has the ability of effectively classify different objects. As a result, its pre-trained version is usable and it is one of the highly used CNN architectures to solve agriculture-related tasks. The graphical representation of MobileNet is given in Figure 3.9.

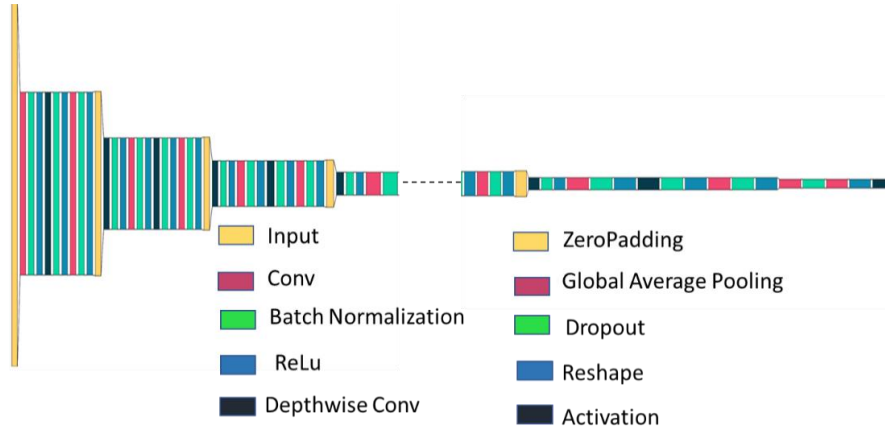


Figure 3.9. MobileNet architecture

### 3.3.4. DenseNet201

A network characterized by its densely arranged layers known as DenseNet201 is another popular CNN architecture characterized by dense connectivity or in other words each layer is connected to every other layer where features from subsequent layers are concatenated and reused unlike standard convolution and pooling mechanism (Huang et al., 2017). In agricultural applications like plant disease diagnosis, DenseNet201 models are widely used (Bakr et al., 2022; Dash et al., 2022). The diagram representation of DenseNet201 has been provided in Figure 3.10. Thus, in our case, VGG16, VGG19, MobileNet and DenseNet201 are adopted by changing the output layer and modifying the last fully connected layer.

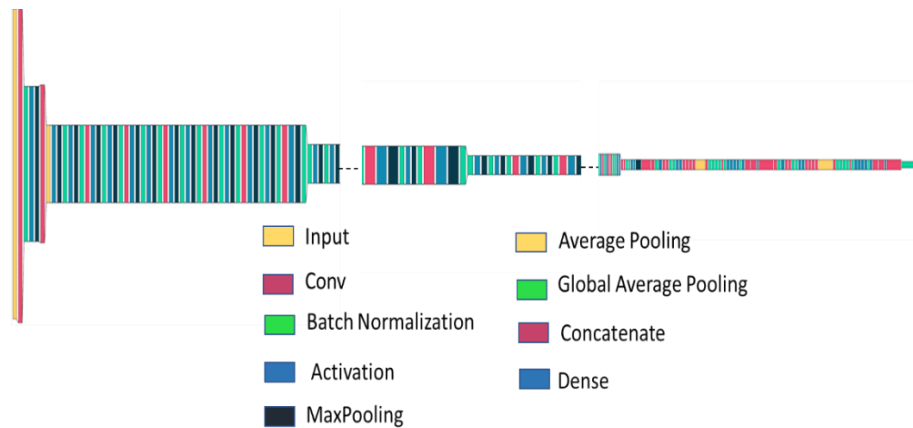


Figure 3.10. DenseNet201 architecture

### 3.3.5. Proposed CNN model

Apart from using transfer learning, a custom model was also trained from scratch to investigate how training a CNN model can handle primary data to detect tomato leaf

diseases. The configuration of the proposed model is as follows. It has four convolutional layers (conv2D), four Max-Pooling layers, one flattened layer, one dropout, and a dense layer of 256 neurons as given in Figure 3.11. The hyperparameters used both in the transfer learning and training of a model from scratch are also given in Table 3.2.

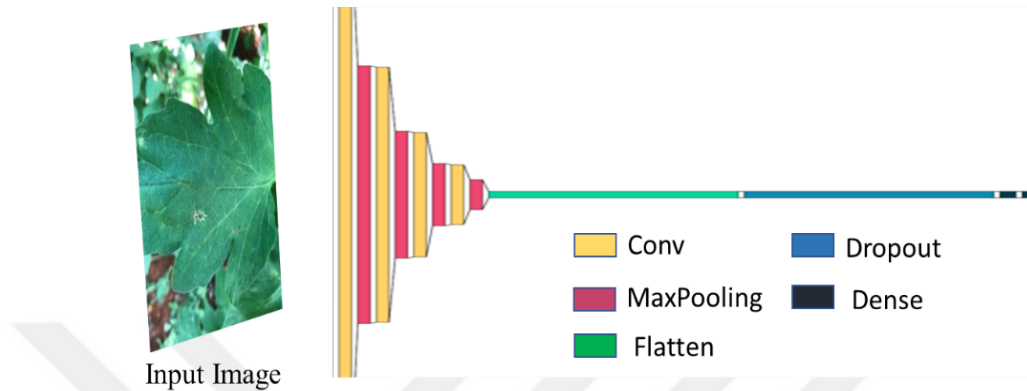


Figure 3.11. The architecture of the proposed CNN model

The hyperparameters used to train all models are given in Table 3.2. Image input shape was fixe to 250\*250, batch size 64, number of epochs 200, a patience call-back of 100 to prevent overfitting while monitoring validation loss that allows to stop training whenever improvement in the validation loss for consecutive 100 epochs, a learning rate of 0.001, loss function of multiclass tasks which is categorical cross-entropy and adam optimizer were used for both the transfer learning and proposed models.

Table 3.2. Training configurations used

| Hyperparameters | Specification             |
|-----------------|---------------------------|
| Input Shape     | 250x250                   |
| Batch size      | 64                        |
| Epochs          | 200                       |
| Patience        | 100                       |
| Learning rate   | 0.001                     |
| Loss            | Categorical cross-entropy |
| Optimizer       | Adam                      |

### 3.4. Performance Metrics

The effectiveness of a model is evaluated using widely used ML model performance metrics given below:

Accuracy: the percentage of correct classification out of the test data. Accuracy is The proportion of accurate predictions made relative to the total number of data points collected, known as accuracy. In the scientific literature, it is known as recognition or success rate. Its expression is given in equation (1).

Precision: is the metric that shows the proportion of true positives out of all the total positive predictions. The mathematical expression of precision is provided in equation (2). Recall: the proportion of true positives compared to actual positive prediction. Its expression is represented in equation (3). F1-score: the weighted harmonic mean of precision and recall were used. Recall is expressed as given in equation (4). The elements of the equation used to measure classification performance are: TP is true positive, TN is true negative, FP is false positive and FN is false negative Figure 3.12 depicts the overall classification model development starting from image acquisition up to performance evaluation by using performance metrics.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1 - Score} = 2 * \frac{\text{Pre} * \text{Recall}}{\text{Pre} + \text{Recall}} \quad (4)$$

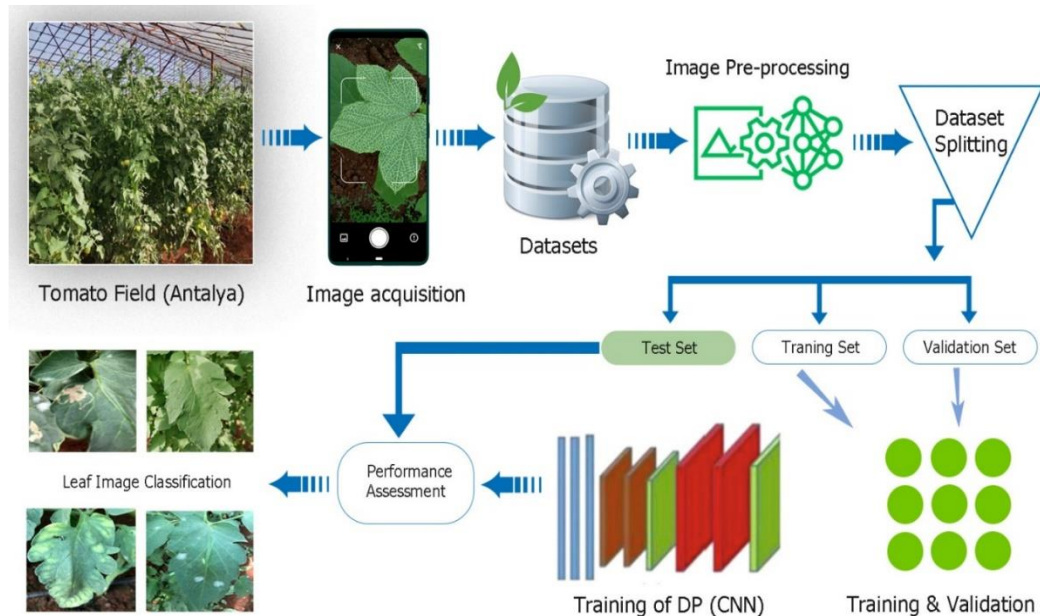


Figure 3.12. Overall process of tomato disease detection model

## 4. RESULTS AND DISCUSSIONS

### 4.1. Results of the Training

#### 4.1.1. VGG16

The model training of VGG16 through transfer learning involved, taking the convolutional base and adding two fully connected layers with 256 neurons each. Figure 4.1 and Figure 4.2 show the training and validation accuracy and training and validation loss respectively, trained for 200 epochs. Training and validation curves are illustrated in Figure 4.1. The number of training epochs is given on the x-axis and the accuracy on the y-axis. The training and validation curves started to plateau from the first 15 epochs and continued up to around 140 epochs. Although the validation curve seems to fluctuate, its overall progress was found and the potential risk of overfitting seems significantly low.

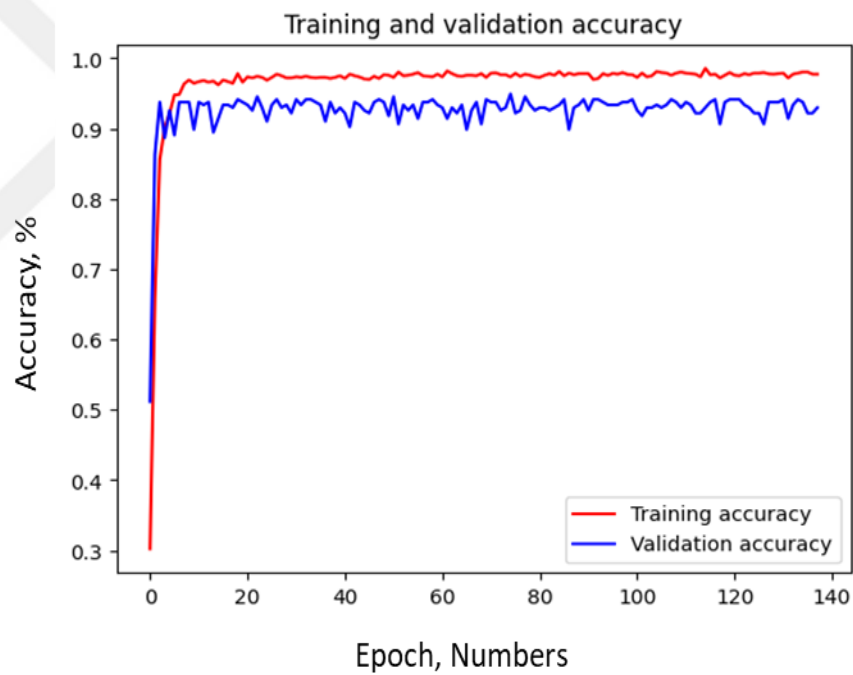


Figure 4.1. Training and Validation accuracy curve of VGG16

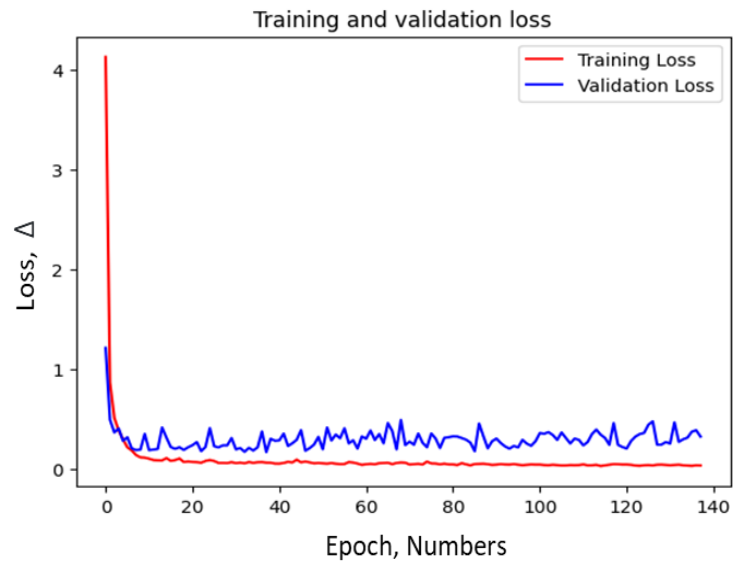


Figure 4.2. Training and Validation loss curve of VGG16

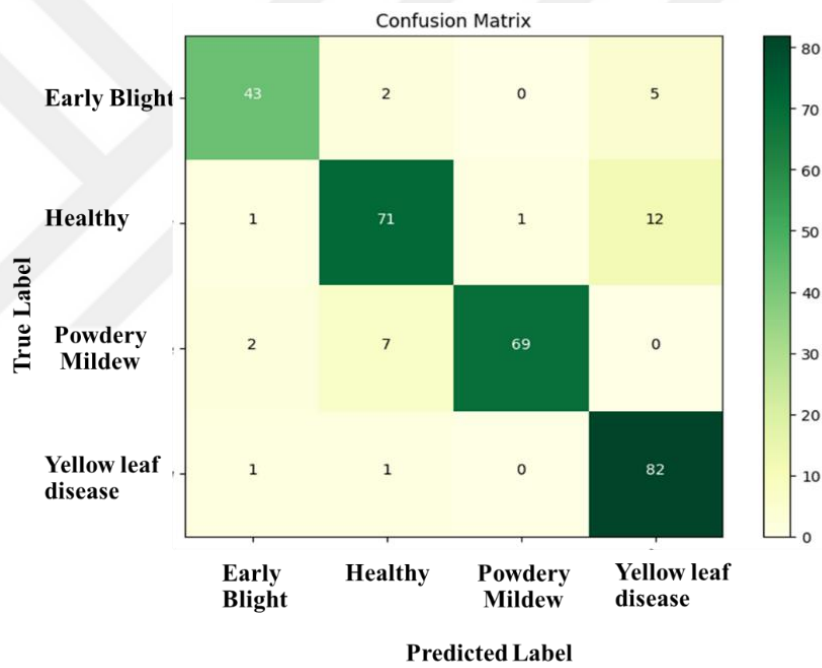


Figure 4.3. Confusion matrix of VGG16  
Table 4.1. Classification Report for VGG16 accuracy 89.23%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early blight        | 0.9149    | 0.8600 | 0.8866   | 50      |
| Healthy             | 0.8765    | 0.8353 | 0.8554   | 85      |
| Powdery mildew      | 0.9857    | 0.8846 | 0.9324   | 78      |
| Yellow leaf disease | 0.8283    | 0.9762 | 0.8962   | 84      |
| Average             | 0.9014    | 0.8890 | 0.8927   | 297     |

The training process of the VGG16 was done by freezing the convolutional base and adding new fully connected 256 neurons and the last output layer of 4 classes. As we can see in the training curve, the training process was let to continue till validation accuracy stopped improving until around the 140<sup>th</sup> epoch, and the test accuracy was 89.23%. As represented in the confusion matrix (Figure 4.3), maximum misclassifications were observed among these classes such as Powdery mildew Healthy and Powdery mildew and Yellow leaf disease. Out of total test datasets 10 were imperfectly classified as healthy and 14 healthy images were misclassified as diseased. From the total healthy instances, fourteen were misclassified as diseased and ten diseased instances (from three diseased labels) were wrongly predicted as healthy. Although the overall accuracy seems promising, after inspecting the misclassifications, it concludes that the model needs further improvement by having additional data. Table 4.1 presents the performance metrics of the VGG16 model in the test data for which precision, recall and F1-score were 90.14%, 88.90% and 89.27% respectively.

#### **4.1.2. VGG19**

For the VGG19 architecture, a similar increase and progress of the training and validation accuracies were observed as shown in Figure 4.4, and its corresponding training and validation loss experienced an identical pattern as shown in Figure 4.5. The training accuracy was greater than the validation accuracy throughout the training epochs starting from the first ten epochs. The validation accuracy curve also experienced fluctuation and drift especially from around the 40<sup>th</sup> epoch. The gap which is greater than in VGG16 suggests there may be a potential of overfitting.

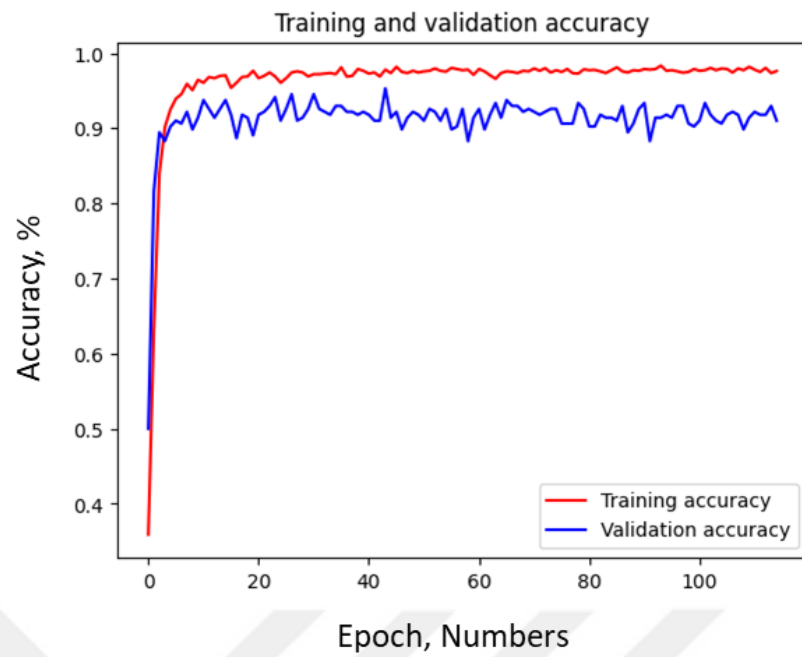


Figure 4.4. Training and validation accuracy curve of VGG19

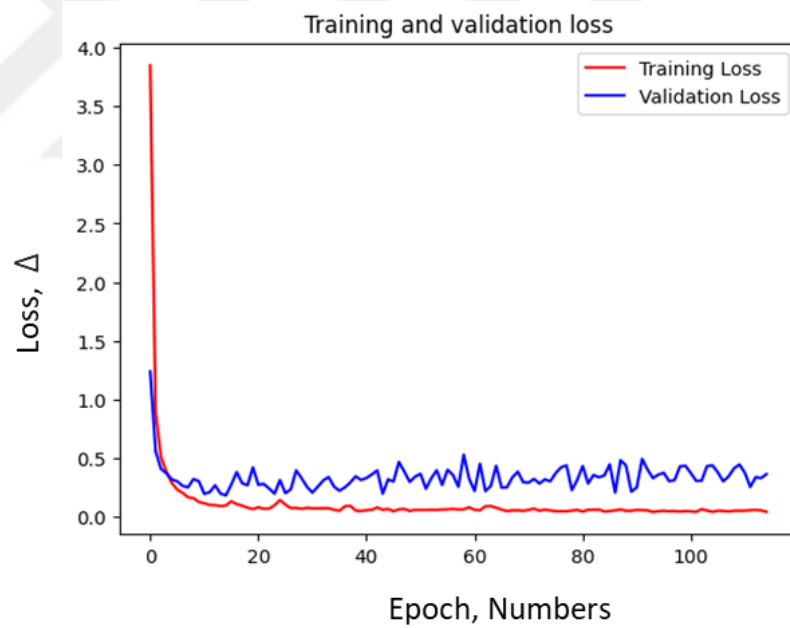


Figure 4.5. Training and validation loss of VGG19

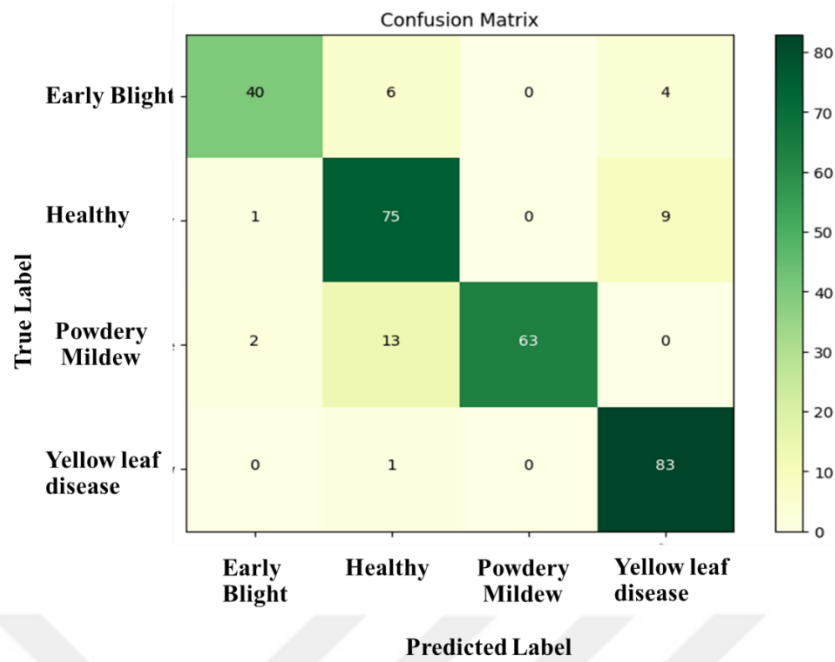


Figure 4.6. Confusion matrix of VGG19  
Table 4.2. Classification report for VGG19, accuracy 87.88%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early Blight        | 0.9302    | 0.8000 | 0.8602   | 50      |
| Healthy             | 0.7895    | 0.8824 | 0.8333   | 85      |
| Powdery mildew      | 1.0000    | 0.8077 | 0.8936   | 78      |
| Yellow Leaf Disease | 0.8646    | 0.9881 | 0.9222   | 84      |
| Average             | 0.8961    | 0.8695 | 0.8773   | 297     |

Being a model of the same family the VGG19 model which only differs from its predecessor VGG16, was trained with a similar technique. The training curves show that the training stopped in earlier epochs than it took for VGG16 for which the classification rate was 87.88%. Examining the confusion matrix (Figure 4.6) for the VGG19 model, in total from three diseased classes, 20 images were misclassified as healthy and 10 healthy images were wrongly labeled as diseased. Comparing the first two models namely VGG16 and VGG19, it shows that, having the advantage of being lighter than VGG19, VGG16 showed relatively better performance. The performance metrics were 89.61%, 86.95% and 87.73% for precision. Recall and F1-score respectively as provided in Table 4.2.

### 4.1.3. InceptionV3

A similar model training of InceptionV3 architecture is also shown in Figures 4.7 and 4.8. The training accuracy was more stable than the validation accuracy which was less stable and showed more variability than the training. The training was much shorter which lasted for 50 epochs where validation loss stopped earlier and the gap was considerably greater than that of VGG16 and VGG19.

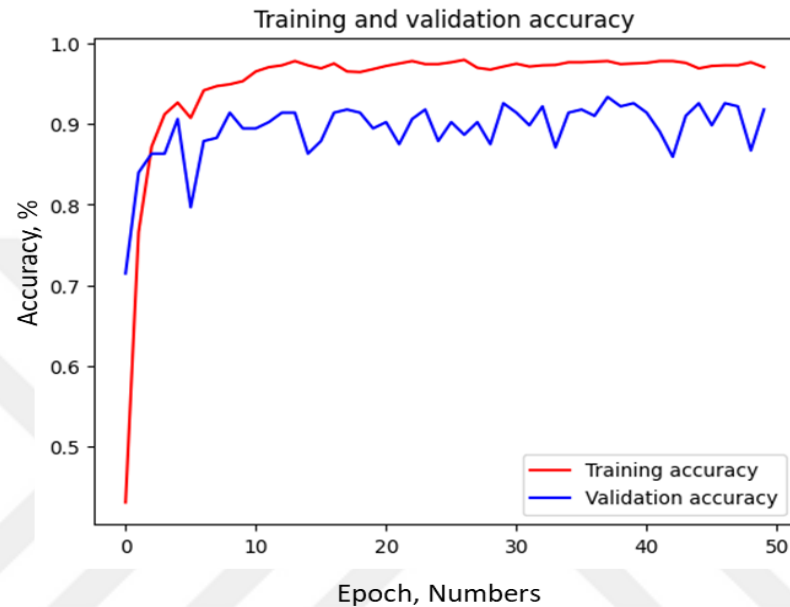


Figure 4.7. Training and validation accuracy curve of InceptionV3

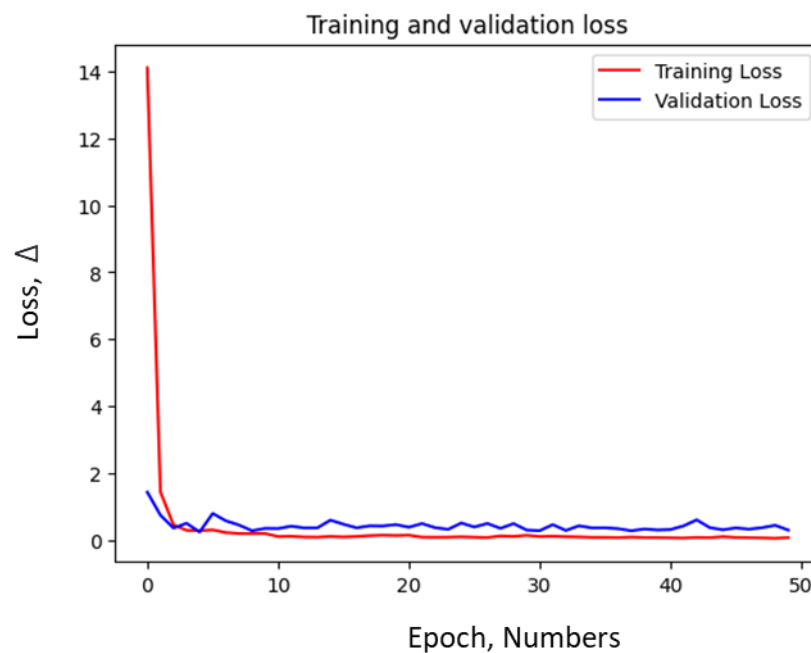


Figure 4.8. Training and validation loss curve of the InceptionV3 model

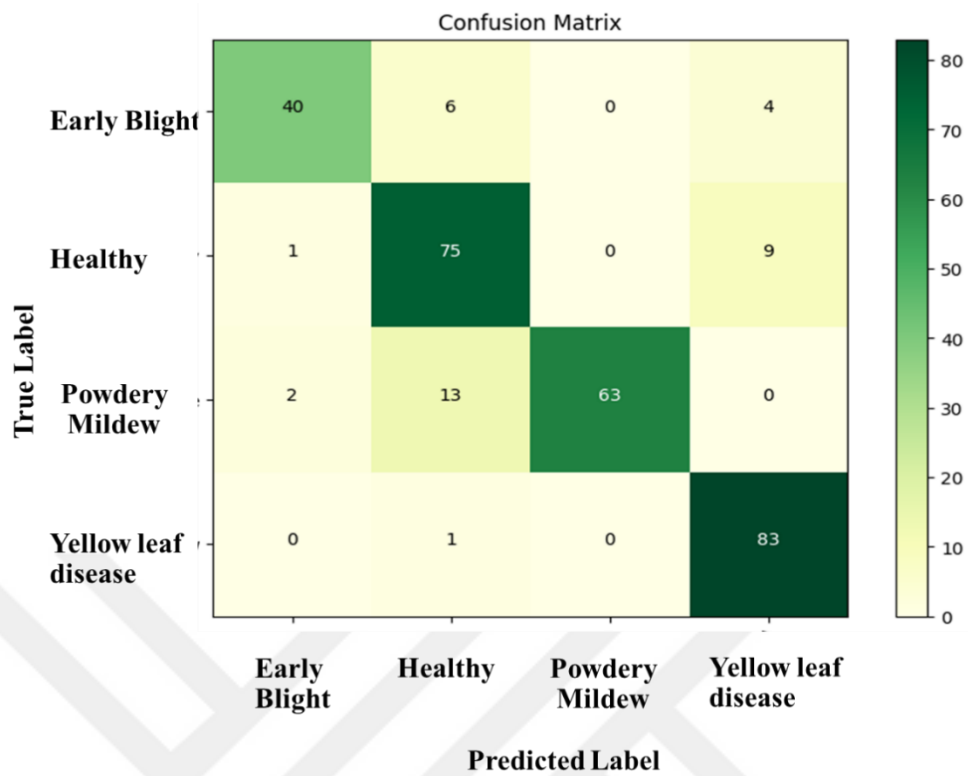


Figure 4.9. Confusion Matrix of InceptionV3  
Table 4.3. Classification report for InceptionV3, accuracy 83.5%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early blight        | 0.9744    | 0.7600 | 0.8539   | 50      |
| Healthy             | 0.7500    | 0.8471 | 0.7956   | 85      |
| Powdery mildew      | 0.9355    | 0.7436 | 0.8286   | 78      |
| Yellow leaf disease | 0.8000    | 0.9524 | 0.8696   | 84      |
| Average             | 0.8650    | 0.8258 | 0.8369   | 297     |

When investigating the InceptionV3 model the training process was characterized by a non-smooth curve and the test accuracy was 83.5%. The training stopped much earlier than VGG16 and VGG19 for a given training epoch. Examining the confusion matrix, the performance was worse than the preceding two models. Out of the total test data, 24 diseased instances were classified as healthy, and 13 healthy instances as diseased (Figure 4.9). Similarly, as shown in Table 4.3 performance metrics of precision, recall and F1-score were 86.5%, 82.58% and 83.69% respectively which are relatively lower values.

#### 4.1.4. MobileNet

The accuracy and loss curves of the MobileNet are given in Figure 4.10 and Figure 4.11. The training lasted upto epoch 110, the gap between training and validation accuracy curves are relatively smaller than each other. The training curve exhibited stabilized progress, whereas the validation curve was variable from epoch 10 up to the 60<sup>th</sup> epoch. However, the gap between the training and validation curves is not significant which indicates the model can generalize fairly well to unseen data.

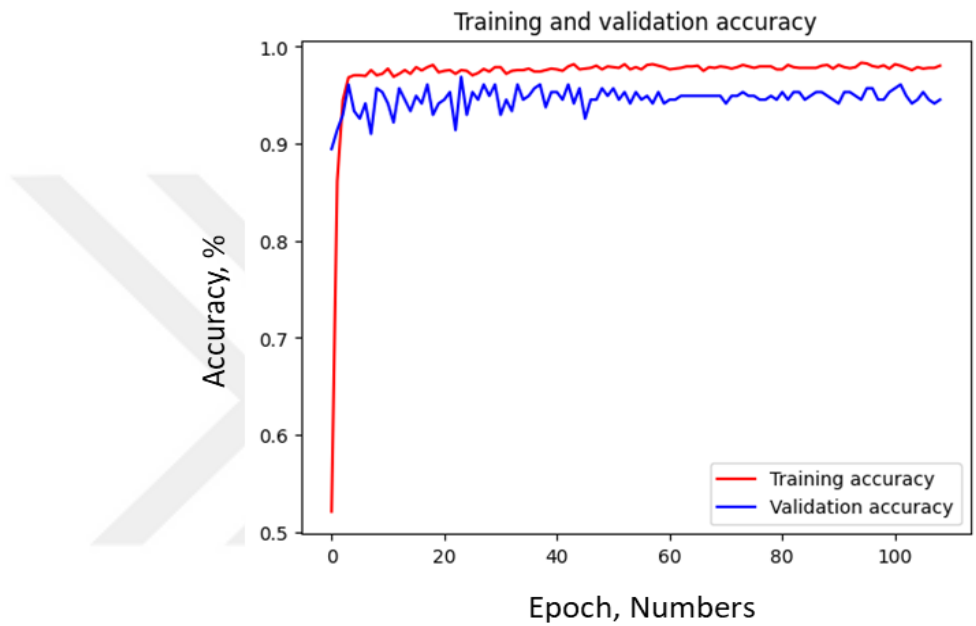


Figure 4.10. Training and validation accuracy curve of MobileNet model

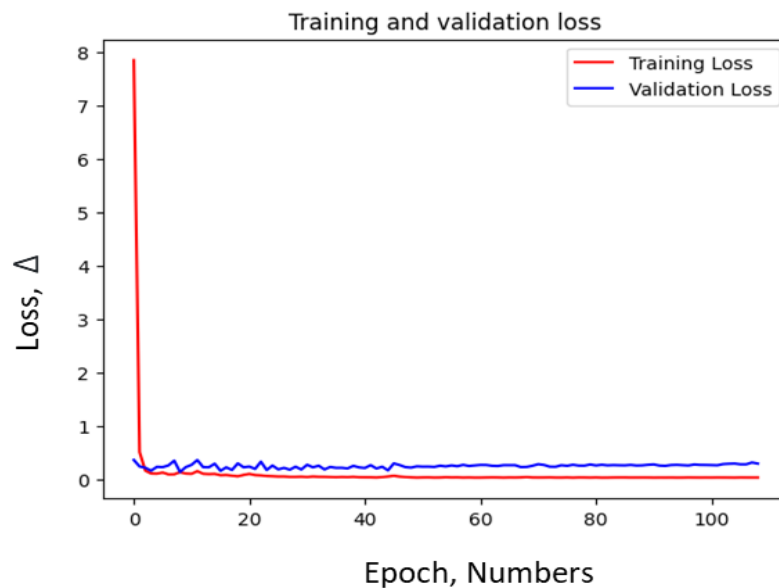


Figure 4.11. Training and validation loss of MobileNet model

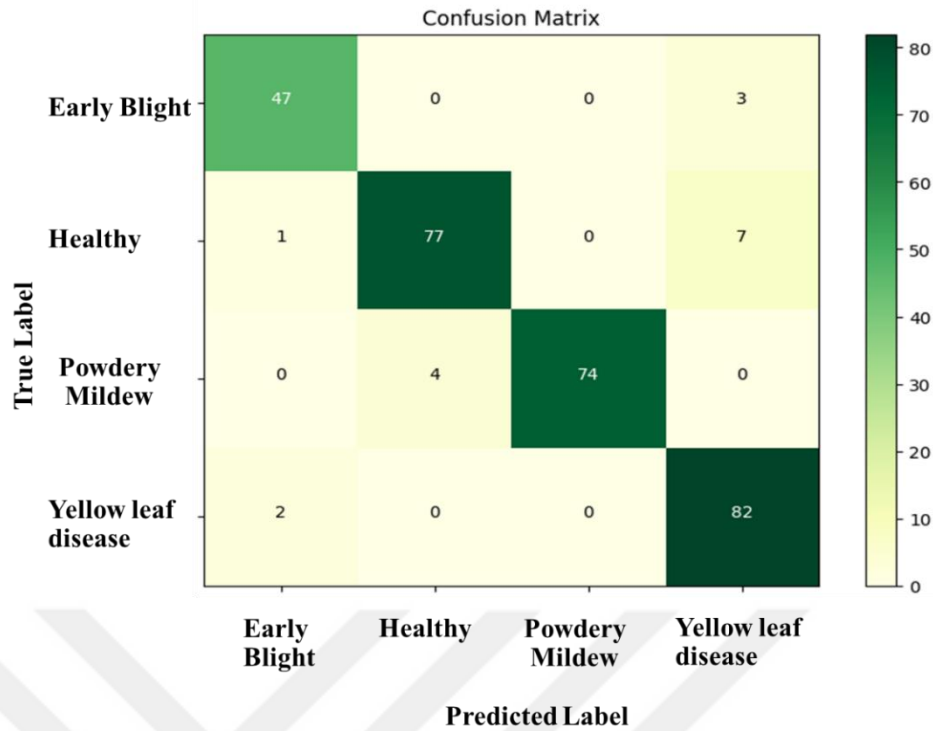


Figure 1.12. Confusion matrix of MobileNet

Table 4.4. Classification report for MobileNet, accuracy 94.28%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early Blight        | 0.9400    | 0.9400 | 0.9400   | 50      |
| Healthy             | 0.9506    | 0.9059 | 0.9277   | 85      |
| Powdery mildew      | 1.0000    | 0.9487 | 0.9737   | 78      |
| Yellow leaf disease | 0.8913    | 0.9762 | 0.9318   | 84      |
| Average             | 0.9455    | 0.9427 | 0.9433   | 297     |

The MobileNet a lightweight CNN architecture designed for mobile applications showed to best accuracy of 94.28%. All the other performance metrics, precision the proportion of true positives of a class among all positive predictions, the proportion of true positives among all actual true positives and F1-score the ratio of the mean of precision and recall were, 94.55%, 94.27 and 94.33% respectively and were all more than 94% which are comparatively much better than all models as shown in Table 4.4. In the confusion matrix, out of 50 early blight instances, 3 misclassifications were observed 1 as healthy and 2 as yellow leaf disease. The class “Healthy” had 85 total instances out of which 7 were wrongly mislabeled as yellow leaf disease and 1 as Early

Blight. From the third class out of 78 Powdery mildew instances 4 were misclassified as healthy and finally, in the fourth class out of 84 total instances, 2 were misclassified as Early Blight as shown in Figure 1.12. Although, there is misclassification among the three diseased classes, taking the whole work as two-class problems of healthy and diseased, only 4 misclassifications were observed. Thus, the efficiency of the model to detect diseased labels is over 99%.

#### 4.1.5. DenseNet201

The accuracy and loss curves of the DenseNet201 model are presented in Figures 4.13 and 4.14 respectively. The curves have relatively shown more variability than previously mentioned models (VGG16, VGG19, InceptionV3, and MobileNet). Especially, the validation curve was not stable throughout the training. The overall training took 70 epochs until it was stopped by an early stopping callback. Although the accuracy seem higher which is more than 90%, the gap suggests a potential overfitting.

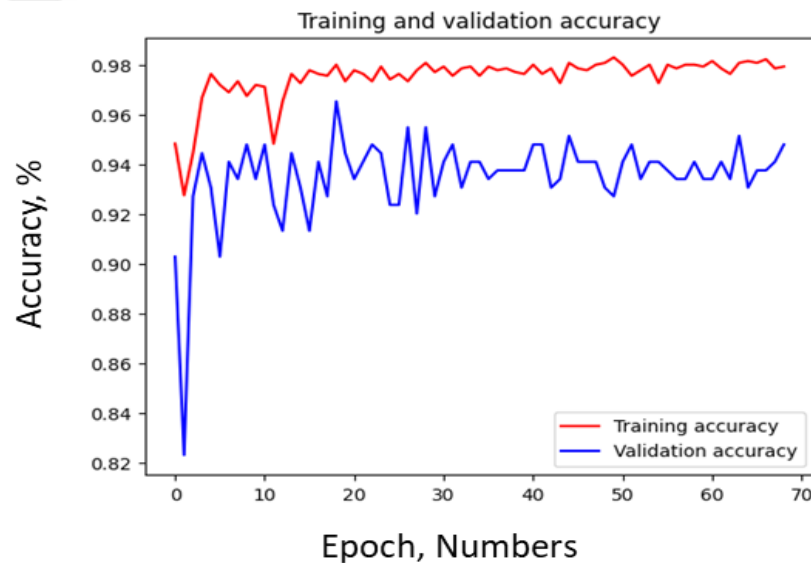


Figure 4.13. Training and validation accuracy curve of the DenseNet201 model

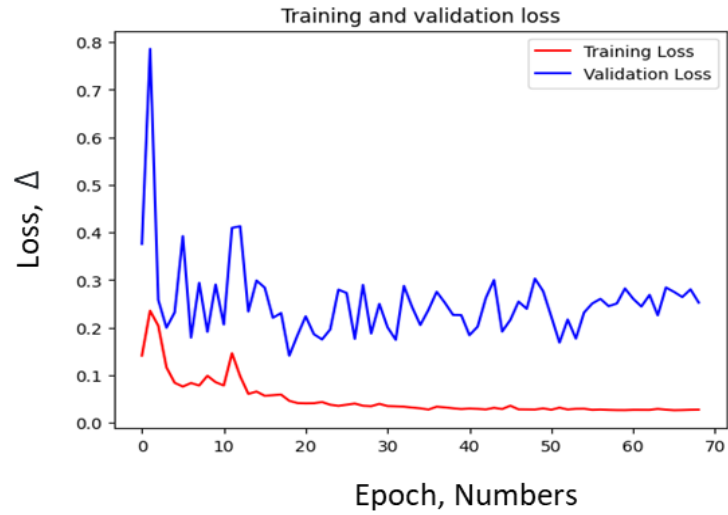


Figure 4.14. Training and validation loss curve of the DenseNet201 model

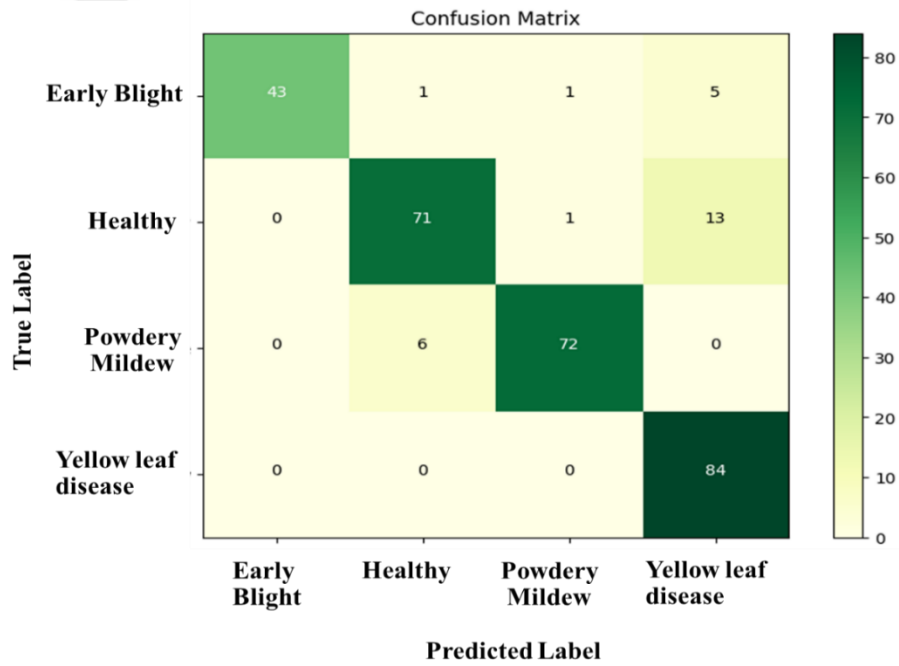


Figure 4.15. Confusion matrix of DenseNet201

Table 4.5. Classification report for DenseNet201 Accuracy 90.91%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early blight        | 1.0000    | 0.8600 | 0.9247   | 50      |
| Healthy             | 0.9103    | 0.8353 | 0.8712   | 85      |
| Powdery mildew      | 0.9730    | 0.9231 | 0.9474   | 78      |
| Yellow leaf disease | 0.8236    | 1.000  | 0.9032   | 84      |
| Average             | 0.9267    | 0.9046 | 0.9116   | 297     |

The DenseNet201 model is characterized by dense layers and as can be seen from the learning curves and thus the process was not as smooth as in the previous models. The validation accuracy took several ups and downs before the lowest accuracy was achieved as per the training configuration. However, the final test accuracy was higher than the VGG16, VGG19 and InceptionV3 models having a test accuracy of 90.91%. The performance metrics namely precision, recall and F1-score were 92.67%, 90.46 and 91.16% respectively as shown in Table 4.5. Examining the classification results, from the Early blight class, one was misclassified as healthy, one as Powdery mildew and five instances as yellow leaf disease. From the second class which is healthy, one instance was mislabeled as Powdery mildew and thirteen instances were incorrectly classified as yellow leaf diseased the confusion matrix shows the actual and predicted results of the model for test data (Figure 4.15). Out of total powdery mildew labels, six instances were misclassified as healthy. And from the final class of yellow leaf disease, all classes were correctly predicted. Moreover, when the system assesses the misclassifications between overall healthy and diseased labels, fourteen healthy instances of the model were labeled as diseased, whereas seven diseased were classified as healthy.

#### **4.1.6. Proposed Model**

The newly proposed model's accuracy and loss curves are also given in Figure 4.16 and Figure 4.17. The training curve increased steadily since early epochs and starting from epoch 25, it was stabilized until the end of the training. The validation accuracy curve has shown variability and drop and ultimately started to drift away until it was stopped by an early stopping call back. Overall characteristics was different from all transfer learning base models as it was trained from scratch with a relatively smaller dataset.

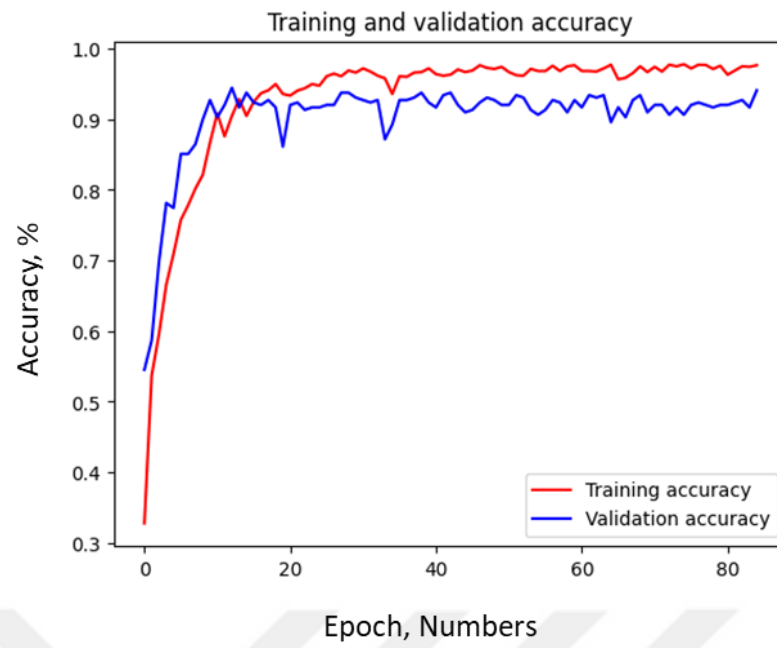


Figure 4.16. Training and Validation accuracy curve of the proposed model

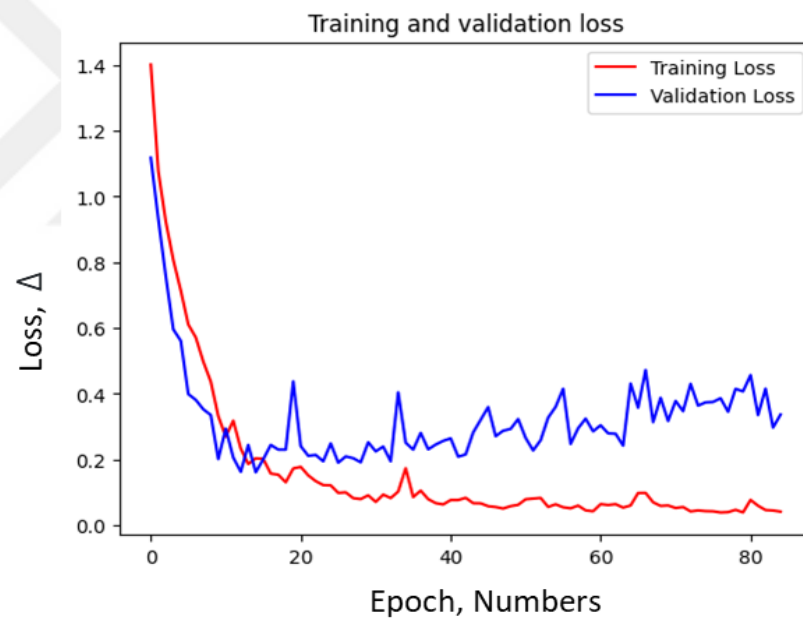


Figure 4.17. Training and Validation loss curve of the proposed model

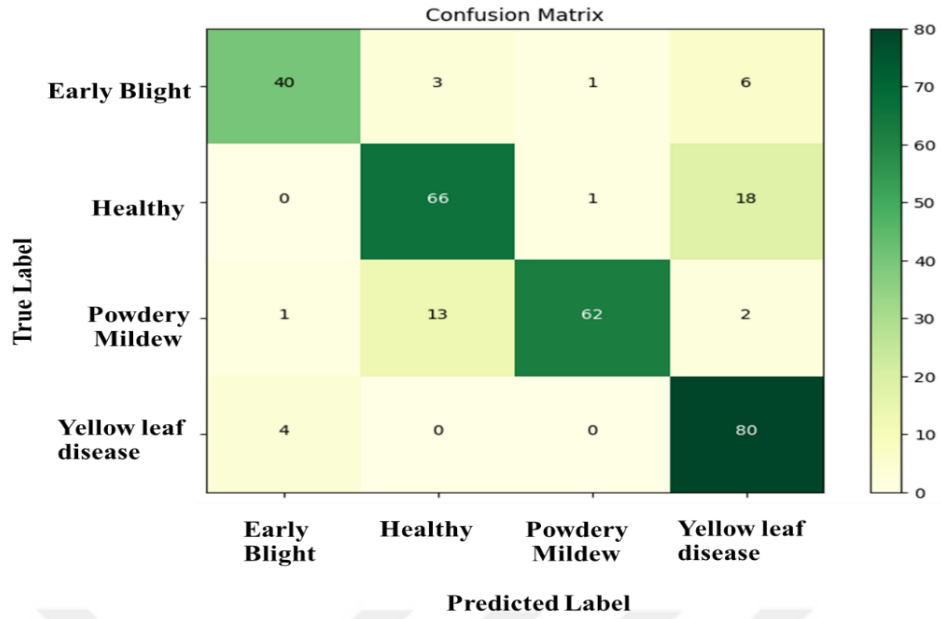


Figure 4.18. Confusion matrix of the proposed model

Table 4.6. Classification report for the proposed model accuracy 83.50%

| Class               | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Early blight        | 0.889     | 0.8000 | 0.8421   | 50      |
| Healthy             | 0.8049    | 0.7765 | 0.7904   | 85      |
| Powdery mildew      | 0.9688    | 0.7949 | 0.8732   | 78      |
| Yellow leaf disease | 0.7547    | 0.9524 | 0.8421   | 84      |
| Average             | 0.8543    | 0.8309 | 0.8370   | 297     |

The newly proposed model having four convolutional layers, four MaxPooling layers, one flattened layer, one dropout, and a dense layer of 256 neurons was also learning curve is was relatively different from the models trained using transfer learning. However, any improvement in accuracy was not observed and the training could not proceed beyond 80 epochs and training stopped at that point. As the model was tested with separate test data, 83.5% accuracy was observed, although it is reasonable for a model trained from zero, it is significantly lower than other models, especially MobileNet. The precision, recall and F1-score performance metrics were 85.435%, 83.09% and 83.70% respectively as shown in Table 4.6. Its confusion matrix (Figure 4.18) shows a total of 49 misclassifications. In the first class of early blight three, one and six misclassifications were observed as healthy, powdery mildew and yellow leaf disease respectively. In the second class which is healthy, one was

misclassified as powdery mildew and eighteen instances were wrongly predicted as yellow leaf disease. In the third class, one thirteen and two mispredictions were observed as early blight, healthy and yellow leaf disease respectively and in the fourth class of yellow leaf disease only four misclassifications were observed as early blight. When we compare the total misclassification between healthy and diseased as a whole, nineteen healthy instances were misclassified as diseased and sixteen healthy instances were misclassified as diseased. The misclassification observed in the confusion matrix of all models might be attributed to the inter-class similarity which may have occurred due to the light conditions, camera angle, distance and other unknown reasons during image acquisition. Table 4.7. shows the summary of all the performance metrics of all the models.

Table 4.7. Performance metrics of all models

| Model          | Accuracy | Precision | Recall | F1-Score |
|----------------|----------|-----------|--------|----------|
| VGG16          | 0.8923   | 0.9014    | 0.8890 | 0.8927   |
| VGG19          | 0.8350   | 0.8650    | 0.8258 | 0.8369   |
| InceptionV3    | 0.8788   | 0.8961    | 0.8695 | 0.8773   |
| MobileNet      | 0.9428   | 0.9455    | 0.9427 | 0.9433   |
| DenseNet201    | 0.9091   | 0.9267    | 0.9046 | 0.9116   |
| Proposed model | 0.830    | 0.8543    | 0.8309 | 0.8370   |

## 4.2. Discussion

It is obvious that plant diseases especially leave disease are the reason for significant production loss and damage. Farmers and producers have been using different mechanism to control disease and minimize loss and decrease in productivity. In line with the quest for sustainability, the aim of this research focused on applying recent technologies such as AI to facilitate early plant disease detection to decrease loss and cost of protection and at the same time increase productivity. As a recap, the results of the study demonstrated that it is possible to develop CNN models to detect tomato leaf diseases having very limited data. The adoption of transfer learning via popular CNN architectures especially MobileNet came up with a remarkable accuracy of 94.2% to recognize and classify tomato leaf diseases. Furthermore, this section interprets and discusses the results of this study with relevant works and trending topics as summarized in Table 4.8.

Research results related to plant disease detection in general and tomato disease detection in particular, focus on the systematic recognition of the target diseases. Firstly, researchers aim to increase the number of prevalent diseases to be detected in the model development method in the purpose of including all possible diseases in the study area as shown in the work of (Karthik et al., 2020) which focused on thirty-eight classes. Second, as part of the model development, image data collection or generation which has a crucial part. Majority of research works, despite of their high accuracies, are based on public data (Jiang et al 2020; Shah et al., 2021). Public data often belong to certain types of plants, diseases and places, and usually suffer from an imbalance among different labels. Third, the researchers' target is to select a suitable model such as popular CNN models to develop a classification recognition model (Jiang et al., 2019).

In connection with the above studies, the scope and results of this study imply that it is within the framework of the research works conducted regarding plant disease detection by using AI tools. Concentrating on, tomato diseases in Antalya and specifically from greenhouses, primary images were captured to create a new tomato disease image dataset. Because AI-based disease detection tools need to be target-focused and models developed for a wider purpose often fail to generalize. Thus, as a primary purpose, this research aimed to develop a more focused and customized application by collecting images of tomato leaves categorized into four classes. To sum up, compared to previously conducted research, although with fewer classes, this strategy, has the potential to provide a targeted solution that is immediately applicable to detect diseases in greenhouses. The performances of all models trained using the relatively small dataset are within the range of high metrics compared to the literature (Jiang et al. 2019; Ouhami et al., 2020). Especially, test accuracies from DenseNet201 and MobileNet, were 90.91% and 94.28% respectively indicating that, although with limited images the results are as high as models trained with larger datasets. This process is evident that datasets tailored to specific tasks can provide accurate and relevant results. It can be correspondingly inferred that custom-made datasets have the potential to overcome data imbalances, noise and irrelevance.

In terms of efficiency, MobileNet was found to be the best model. Its inherent characteristics proved to be effective. Sagar and Dheeba (2020) utilized pre-trained architectures to recognize 38 plant diseases employing VGG16, ResNet50, InceptionV3, InceptionResNet and DenseNet169 and obtained results from ResNet50

with an accuracy of 0.982, precision of 0.94, recall of 0.94 and F1-score of 0.94. In their investigation to detect tomato disease implemented DenseNet and VGG16 through transfer learning (Ouhami et al., 2020). The training depended on images of infected plant leaves six pest and disease classes. The performances were, 95.65%, 94.93% and 90.58% for DensNet161, DensNet121 and VGG16 respectively.

Menon et al. (2021) studied a multi-class classification by training a model that classifies several plant diseases over two different datasets. For both training and validation sets, VGG16 and MobileNetV2 architectures performed well and demonstrated the highest accuracy with an accuracy of 92% for each set. Also, Moyazzoma et al. (2021) used a pre-trained feature extraction method that is MobileNetv2 which is convenient for mobile devices, the model's performance was promising with an accuracy of 90.38%. Hassan et al. (2021) established models trained with a secondary dataset containing 14 plant species and 38 classes of diseased and healthy plants. MobileNet was among the best models having an accuracy of 97.02%

Peyal et al. (2021), employed the transfer learning technique to compare three widely used CNN models -GG16, VGG19, and InceptionV3, to detect tomato leaf disease. Using these models on a dataset of 11,000 images divided into nine categories revealed that Inception was the most accurate and efficient model. Additionally, Thangaraj et al. (2021) suggested a deep CNN model based on transfer learning to detect tomato leaf disease in their work. The model made use of both stored and real-time photos of tomato plants.

Khasawneh et al. (2022) classified nine tomato diseases, including the healthy condition, using deep transfer learning to automatically detect tomato diseases. The technique used in this study involved feeding leaf pictures into CNN models, which do not call for feature extraction or preprocessing. The average accuracy of the results, which were thoroughly evaluated using a variety of data-splitting strategies and measures, was 99.4%. Table 4.8. provides the summary of trending research findings of plant disease detection, what data was used, which model was trained, its accuracy and the respective authors.

Paymode and Malode (2022), aimed to develop a model for detecting leaf diseases in tomatoes and grapes. Using crop leaf images, a popular network VGG was optimized

by applying a transfer learning technique. For grapes and tomatoes, respectively, the methods yielded a classification accuracy of 98.4% and 95.71%.

Freshly, Saeed et al. (2023), used transfer learning approaches and widely used architectures, specifically InceptionV3 and ResNetV2, to identify tomato leaf diseases by classifying photos of healthy and diseased tomato leaves. A total of 5225 photos custom and public, were used to train the models. After testing the models at various dropout rates, the Inception ResNetV2 model with a 15% dropout rate performed the best, with an accuracy of 99.22%. The high efficiency suggests that CNN models can be applied to diagnose tomato diseases in lab and field settings.

Having the advantage of being lightweight and efficient it can be used in real situations where computational resources are limited to make it more accessible to farmers in developing regions as mobile and other handheld devices to automate disease detection in greenhouses to facilitate continuous monitoring of tomato plants through IoT systems. Thus, earlier research has shown how the transfer learning strategy could enhance disease identification in the agriculture industry.

Even though the generalizability of the results is affected by having limited availability of diverse data, variations in lighting conditions, and inability of real-time acquisition and detection and limited time and resources, expanding and supporting previous research endeavors, the findings of this investigation have further validated the suitability of transfer learning. This study, in particular, highlights the effectiveness of the lightweight MobileNet model for application in a specific location in Antalya, even when using a tremendously limited dataset. Through the power of transfer learning, this research not only reinforces earlier findings but also provides additional evidence of how this approach can be tailored to specific agricultural contexts, in this case, and procedure improving disease detection with limited data resources.

MobileNet is a lightweight and efficient model that enables it to converge faster with the available data. Depending on the results of this research it can be said that its inherent simplicity was the reason that enabled it to outperform other models which are deep and complex. In addition, MobileNet has fewer trainable parameters comparing to other models which enables it to generalize better than other models with larger number of parameters. In conclusion, it can be inferred that its efficient architecture and specific

regularizations and the suitability that might have been obtained from the particular data contributed to the overall success of the MobileNet CNN model.

Table 4.8. Trending results of disease detection algorithms

| <b>Disease Type</b> | <b>Data</b> | <b>Models</b>                | <b>Accuracy</b> | <b>References</b>          |
|---------------------|-------------|------------------------------|-----------------|----------------------------|
| Plant diseases      | Public      | DenseNet161                  | 95.65%          | (Sagar and Dheeba, 2020)   |
| Four rice plant     | Field       | Applied Models               | 91.37           | (Shrivastava et al, 2019)  |
| Plant leaf          | Public      | VGG16, ResNet50, InceptionV3 | 92%             | (Chen et al, 2020)         |
| Tomato disease      | Public      | DensNet161, DensNet121       | 95.65%          | (Ouhami et al, 2020)       |
| 38 categories       | Public      | EfficientNetB0               | 99.56%          | (Hassan et al, 2021)       |
| Several plants      | Public      | VGG16 and MobileNetV2        | 92%             | (Menon et al, 2021)        |
| Plant disease       | Field       | MobileNetv2                  | 90.38%          | (Moyazzoma et al, 2021)    |
| Tomato Disease      | Public      | CNN                          | 99.4%           | (Khasawneh et al, 2022)    |
| Tomatoes and Grapes | Public      | VGG                          | 98.4%           | (Paymode and Malode, 2022) |
| Tomato leaf         | Public      | CNN                          | 99.22%          | (Saeed et al, 2023)        |

## 5. CONCLUSIONS AND RECOMMENDATIONS

The existence of disease is uncertain because of the diversity of tomato diseases and pests, consequently, the forms of pesticides are diverse and the cost of treatment is usually higher than that of prevention which adversely results in a significant economic crisis. Thus, early diagnosis and treatment of pests and diseases are prerequisites to treating immediately to prevent the disease from infestation. The research aimed to answer the question of developing a disease-detection AI model with a very limited image data. To address it, leaf images of healthy and three types of diseases were captured and successful models were trained. It can be concluded that having a limited dataset using a transfer learning mechanism makes use of weights of pre-trained algorithms for plant disease detection models that can be tailored for a specific task. The classification accuracy of 94.28%, designates the capability of the lightweight MobileNet model to effectively handle the task of early plant disease detection and subsequently reduce pesticide application frequencies.

Principally detection and prevention can effectively delay the spread of pests and diseases. The application of pesticides is one of the plant protection measures taken to tackle this issue. Nevertheless, with all its benefits and advantages it comes up with certain negative effects. In addition, fertilizers, pesticides, and other chemicals are the main inputs utilized in field cultivation in general and greenhouse cultivation in particular. According to topical studies from the province of Antalya, it is conveyed that pesticides and fertilizers are widely used in greenhouse agriculture. Moreover, farmers overuse fertilizer in greenhouse production, which is the cause of soil salinity issues and nitrate poisoning of underground water bodies. Excessive utilization of these inputs is destructive and hazardous to human health and pollutes the ecosystem. Many academics and researchers claim that Non-destructive identification and early treatment of plant diseases are essential to the growth of precision and ecological agriculture.

In recent years, plant disease detection has been applied technology and internet-based. Due to the fast growth of powerful algorithms and the enhancement of computational ability, AI-based diagnosis is gaining popularity and applicability. The application of CNN models which makes the practice of images as input and capture representation belonging to each label is one of the most efficient models in the computer vision world. Its potential is profiting industries from medicine to automotive and agriculture in numerous ways to control the operation. CNN models are trained in

two strategies, one from scratch and the other through a technique entitled transfer learning. Transfer learning is linked to utilizing pre-trained popular CNN models, which are categories of arranged ML algorithms trained for image recognition and classification. This strategy provides a perfect condition for these models to gain easily and directly learn powerful features from input images, without the need for manual feature extraction, these networks have gained widespread acceptance.

This research presents the development and assessment of disease detection models. The study was applied to greenhouse environments, particularly emphasizing tomato plant leaves. The study's targets were to validate the efficiency of several disease detection methods and their potential to identify plant diseases by making custom-made image datasets. The objectives of this research were, first collecting and creating a dataset of images of healthy and three labels of diseased tomato leaves from a greenhouse situated in Antalya -Türkiye. Second, to train a disease classification model using the custom-made image dataset, and finally to evaluate the performance of transfer learning with popular CNN architectures (VGG16, VGG19, InceptionV3, DenseNet201, MobileNet) and compare transfer learning-based models with custom CNN model.

The data collection process involved the acquisition of images of healthy and diseased tomato leaves (four classes) from a greenhouse. As a result, a relatively small dataset that included pictures of three distinct diseases and healthy leaves was created. Following many significant namely stages collection, preprocessing and labeling of the images was a crucial stage, before further development and analysis. The actual dataset was organized and developed for the disease classification model was carried out using the custom-designed image dataset separated into folders to facilitate model training and testing. A transfer learning strategy was adopted to take advantage of the weights of pre-trained models. Popular CNN architectures, namely VGG16, VGG19, InceptionV3, DenseNet201, and MobileNet, were trained and finally, a custom-made CNN model was trained from scratch. All models were trained for an equal number of epochs, and the training process included steps such as early stopping and saving the best model as checkpoints to ensure optimal performance.

Evaluating the performance metrics, it was demonstrated that significant accuracy differences exist among the models in classifying tomato leaf images. Out of all the

models, DenseNet201 achieved an accuracy of 90.91%, while MobileNet outperformed all other models with a classification accuracy of 94.28%. These results highlighted the potential of the MobileNet models in accurately detecting tomato diseases. During the validation of the models using test data misclassification was observed. The misclassification from different models may be due to high similarity among dissimilar classes, light variations, camera position, and additional unidentified details when the images were captured.

Therefore, it can be inferred that; here the results of the research demonstrated the feasibility of using advanced CNN architectures for disease detection in greenhouse environments locally and globally. More briefly the creation of a labeled database and the development of a custom-image-designed dataset provided a robust foundation to train the model for disease detection. Besides, transfer learning proved to be an effective approach with MobileNet emerging as the most accurate model. Furthermore, the findings highlight the potential of using limited data to train suitable CNN models for early and precise disease detection, which contributes to the increase of efficiency and the creation of eco-friendly production. Thus, based on the recent literature results and the findings of this research, it can be inferred that CNN models appear to be promising and innovative in detecting and recognizing various plant diseases. Besides, with the ever-increasing computational resources and with the widespread application of AI in every aspect of agricultural business, it is obvious that research results that have so far proved the efficiency and applicability of computer vision will be realized sooner both in laboratory and commercial schemes.

As a recommendation, this study suggests that further research needs to be conducted regarding the development of AI-based integration models in real-time monitoring systems in greenhouses. Creating a larger image dataset increases prevalent types of diseases and crops to develop more generalized and inclusive solutions to handle the case. Moreover, applying these models in diverse agroecological zones needs to be studied to gain a valuable understanding of their robustness, especially for monitoring. Additionally, such kind of research requires the cooperation and participation of different stakeholders, such as research centers, agricultural organizations, and farm business enterprises to broaden and deepen the scope of the project and facilitate its implementation in the physical greenhouse industry. Finally, this thesis demonstrates the study works undertaken to investigate the potential of CNN

in general and transfer learning in specific in disease detection in greenhouse settings. It emphasizes the promising role of advanced ML techniques in promoting sustainable agriculture, demonstrating the important impact of early and precise disease detection on farming practices.



## REFERENCES

- Abade, A., Ferreira, P. A., and Vidal, F. de B. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Computers and Electronics in Agriculture*, 185, 106125. <https://doi.org/10.1016/j.compag.2021.106125>.
- Abraham, E. R., Mendes dos Reis, J. G., Vendrametto, O., Oliveira Costa Neto, P. L. D., Carlo Toloi, R., Souza, A. E. D., and Oliveira Morais, M. D. (2020). Time series prediction with artificial neural networks: An analysis using Brazilian soybean production. *Agriculture*, 10 (10), 475.
- Ahmed, A. A., and Reddy, G. H. (2021). A mobile-based system for detecting plant leaf diseases using deep learning. *AgriEngineering*, 3(3), 478-493.
- Aji, G. K., Hatou, K., and Morimoto, T. (2020). Modeling the dynamic response of plant growth to root zone temperature in hydroponic chili pepper plant using neural networks. *Agriculture*, 10(6), 234.
- Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M. and Alrahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications*, 17(1), 31-38.
- Alhnaity, B., Pearson, S., Leontidis, G., and Kollias, S. (2020). Using deep learning to predict plant growth and yield in greenhouse environments. *Acta Horticulturae*, 1296, 425-432. <https://doi.org/10.17660/ActaHortic.2020.1296.55>.
- Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 1-74.
- Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., and Stefanovic, D. (2019). Solving current limitations of deep learning-based approaches for plant disease detection. *Symmetry*, 11(7), 939. <https://doi.org/10.3390/sym11070939>.
- Ayikpa, K. J., Mamadou, D., Gouton, P., and Adou, K. J. (2022). Experimental Evaluation of Coffee Leaf Disease Classification and Recognition Based on Machine Learning and Deep Learning Algorithms. *Journal of Computer Science*, 18(12), 1201–1212. <https://doi.org/10.3844/jcssp.2022.1201.1212>
- Bacco, M., Barsocchi, P., Ferro, E., Gotta, A., and Ruggeri, M. (2019). The digitization of agriculture: A survey of research activities on smart farming. *Array*, 3–4, 100009. <https://doi.org/10.1016/j.array.2019.100009>
- Bakr, M., Abdel-Gaber, S., Nasr, M. and Hazman, M. (2022). DenseNet Based Model for Plant Diseases Diagnosis. *European Journal of Electrical Engineering and Computer Science*. 6, 5 (Sep. 2022), 1–9. DOI:<https://doi.org/10.24018/ejece.2022.6.5.458>
- Barbedo, J. G. A. (2018). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and electronics in agriculture*, 153, 46-53.

- Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, 180, 96-107.
- Bir P., Kuar R., Singh G. (2020). Transfer learning-based tomato disease detection for mobile application. IEEE International Conference on Computing, Power and Communication Technologies (GUCON) : Galgotias University, Greater Noida, UP, India. Oct 2-4, 2020.
- Brahimi, M., Boukhalfa, K., and Moussaoui, A. (2017). Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence*, 31(4), 299-315.
- Briassoulis, D., Waaijenberg, D., Gratraud, J., and Von Eslnr, B. (1997). Mechanical properties of covering materials for greenhouses: Part 1, general overview. *Journal of Agricultural Engineering Research*, 67(2), 81-96.
- Carlini, M., Honorati, T., and Castellucci, S. (2012). Photovoltaic greenhouses: Comparison of optical and thermal behavior for energy savings. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2012/743764>
- Castellano, S., Mugnozza, G. S., Russo, G., Briassoulis, D., Mistrionis, A., Hemming, S., and Waaijenberg, D. (2008). Plastic nets in agriculture: A general review of types and applications. *Applied engineering in agriculture*, 24(6), 799-808.
- Cetinic, E., Lipic, T., and Grgic, S. (2018). Fine-tuning convolutional neural networks for fine art classification. *Expert Systems with Applications*, 114, 107-118.
- Chakraborty, K. K., Mukherjee, R., Chakraborty, C., and Bora, K. (2022). Automated recognition of optical image-based potato leaf blight diseases using deep learning. *Physiological and Molecular Plant Pathology*, 117, 101781
- Chen, C. J., Huang, Y. Y., Li, Y. S., Chang, C. Y., and Huang, Y. M. (2020). An AIoT-based smart agricultural system for pest detection. *IEEE Access*, 8, 180750-180761.
- Chen, J., Zhang, D., Nanekaran, Y. A., and Li, D. (2020). Detection of rice plant diseases based on deep transfer learning. *Journal of the Science of Food and Agriculture*. <https://doi.org/10.1002/jsfa.10365>
- Chen, J., Chen, J., Zhang, D., Sun, Y., and Nanekaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, 173, 105393. <https://doi.org/10.1016/j.compag.2020.105393>
- Chen, J., Zhang, D., Zeb, A., and Nanekaran, Y. A. (2021). Identification of rice plant diseases using lightweight attention networks. *Expert Systems with Applications*, 169, 114514.
- Cieniawska, B., and Pentos, K. (2021). The average degree of coverage and coverage unevenness coefficient as parameters for spraying quality assessment. *Agriculture*, 11(2), 151.
- Czosnek, H., and Laterrot, H. (1997). A worldwide survey of tomato yellow leaf curl viruses. *Archives of Virology*, 142, 1391-1406. <https://doi.org/10.1007/s007050050168>

- Dalai, S., Tripathy, B., Mohanta, S., Sahu, B., and Palai, J.B. (2020). Greenhouses: Types and Structural Components. *Protected Cultivation and Smart Agriculture*.
- Dan, L. I. U., Xin, C., Chongwei, H., and Liang, J. I. (2015). Intelligent agriculture greenhouse environment monitoring system based on IOT technology. *In 2015 International Conference on Intelligent Transportation, Big Data and Smart City* (487-490). IEEE. December
- Dash, A., Sethy, P. K., and Behera, S. K. (2023). Maize disease identification based on an optimized support vector machine using a deep feature of DenseNet201. *Journal of Agriculture and Food Research*, 14. <https://doi.org/10.1016/j.jafr.2023.100824>
- DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E. L., Yosinski, J., Gore, M. A., Nelson, R. J., and Lipson, H. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology*, 107(11), 1426–1432. <https://doi.org/10.1094/PHYTO-11-16-0417-R>
- Dyrmann, M., Karstoft, H., and Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems Engineering*, 151, 72-80.
- Elavarasan, D., Vincent PM, D. R., Srinivasan, K., and Chang, C. Y. (2020). A hybrid CFS filter and RF-RFE wrapper-based feature extraction for enhanced agricultural crop yield prediction modeling. *Agriculture*, 10(9), 400.
- Engindeniz, S., and Ozturk Cosar, G. (2013). An economic comparison of pesticide applications for processing and table tomatoes: A case study for Turkey. *Journal of Plant Protection Research*, 53(3), 230-237. <https://doi.org/10.2478/jppr-2013-0035>.
- Escamilla-García, A., Soto-Zarazúa, G.M., Toledano-Ayala, M., Rivas-Araiza, E. and Gastélum-Barrios, A. (2020). Applications of artificial neural networks in greenhouse technology and overview for smart agriculture development. *Applied Sciences*, 10(11), p.3835.
- FAO. (2021). Climate change fans spread of pests and threatens plants and crops, new FAO study.
- Fuentes, A., Yoon, S., Youngki, H., Lee, Y., Park, D.S. (2016). Characteristics of tomato plant diseases—a study for tomato plant disease identification. *Proc Int Symp Inf Technol Converg.*; 1:226–31.
- García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., and Vázquez-Peña, M. A. (2020). Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicles. *Agriculture*, 10(7), 277.
- Gupta, A., Gonzalez-Rojas, Y., Juarez, E., Crespo Casal, M., Moya, J., Falci, D. R., ... and Shapiro, A. E. (2021). Early treatment for Covid-19 with SARS-CoV-2 neutralizing antibody sotrovimab. *New England Journal of Medicine*, 385(21), 1941-1950.

- Harjunowibowo, D., Ding, Y., Omer, S. A., and Riffat, S. (2018). Recent active technologies of greenhouse systems – A comprehensive review. *Bulgarian Journal of Agricultural Science*, 24(1), 158-170.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
- Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017; (4700–4708).
- Jiang, P., Chen, Y., Liu, B., He, D., and Liang, C. (2019). Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, 7, 59069-59080.
- Jiang, H., Xue, Z. P., and Guo, Y. (2020, June). Research on plant leaf disease identification based on a transfer learning algorithm. In *Journal of Physics: Conference Series*, 1576 (1). IOP Publishing.
- Jiang, Z., Dong, Z., Jiang, W., and Yang, Y. (2021). Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. *Computers and Electronics in Agriculture*, 186, 106184.
- Kamilaris, A., and Prenafeta-Boldú, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science*, 156(3), 312-322.
- Kamilaris, A., and Prenafeta-Boldú, F.X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, (70-90).
- Karmokar, B. C., Ullah, M. S., Siddiquee, M. K., and Alam, K. M. R. (2015). Tea leaf disease recognition using neural network ensemble. *International Journal of Computer Applications*, 114(17).
- Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., and Menaka, R. (2020). Attention embedded residual CNN for disease detection in tomato leaves. *Applied Soft Computing*, 86, 105933.
- Kaplan M, Aktas M, Gunes A, Alpaslan M, Sonmez S 2000. An assessment of fertilizer production and consumption in Turkey. *Proceedings of the 5<sup>th</sup> Technical Congress of Turkish Agricultural Engineers 2*: 881–900.
- Kerkech, M., Hafiane, A., and Canals, R. (2018). Deep learning approach with colorimetric spaces and vegetation indices for vine disease detection in UAV images. *Computers and electronics in agriculture*, 155, 237-243.
- Kerkech, M., Hafiane, A., and Canals, R. (2020). Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Computers and Electronics in Agriculture*, 174, 105446.
- Keskin, G., Tatlidil, F. F., and Dellal, I. (2010). An analysis of tomato production cost and labor force productivity in Turkey. *Bulgarian Journal of Agricultural Science*, 16(6), 692-699.

- Khakimov, A., Salakhutdinov, I., Omonlikov, A., and Utaganov, S. (2022). Traditional and current-prospective methods of agricultural plant diseases detection: A review. *IOP Conference Series: Earth and Environmental Science*, 951(1), 012002. <https://doi.org/10.1088/1755-1315/951/1/012002>.
- Khasawneh, N., Faouri, E., and Fraiwan, M. (2022). Automatic Detection of Tomato Diseases Using Deep Transfer Learning. *Applied Sciences* (Switzerland), 12(17). <https://doi.org/10.3390/app12178467>
- Kutlar, I. (2019). Analysis of the sources of agricultural information available to greenhouse tomato growers in Turkey. *Fresenius Environmental Bulletin*, 28(9), 6825-6835.
- Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J., and Kazimi, Z. A. (2022). Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. *Plants*, 11(17), 2230.
- Li, G., Ma, Z. and Wang, H. (2012). Image recognition of grape downy mildew and grape Powdery mildew based on support vector machine. In *Computer and Computing Technologies in Agriculture V: 5th IFIP TC 5/SIG 5.1 Conference, CCTA 2011, Beijing, China, October 29-31, 2011, Proceedings, Part III 5* (pp. 151-162). Springer Berlin Heidelberg.
- Li, Y., and Chao, X. (2020). ANN-based continual classification in agriculture. *Agriculture*, 10(5), 178.
- Li, Y., and Chao, X. (2021). Semi-supervised few-shot learning approach for plant disease recognition. *Plant Methods*, 17, 1-10.
- Li, Y., and Yang, J. (2020). Few-shot cotton pest recognition and terminal realization. *Computers and Electronics in Agriculture*, 169, 105240.
- Li, Y., and Yang, J. (2021). Meta-learning baselines and database for few-shot classification in agriculture. *Computers and Electronics in Agriculture*, 182, 106055.
- Li, Y., Nie, J., and Chao, X. (2020). Do we need deep CNN for plant disease identification? *Computers and Electronics in Agriculture*, 178, 105803.
- Liang, W. J., Zhang, H., Zhang, G. F., and Cao, H. X. (2019). Rice blast disease recognition using a deep convolutional neural network. *Scientific reports*, 9(1).
- Lu, Y., Yi, S., Zeng, N., Liu, Y., and Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267, 378–384. <https://doi.org/10.1016/j.neucom.2017.06.023>
- Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., ... and Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for sustainable development*, 35, 1-25.
- Maind, S. B., and Wankar, P. (2014). Research paper on basics of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1), 96-100.
- Manish, K. (2017). Estimates of genetic variability, heritability and genetic advance for yield and yield component traits in thermo tolerant tomato (*Solanum*

lycopersicum L.) genotypes. *International Journal of Agriculture Sciences*, ISSN, 0975-3710.

- Menon, V., Ashwin, V., and Deepa, R. K. (2021). Plant Disease Detection using CNN and Transfer Learning. ICCISc 2021 - 2021 *International Conference on Communication, Control and Information Sciences*, Proceedings. <https://doi.org/10.1109/ICCISc52257.2021.9484957>
- Miranda, J., Ponce, P., Molina, A., and Wright, P. (2019). Sensing, smart and sustainable technologies for Agri-Food 4.0. *Computers in Industry*, 108, 21-36.
- Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.
- Moyazzoma, R., Hossain, M. A. A., Anuz, M. H., and Sattar, A. (2021). Transfer Learning Approach for Plant Leaf Disease Detection Using CNN with Pre-Trained Feature Extraction Method Mobilnetv2. *International Conference on Robotics, Electrical and Signal Processing Techniques*, 526–529. <https://doi.org/10.1109/ICREST51555.2021.9331214>
- Mukhtar, H., Khan, M. Z., Khan, M. U. G., and Younis, H. (2021). Wheat disease recognition through one-shot learning using field images. In *2021 International Conference on Artificial Intelligence (ICAI)* (229-233). IEEE.
- Mukti I.Z. and Biswas D. (2019). Transfer Learning Based Plant Disease Detection Using ResNet50. *4th International Conference on Electrical Information and Communication Technology (EICT)*. Kuhlna Bangladesh.
- Moid, M. A., and Chaurasia, M. A. (2021). Transfer Learning-based Plant Disease Detection and Diagnosis System using Xception. Proceedings of the *5th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, I-SMAC 2021, 451–455. <https://doi.org/10.1109/I-SMAC52330.2021.9640694>
- Mokhtar, U., Ali, M. A., Hassenian, A. E., and Hefny, H. (2015). Tomato leaves disease detection approach based on support vector machines. In *2015 11th International Computer Engineering Conference (ICENCO)* (246-250). IEEE.
- Narvekar, C., and Rao, M. (2020). Flower classification using CNN and transfer learning in CNN-Agriculture Perspective. In *2020 3rd international conference on intelligent sustainable systems (ICISS)* (660-664). IEEE.
- Nashwa, A., and Abo-Elyousr, M. (2012). Evaluation of various plant extracts against the Early blight disease of tomato plants under greenhouse and field conditions. *Plant Protect. Sci.*, 48: 74-79.
- Niedbała, G., Kurasiak-Popowska, D., Stuper-Szablewska, K., and Nawracała, J. (2020). Application of artificial neural networks to analyze the concentration of ferulic acid, deoxynivalenol, and nivalenol in winter wheat grain. *Agriculture*, 10(4).
- Ozkan, B. (2000). Development constraints analysis for gender growing greenhouse vegetables in Antalya, Turkey. In *XIVth International Symposium on Horticultural Economics* 536 (411-418).

- Olaoye, F., and Potter, K. (2024). Deep Learning Algorithms and Applications. *EasyChair*. (12551).
- Ouhami, M., Es-saady, Y., Hajji, M.E., Hafiane, A., Canals, R., and Yassa, M.E. (2020). Deep Transfer Learning Models for Tomato Disease Detection. *Image and Signal Processing*, 12119, 65 - 73.
- Özaydin, G., and Direk, M. (2022, December). Tomato production potential and marketing structure in Turkey. Paper presented at the *Tomato Production Potential and Marketing Structure in Turkey Conference*, Tashkent. Mithat Direk's Lab.
- Patil, B., and Bodhe, S. K. (2011). Leaf disease severity measurement using image processing. *International Journal of Engineering and Technology*, 3(5), 297-301.
- Patil, K., Saini, S., and Kumar, K. (2012). A comprehensive review on roughness geometries and investigation techniques used in artificially roughened solar air heaters. *International Journal of Renewable Energy Research*, 2(1), 1-15.
- Paymode, A. S., and Malode, V. B. (2022). Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG. *Artificial Intelligence in Agriculture*, 6, 23–33. <https://doi.org/10.1016/j.aiia.2021.12.002>
- Pentoś, K., Pieczarka, K., and Serwata, K. (2021). The relationship between soil electrical parameters and compaction of sandy clay loam soil. *Agriculture*, 11(2), 114.
- Petrellis, N. (2019). Plant disease diagnosis with color normalization. In 2019 *8th International Conference on Modern Circuits and Systems Technologies (MOCAST)* (1-4). IEEE.
- Pérez-Alonso, J., Pérez-García, M., Pasamontes-Romera, M., and Callejón-Ferre, A. J. (2012). Performance analysis and neural modeling of a greenhouse-integrated photovoltaic system. *Renewable and Sustainable Energy Reviews*, 16(7), 4675-4685.
- Pérez-Pérez, B. D., Garcia Vazquez, J. P., and Salomón-Torres, R. (2021). Evaluation of convolutional neural networks' hyperparameters with transfer learning to determine to sort of ripe medjool dates. *Agriculture*, 11(2), 115.
- Peyal, H. I., Shahriar, S. M., Sultana, A., Jahan, I., and Mondol, M. H. (2021, July 8). Detection of tomato leaf diseases using transfer learning architectures: A comparative analysis. 2021 *International Conference on Automation, Control and Mechatronics for Industry 4.0, ACMI*. <https://doi.org/10.1109/ACMI53878.2021.9528199>
- Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A., and Echazarra, J. (2019). Crop conditional Convolutional Neural Networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. *Computers and Electronics in Agriculture*, 167, 105093.
- Ponce, P., Molina, A., Cepeda, P., Lugo, E. and MacCleery, B. (2014). Greenhouse design and control. CRC press.

- Reed., R., and Dubuf, H. (1993). A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image understanding*, 57(3), 359-372.
- Revathi, P., and Hemalatha, M. (2014). Identification of cotton diseases based on cross information gain deep forward neural network classifier with PSO feature selection. *International Journal of Engineering and Technology*, 5(6), 4637
- Ristaino, J. B., Anderson, P. K., Bebber D, D. P., Brauman E, K. A., Cunniffe, N. J., Fedoroff, N. v, Finegold, C., Garrett, K. A., Gilligan, C. A., Jones K, C. M., Martin, M. D., Macdonald, G. K., Neenan, P., Records, A., Schmale, D. G., Tateosian, L., and Wei, Q. (2021). The persistent threat of 47 emerging plant disease pandemics to global food security.  
<https://doi.org/10.1073/pnas.2022239118/-/DCSupplemental>
- Roble, C. (2019). Global Greenhouse Statistics.
- Roser, M., Ritchie, H., and Ortiz-Ospina, E. (2019). World Population Growth.  
[https://ourworldindata.org/world-population-growth?source\\_d904819ea029](https://ourworldindata.org/world-population-growth?source_d904819ea029)
- Ruiz-Garcia, L., Lunadei, L., Barreiro, P., and Robla, J. I. (2009). A review of wireless sensor technologies and applications in agriculture and food industry: state of the art and current trends. *sensors*, 9(6), 4728-4750.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Berg, A.C., 2015. Imagenet large-scale visual recognition challenge. *Int. J. Comput. Vision* 115 (3), 211–252.
- Sabrol, H., and Satish, K. (2016, April). Tomato plant disease classification in digital images using classification tree. In 2016 *international conference on communication and signal processing (ICCSP)* (1242-1246). IEEE.
- Saeed, A., Abdel-Aziz, A. A., Mossad, A., Abdelhamid, M. A., Alkhaled, A. Y., and Mayhoub, M. (2023). Smart Detection of Tomato Leaf Diseases Using Transfer Learning-Based Convolutional Neural Networks. *Agriculture (Switzerland)*, 13(1). <https://doi.org/10.3390/agriculture13010139>
- Sagar, A., and Dheeba, J. (2020). On using transfer learning for plant disease detection. *BioRxiv*, 2020-05.
- Sahu, C. K., Sethy, P. K., and Behera, S. K. (2018). Sensing technology for detecting insects in a paddy crop field using an optical sensor. In *Information and Communication Technology for Sustainable Development: Proceedings of ICT4SD 2016*, 1 (191-199). Springer Singapore.
- Sammons, P.J., Furukawa, T. and Bulgin, A. (2005). Autonomous pesticide spraying robot for use in a greenhouse. In Australian Conference on Robotics and Automation (Vol. 1, No. 9). *Canberra, Australia: Commonwealth Scientific and Industrial Research Organisation*. December.
- Sankaran, S., Mishra, A., Ehsani, R., and Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and electronics in agriculture*, 72(1), 1-13.
- Sardogan, M., Tuncer, A. and Ozen, Y. (2018). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 *3rd International*

- Conference on Computer Science and Engineering (UBMK)* (pp. 382-385). IEEE. September.
- Sarma, K. K., Das, K. K., Mishra, V., Bhuiya, S., and Kaplun, D. (2022). Learning-aided system for agriculture monitoring designed using image processing and IoT-CNN. *IEEE Access*, 10, 41525-41536.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, (85-117).
- Sehgal, A., and Mathur, S. (2019). Plant disease classification using soft computing supervised machine learning. In *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (75-80). IEEE.
- Shah, T. M., Nasika, D. P. B., and Otterpohl, R. (2021). Plant and weed identifier robot as an agroecological tool using artificial neural networks for image identification. *Agriculture*, 11(3), 222.
- Shrivastava, V. K., Pradhan, M. K., Minz, S., and Thakur, M. P. (2019). Rice plant disease classification using transfer learning of deep convolution neural network. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(3/W6), 631–635. <https://doi.org/10.5194/isprs-archives-XLII-3-W6-631-2019>
- Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. In: *Int. Conf. Learn. Represent.* 1–14
- Singh, K.K. (2018). “An Artificial Intelligence and Cloud-Based Collaborative Platform for Plant Disease Identification, Tracking and Forecasting for Farmers”, *IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, Bangalore, India, 23-24 Nov, 49-56
- Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., and Stefanovic, D. (2016). Deep neural networks-based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, 2016(1), 3289801.
- Strange, R., and Scott, P. (2005). Plant disease: A threat to global food security. *Annual Review of Phytopathology*, 43(1), 83-116. <https://doi.org/10.1146/annurev.phyto.43.113004.133839>.
- Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30, pp. 2818–2826.
- Takayama, M., and Ezura, H. (2015). How and why does tomato accumulate a large amount of GABA in the fruit? *Frontiers in plant science*, 6, 612. <https://doi.org/10.3389/fpls.2015.00612>
- Tampakaki, A. P., Skandalis, N., Gazi, A. D., Bastaki, M. N., Sarris, P. F., Charova, S. N., Kokkinidis, M., and Panopoulos, N. J. (2010). Playing the “Harp”: Evolution of our understanding of hrp/hrc genes. *Annual Review of Phytopathology*, 48, 347-370. <https://doi.org/10.1146/annurev-phyto-073009-114407>

- Tarek, S., and Nasreldin, T. I. (2023). Towards applying smart mobility solutions in Egypt: An integrative framework and a case study application. *Ain Shams Engineering Journal*, 14(7), 101987.
- Thakur, P., Chug, A., and Singh, A. P. (2021). Plant Disease detection of Bell Pepper Plant Using Transfer Learning over different Models. Proceedings of the 8th International Conference on Signal Processing and Integrated Networks, SPIN 2021, 384–389. <https://doi.org/10.1109/SPIN52536.2021.9565945>
- Thangaraj, R., Anandamurugan, S., and Kaliappan, V. K. (2021). Automated tomato leaf disease classification using transfer learning-based deep convolution neural network. *Journal of Plant Diseases and Protection*, 128(1), 73-86.
- Thangaraj, R., Anandamurugan, S., Pandiyan, P., and Kaliappan, V. K. (2022). Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion. In *Journal of Plant Diseases and Protection* (Vol. 129, Issue 3, pp. 469–488). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s41348-021-00500-8>
- Tunio, M. H., Jianping, L., Butt, M. H. F., and Memon, I. (2021). Identification and Classification of Rice Plant Disease Using Hybrid Transfer Learning. 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2021, 525–529. <https://doi.org/10.1109/ICCWAMTIP53232.2021.9674124>
- TÜİK (2023). Türkiye İstatistik Kurumu. <https://data.tuik.gov.tr/Bulten/Index?p=Bitkisel-Uretim-Istatistikleri-2021-37249>.
- United Nations. (2022). Our growing population. <https://www.un.org/en/globalissues/population#:~:text=The%20world's%20population%20is%20expected,nearly%2011%20billion%20around%202100>
- Von Zabeltitz, C. Greenhouse structures. In *Integrated Greenhouse Systems for Mild Climates*; Springer: Berlin/Heidelberg, Germany, 2011, 59–135.
- Wang, H., Li, G., Ma, Z., and Li, X. (2012). Image recognition of plant diseases based on backpropagation networks. In Proceedings of the 2012 5th International Congress on Image and Signal Processing, (894-900). Chongqing, China. <https://doi.org/10.1109/CISP.2012.6469998>
- Wang, M., Wang, J., and Chen, L. (2020). Mapping paddy rice using weakly supervised long short-term memory network with time series Sentinel optical and SAR images. *Agriculture*, 10(10), 483.
- Wang, X., and Liu, J. (2021). Tomato anomalies detection in greenhouse scenarios based on YOLO-dense. *Frontiers in Plant Science*, 12, p.533.
- Wang, G.Li , and Z.Ma, X.Li. (2012). “Application of Neural Networks to Image Recognition of Plant Diseases”, *International Conference on Systems and Informatics (ICSAI2012)*, Yantai, China, 19-20 May 2159-2164.
- Xiaoling, X., Cui, X., and Bing, N. (2017). Inception-v3 for flower classification. 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, 2017, pp. 783-787, doi: 10.1109/ICIVC.2017.7984661.

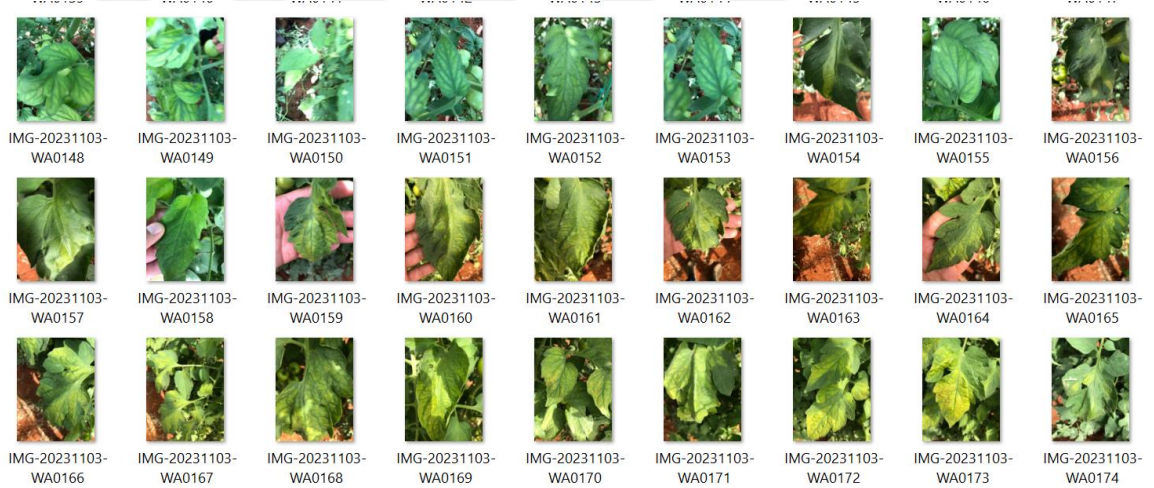
- Xie, J., Yu, J., Chen, B., Feng, Z., and Siddique, K. H. M. (2017). Facility cultivation systems: a Chinese model for the planet. *Adv. Agron.* 145, 1–42. doi: 10.1016/bs.Agron.2017.05.005
- Xu, Z., Shou, W., Huang, K., Zhou, S., Li, G., and Tang, G. (2000). The current situation and trend of tomato cultivation in China. *Acta Physiol Plant.*;22(3):379–82.
- Yang, N. W., Zang, L. S., Wang, S., Guo, J. Y., Xu, H. X., Zhang, F., and Wan, F. H. (2014). Biological pest management by predators and parasitoids in the greenhouse vegetables in China. *Biological control*, 68, 92-102.
- Yang, J., Guo, X., Li, Y., Marinello, F., Ercisli, S., and Zhang, Z. (2022). A survey of few-shot learning in smart agriculture: developments, applications, and challenges. *Plant Methods*, 18(1), 28.
- Yarak, K., Witayangkurn, A., Kritiyutanont, K., Arunplod, C., and Shibasaki, R. (2021). Oil palm tree detection and health classification on high-resolution imagery using deep learning. *Agriculture*, 11(2), 183.
- Yasmeen, U., Khan, M. A., Tariq, U., Khan, J. A., Yar, M. A. E., Hanif, C. A., ... and Nam, Y. (2021). Citrus disease recognition using a deeply improved genetic algorithm. *Comput. Mater. Contin*, 71, 3667-3684.
- Yilmaz, I., Sayin, C., and Ozkan, B. (2005). Turkish greenhouse industry: past, present, and future. *New Zealand Journal of Crop and Horticultural Science*, 33(3)
- Zeng, M., Gao, H., and Wan, L. (2021). Few-shot grape leaf disease classification based on generative adversarial network. *In Journal of Physics: Conference Series*, 1883(1)
- Zhou, F.H., and Liao, Z.Z. (2013). A particle swarm optimization algorithm. *In Applied Mechanics and Materials*, 303, 1369-1372. Trans Tech Publications Ltd.
- Zhao, Z.-Q., Zheng, P., Xu, S.-T., and Wu, X. (2018). Object detection with deep learning: A review (Version 2). arXiv. <https://arxiv.org/abs/1807.05511>
- Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., and Guo, Y. (2018). Deep learning for smart agriculture: Concepts, tools, applications, and opportunities. *International Journal of Agricultural and Biological Engineering*, 11(4), 32-44.

## APPENDIX

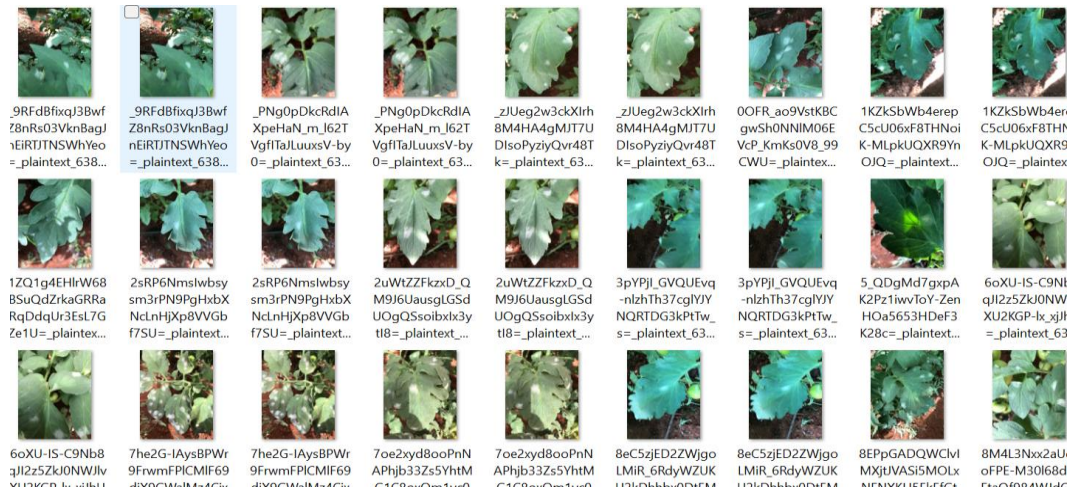
### Early blight image dataset



### Yellow Disease Image dataset



## Powdery mildew image dataset



## Healthy image dataset



## Sample of Python codes

### ✓ VGG16

```
[ ] # Instantiating vgg16
import tensorflow as tf
from keras import preprocessing
from keras.preprocessing.image import ImageDataGenerator

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D
from tensorflow.keras.layers import MaxPooling2D
from tensorflow.keras.layers import Flatten
from tensorflow.keras.layers import Dropout
from tensorflow.keras.layers import Dense
from tensorflow.keras.callbacks import EarlyStopping, ModelCheckpoint

from keras.applications import VGG16

conv_base = VGG16(weights='imagenet',
                    include_top=False,
                    input_shape=(400, 400, 3))

# Adding a densely connected classifier on top of the convolutional base

from keras import models
from keras import layers
```

Found 873 images belonging to 10 classes.  
4/4 [=====] - 20s 2s/step

|   | Accuracy | Precision | Recall   | Specificity | F1-score | AUC-ROC  | AUC-PR   | Cohen's kappa | MCC      |
|---|----------|-----------|----------|-------------|----------|----------|----------|---------------|----------|
| 0 | 0.924399 | 0.926582  | 0.924399 | 0.987805    | 0.923771 | 0.997691 | 0.981203 | 0.915888      | 0.916261 |

```
import os
print(os.getcwd())
from google.colab import files
files.download('Resnet_best_model.h5')
```

/content

### ✓ MobileNet

```
[ ] # Instantiating MoblileNet
import tensorflow as tf
from keras import preprocessing
from keras.preprocessing.image import ImageDataGenerator

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D
```

```

Picture: Kasel37/20201002_102744_output.jpg  Misprediction ==> starkspurgoldendelicious
Picture: Kasel37/20201002_102839_output.jpg  Misprediction ==> Breaburn
Picture: Kasel37/20201002_102940_output.jpg  Misprediction ==> Breaburn
Picture: Kasel37/20201002_103005_output.jpg  Misprediction ==> Breaburn
Picture: Kasel37/20201002_103017_output.jpg  Misprediction ==> Redchief
Picture: Mondialgala/20200901_150702_output.jpg  Misprediction ==> Goldenreinders
Picture: Redchief/20201005_133513_output.jpg  Misprediction ==> Scarletsbur
Picture: Redchief/20201005_134307_output.jpg  Misprediction ==> Scarletsbur
Picture: Redchief/20201005_134406_output.jpg  Misprediction ==> Kasel37
Picture: Scarletsbur/20200929_120759_output.jpg  Misprediction ==> Goldenreinders
Picture: Scarletsbur/20200929_120842_output.jpg  Misprediction ==> Redchief
Picture: Starkcrimson/20201009_090551_output.jpg  Misprediction ==> Scarletsbur
Picture: Starkcrimson/20201009_092107_output.jpg  Misprediction ==> Redchief
Picture: starkspurgoldendelicious/20201001_133233_output.jpg  Misprediction ==> Grannysmith
Picture: starkspurgoldendelicious/20201001_134956_output.jpg  Misprediction ==> Grannysmith

```

```
[ ] results_for('best_model.h5', validation_generator, "Validation")
```

### ▼ InceptionV3

```
[ ] # Instantiating Inception
import tensorflow as tf
from keras import preprocessing
```

## Location of the Greenhouses



## CURRICULUM VITAE

I completed my Bachelor of Agricultural Machinery Eng studies at the College of Agriculture at Koya University, Iraq. I received a cum laude degree from Koya University's Faculty of Agricultural Engineering. My GPA was 70.46 out of 100. As a result, I was permitted by the Ministry of Higher Education to be a tutor and a member of our teaching team since 2009 and continuing until now. Moreover, I hold a Master of Engineering Science with a Major in Agricultural Engineering, from February 27, 2012, to January 2, 2014, from Southern Queensland University, Toowoomba, Queensland, Australia. After completing my master's I worked as a lecturer at the College of Engineering manufacturing department delivering courses since April 2014 at Koya University. Besides I am a member of the Iraq Agricultural Engineers Syndicate and a member of Engineer Australia. In 2019, I was selected as a PhD student by Turkish scholarships and it was a great chance and opportunity, I started my PhD study in 2020, at Ondokuz Mayıs University, Faculty of Agriculture in the Department of Agricultural Machinery and Technologies Engineering.

### Contact information:

ORCID ID :  0000-0003-2584-6945.

### Publications:

1. Alotaibi, J. G., Omer, A. K., Salih, N. D., and Yousif, B. F., (2020). Tribological Performance of Materials Under Bio-Three Body Abrasion. *American Journal of Applied Sciences*, 17(1), 161-170.
2. Gürdil, G., Mengstu, M., and Omer, A. K., (2021). Utilization of Agricultural Wastes for Sustainable Development. *Black Sea Journal of Agriculture*, 4(4), 146-152.
3. Omer, A. K., Poh, T. K., Lama, J. A. K., Chin, C. W., and Yousif, B. F., (2017). Corrosion Simulation of High-Stress Three-Body Abrasion. of *the 21st World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2017)*.
4. Omer, A. K, Y. B. Öztekin, (2023). Deep Learning as a New Technique to Detect Plant Leaf Diseases *15th International Congress on Agricultural Mechanization and Energy in Agriculture – ANKAgEng'2023*“, October 29 – November 1, 2023, in Antalya, Türkiye.
5. Omer, A. K Shvan H Radha, (2020). The Modification of Smart Rubber Tire Structure. *ZANCO Journal of Pure and Applied Sciences*, 28(2), 78-81.

### **Won Awards, Incentives and Scholarships**

1. The Human Capacity Development Program in Higher Education (HCDP) fully funded the Master of Agricultural Engineering scholarship in USQ - Australia from 2011 to 2014.
2. Türkiye scholarship fully funded PhD 2019 to 2024.
3. The best paper award at the 21<sup>st</sup> World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2017), United States of America.

