: N° d'ordre
: N° de série

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research

UNIVERSITY OF ECHAHID HAMMA LAKHDAR - EL OUED

FACULTY OF EXACT SCIENCES
Computer Science Department

Thesis
submitted in partial fulfilment of the requirements for the degree of

ACADEMIC MASTER

Field: **Mathematics and Computer Science**
Option: **Computer Science**
Specialty: **Distributed Systems and Artificial Intelligence**

Presented by :

- Bessei Yacine
- Bessei Bachir

Title

Validation of data preprocessing to guarantee high accuracy in deep learning treatments: Agricultural domain as a case of study

Examination Commitee :

M.	Abdelkader Laouid	MAA	Supervisor
M.	Messaoud Abbas	MCB	Chairman
M.	Abdennacer Khelaifa	MAA	Examiner

Academic year: 2022-2023

شكر و تقدير

قال تعالى (**لَئِنْ شَكَرْتُمْ لَأَزِيدَنَّكُمْ**)سورة إبراهيم الآية 07

الحمد الله أولا وآخرا بأن أنار لنا درب العلم والمعرفة ووفقنا لإتمام هذا العمل المتواضع

أتقدم بجزيل الشكر وعظيم الامتنان إلى من تشرفت بإشرافه

الأستاذ الدكتور: **عبد القادر لعويد**

على جميع توجيهاته ونصائحه القيمة

كما أشكر أعضاء لجنة المناقشة على تفضلهم بقبول مناقشة هذه المذكرة

واعترافا لأهل الفضل بفضلهم وأهل الإحسان بجودهم وعملا بالأدب الإسلامي الرفيع بشكر المحسن على إحسانه

أتقدم بالشكر إلى كل **طاقم كلية العلوم الدقيقة** بجامعة الشهيد حمه لخضر الوادي

إلى كل من علمني علما عبر كامل مراحل الدراسة من الابتدائي إلى الجامعة

إلى كل من ساعدني في إنجاز هذا العمل ومد لي يد العون أخص بالذكر **صاحب مزرعة العياشي**

شكرا لكم جميعا

الإهداء

أهدي ثمرة جهدي المتواضع هذا إلى من كلله الله بالهيبة و الوقار ...
إلى من علمني العطاء بدون انتظار إلى من أحمل اسمه بكل افتخار...
إلى الذي غاب لكنه حاضر دائما في قلبي وحياتي.... إلى **أبي الغالي**
رحمه الله، ستبقى دائما حاضرا رغم فراقك.

إلى ذلك الحرف اللامتناهي من الحب والحنان، إلى ملاكي في الحياة، إلى من مارست دور الأب والأم، إلى من كان دعاؤها سر نجاحي وحنانها، **أمي الغالية** والحبيبة حفظها الله.

إلى سندنا وشيخنا وقائدنا سيدي **محمد العيد التماسيني** الذي لم يدخر جهدا في توجيهنا ونصحنا إلى طريق العلم والمعرفة رضي الله عنه وبارك في عمره واجعله ذخرا لنا.

إلى من أرى السعادة بأعينهم وأرتاح وأنا بينهم **إخوتي الأعزة.**

إلى أعز الأصدقاء

إلى كل من قدم لنا يد العون من قريب أو بعيد لإنجاز هذا العمل المتواضع.

Abstract

The agricultural sector occupies a great importance to mankind, and it is constantly striving to strengthen and develop its systems. Similar to other fields, the agricultural industry has adopted deep learning techniques to process agricultural data, with the aim of achieving high quality crop results. In different regions as Europe, North America, and East Asia, these techniques have been used to identify plant diseases, their causes, and even predict crop yields in certain seasons. However, we have noticed a lack of interest in applying these technologies in the desert environment, because it is completely different in terms of non-fertile soil quality, drought, water salinity, extreme temperatures, and so on. Given the above conditions, we propose to use previously developed models based on convolutional neural networks to solve the aforementioned problems. Our focus is to identify and classify tomato leaf diseases using a specially collected dataset from a desert environment.
Keywords : Artificial intelligence, Deep learning, Convolutional Neural Networks, Saharan Agriculture.

Résumé

Le secteur agricole occupe une grande importance pour l'humanité, et il s'efforce constamment de renforcer et de développer ses systèmes. Comme dans d'autres domaines, l'industrie agricole a adopté des techniques d'apprentissage en profondeur pour traiter les données agricoles, dans le but d'obtenir des résultats de récolte de haute qualité. Dans des régions aussi différentes que l'Europe, l'Amérique du Nord et l'Asie de l'Est, ces techniques ont été utilisées pour identifier les maladies des plantes, identifier leurs causes et même prédire les rendements des cultures à certaines saisons. Cependant, nous avons remarqué un manque d'intérêt pour l'application de ces technologies à l'environnement désertique, car l'environnement désertique est complètement différent de la qualité des sols non fertiles, de la sécheresse, de la salinité de l'eau, des températures extrêmes, etc. Compte tenu des conditions ci-dessus, nous proposons d'utiliser des modèles précédemment développés basés sur des réseaux de neurones convolutifs pour combler cette lacune. Notre objectif est d'identifier et de classer les maladies des feuilles de tomate à l'aide d'un ensemble de données spécialement collectées dans un environnement désertique.
Mots clés : Intelligence artificielle, Apprentissage approfondi , Réseaux de neurones convolutifs , Agriculture saharienne.

ملخص

يحتل القطاع الزراعي أهمية كبيرة للبشرية ، ويسعى باستمرار لتعزيز وتطوير أنظمته. على غرار المجالات الأخرى ، تبنت الصناعة الزراعية تقنيات التعلم العميق لمعالجة البيانات الزراعية ، بهدف تحقيق نتائج محاصيل عالية الجودة. في مناطق مختلفة مثل أوروبا وأمريكا الشمالية وشرق آسيا ، تم استخدام هذه التقنيات لتحديد أمراض النبات وتحديد أسبابها وحتى التنبؤ بغلات المحاصيل في مواسم معينة. ومع ذلك ، فقد لاحظنا قلة الاهتمام في تطبيق هذه التقنيات على البيئة الصحراوية ، لأن البيئة الصحراوية مختلفة تماما من نوعية التربة غير الخصبة والجفاف وملوحة المياه ودرجات الحرارة القصوى وغير ذلك. بالنظر إلى الظروف المذكورة أعلاه ، نقترح استخدام نماذج مطورة مسبقًا تعتمد على الشبكات العصبية التلافيفية لمعالجة هذه الفجوة. ينصب تركيزنا على تحديد وتصنيف أمراض أوراق الطماطم باستخدام مجموعة بيانات تم جمعها خصيصًا من بيئة صحراوية.

الكلمات المفتاحية : الذكاء الاصطناعي ، التعلم العميق ، الشبكات العصبية التلافيفية ، الزراعة الصحراوية.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1 General Introduction

This chapter serves as the foundation for our research, providing a comprehensive overview of the selected topic and outlining its significance. We begin by presenting the background and motivation, offering insights into the research area and the driving factors behind our study. Subsequently, we identify the specific problem or knowledge gap that our research aims to address. we formulate research questions that provide a clear direction. Additionally, we define the goals we strive to achieve. Finally, we establish the scope and limitations of our study, delineating boundaries and potential constraints that may impact our findings. By comprehensively addressing these elements, this chapter sets the stage for subsequent sections, ensuring a cohesive and informed exploration of our research topic.

1.1 Background and Motivation

Farming is an essential activity that has played a crucial role in the development and growth of human civilization. It is a practice that involves cultivating crops and raising livestock to provide food and raw materials for various Food Industry. Agriculture has spread throughout the world and has become increasingly popular due to its significance in providing food security for people [3]. In recent years, there has been a significant shift towards the integration of technology in agriculture, with the introduction of artificial intelligence (AI) being one of the most notable advancements. AI in agriculture involves the use of computer algorithms and machine learning techniques to analyze and interpret data collected from various sources, such as imaging devices like phones and cameras [29]. These imaging devices have become crucial tools for farmers, allowing them to collect data on crop health, soil quality, and environmental conditions in real-time. The data collected can then be analyzed by AI algorithms to provide insights and recommendations on the best course of action for improving crop yields and preventing disease outbreaks. The use of AI in agriculture has several benefits, including increased efficiency, reduced costs, and improved crop yields. AI algorithms can analyze large amounts of data in a short amount of time, enabling farmers to make informed decisions quickly. Additionally, AI can help farmers optimize resource use, such as water and fertilizer, by providing recommendations on the most efficient application methods [26].

1.2 Problem Statement

The importance of agriculture cannot be understated. It is an essential sector that provides food and other basic necessities to the world's population. However, agricultural diseases pose a significant threat to this sector. Unfortunately, some farmers remain ignorant of the existence of these diseases, leading to the destruction of crops. Moreover, agricultural diseases can spread rapidly, making it crucial for farmers to identify and manage them in their early stages. The lack of specialization in agriculture by some farmers can further exacerbate this issue, as they may not have the necessary knowledge and skills to detect and manage agricultural diseases. As a result, these farmers may only notice the existence of the disease after it has spread significantly, leading to a significant loss of crops and revenue. Therefore, it is imperative for farmers to be aware of the common agricultural diseases and their symptoms and take proactive measures to manage them before they can cause significant damage.

1.3 Research Questions

The problems and difficulties that we mentioned earlier dictated to us the following questions:

- Is it possible to provide a program that helps the farmer to detect diseases early?
- Can we access a program that gives high quality results?
- Is it possible to provide a program that accepts all kinds of images in all their forms?
- Can we collect data that will help us in the study?

1.4 Objectives

The aim of this project is to develop a system that helps farmers detect agricultural diseases and give them accurate treatment procedures. The following features must be included in such a system:

- The system must recognize all plant varieties.
- The system accepts all types of images in any format.
- In the event of a disease of a particular plant, the system gives a set of recommendations for its treatment.

1.5 Scope and Limitations

The scope of the study is to develop an AI system to detect plant diseases. However, there are several limitations that may affect the results of the study. Firstly, there is a lack of data available for the study, which can make it difficult to train the AI system accurately. Additionally, collecting data from remote areas can be challenging, which further limits the amount of data that can be collected. The imaging device used for data collection may not be the most powerful, which could result in lower quality images that are more difficult to analyze.Furthermore, there are plants that have not yet been cultivated because they are not in their growing season, and this reason limits the amount of data that can be collected on certain species only. Finally, the computer used to process the photos may not be powerful enough to handle the amount of data required for accurate analysis. These limitations should be taken into consideration when interpreting the results of the study.

1.6 Research structure

The structure of this research is as follows:

- **Chapter 2: Literature Review** In this chapter, we will discuss the definition of artificial intelligence and its applications, focusing on its relevance to agriculture. We will also explore the challenges and opportunities in desert agriculture and provide examples of research conducted on artificial intelligence in this field.
- **Chapter 3: Research Methodology** Chapter 3 provides a comprehensive overview of the methodology employed in our research. We will discuss the research design, the AI technologies and tools utilized, and the framework for implementing AI in agriculture.

- **Chapter 4: Application of AI in Sahara Agriculture** This chapter presents an extensive study on the application of artificial intelligence (AI) in the agricultural sector of the Sahara region. It specifically identifies the AI techniques employed in the case study. Furthermore, the chapter presents the results obtained from the implementation of AI and conducts a comprehensive analysis of these results.

- **Chapter 5: Conclusion and Future Vision** In the final chapter, we conclude the thesis by summarizing the key findings and implications. We also provide a clear delineation of the research boundaries and offer a vision for future advancements in this area.

Chapter 2

Literature Review

2 Literature Review

In this chapter, we delve into the fascinating realm of artificial intelligence (AI) and its applications, specifically focusing on its utilization in Saharan agriculture. then we embark on a comprehensive exploration of the role of AI in Saharan agriculture. By understanding the applications, challenges, and opportunities, as well as drawing upon previous research, we aim to pave the way for novel approaches and strategies that harness the power of AI to overcome obstacles of agriculture in the Sahara desert.

2.1 Overview of Artificial Intelligence

Artificial intelligence (AI) is a transformative technology that has revolutionized various industries. It encompasses the development of intelligent systems that can perform tasks that typically require human intelligence. From machine learning algorithms to computer vision and natural language processing, AI has found applications in diverse fields such as healthcare, finance, transportation, and more. This chapter provides an overview of AI, exploring its fundamental concepts and highlighting its wide-ranging applications across different domains.

2.1.1 The concept of artificial intelligence

Artificial intelligence (AI) is a rapidly evolving field of computer science that involves creating machines and computer systems that can perform tasks that typically require human intelligence, such as understanding natural language, recognizing images and speech, making decisions, and learning from experience. AI has the potential to revolutionize many aspects of human life, from healthcare to transportation to entertainment. The concept of artificial intelligence dates back to the mid-20th century, when computer scientists began to develop algorithms and techniques for simulating human intelligence. Early AI systems were designed to perform simple tasks, such as playing chess or solving mathematical problems. However, in recent years, the development of machine learning algorithms and the explosion of data available to train them has enabled AI to perform more complex tasks, such as natural language processing and computer vision [20].

2.1.2 Artificial intelligence applications

Artificial Intelligence, in the form of neural networks and expert systems, has wide-ranging applications in various human activities. Its high precision and low computation time make it a cutting-edge technology. Robot expert systems are already taking over workshop-level jobs in large industries, which means humans are being sidelined into more supervisory roles. Additionally, stock brokerage firms are utilizing Artificial Intelligence to analyze data, make investment analyses, and buy or sell stocks without any human interference. Here are some examples of the applications of Artificial Intelligence [16].

- **Computer Games:** Game AI, also known as Artificial Intelligence in Computer Games, uses algorithmic techniques to enhance the player's experience in computer and video games. The primary objective of Game AI is to create the impression of intelligence in the behavior of Non-Player Characters (NPCs), including opponents, companions, and other entities in the game world. Nowadays, almost all modern computer games utilize some form of artificial intelligence, making it the largest class of commercial products where the public regularly interacts with this technology [11].

- **Heavy industries:** Artificial intelligence (AI) robots are machines that are programmed to perform a wide range of tasks that typically require human-like intelligence, such as decision-making, problem-solving, and pattern recognition. These robots have become increasingly popular in heavy industries, such as manufacturing, construction, and mining, because they are able to perform tasks that are too dangerous or difficult for humans to do. One of the major advantages of AI robots is that they do not need any breaks while working, unlike humans who require rest and breaks to avoid fatigue and exhaustion. This means that robots can work continuously for extended periods of time, thus increasing the overall efficiency and productivity of the industries they are employed in [16].

- **Weather Forecasting:** One application of Artificial intelligence is in predicting weather conditions. Weather prediction is a complex task that requires analyzing vast amounts of data, including temperature, humidity, air pressure, wind speed, and more. Traditionally, weather forecasting has relied on physical models that use mathematical equations to simulate the behavior of the atmosphere. However, these models are limited in their accuracy and are often unable to predict weather events with high certainty [16].

- **Healthcare** AI has the potential to revolutionize healthcare by providing faster and more accurate diagnoses, identifying patterns in medical data that can lead to new treatments, and providing personalized care to patients. For example, AI algorithms can analyze medical images to detect cancerous tumors or identify abnormal patterns in brain scans that may indicate a neurological disorder [25].

2.2 Applications of AI in Agriculture

According to agricultural organization, the world population is increasing at a very fast rate. With increasing population, the need for food also increases briskly. Due to this context, we need to use various technological solutions to make farming more convenient and efficient. These latest technologies comprise of Artificial Intelligence, Machine Learning and Deep Learning. Artificial Intelligence technology is supporting numerous sectors of agriculture to boost the productivity and efficiency.In this title, we will explore some of the applications of artificial intelligence in agriculture, including how it is currently used and its potential in the future[36].

- **Agricultural Robots:** As the world population continues to grow, the agricultural industry will face challenges in meeting the increasing demand for food. One of the biggest challenges will be the shortage of labor in many countries. To address this issue, agricultural machinery needs to be more efficient, cost-effective, and advantageous. Robotics has been integrated into various aspects of agriculture, including aerial imaging, autonomous navigation, indoor harvesting, fruit harvesting, spraying and weeding, and plant diagnosis. For example, drones provide real-time high-quality imaging to help identify and monitor crops, assess their progress, and determine when they are ready for harvesting. Driver-less tractors offer precision and safety, while harvesting robots use machine vision to identify ripe fruits and harvest them efficiently. Blue River Technology helps to reduce the use of herbicides and chemicals by precisely identifying and spraying only the regions where weeds are present [40]. Additionally, the Plantix app uses AI to provide farmers and gardeners with information about crop diseases, pest infestations, and nutrient deficiencies by uploading a picture of the crop [42]. These advancements in agricultural robotics not only address the labor shortage but also improve crop yields and reduce costs [36].

- **Monitoring crop and soil:** Monitoring crop and soil quality is essential for ensuring food security and maintaining a healthy economy. Reliable and timely information on the status of crops and soil is crucial for decision-making in improving crop production. Efficient monitoring can identify potential defects and nutrient deficiencies in the soil and crop, and thus, play a crucial role in crop production. Recent research shows that companies are adopting computer vision and deep learning algorithms to process data collected by drones for crop and soil monitoring. Agricultural apps such as Farm at Hand and Xarvio utilize computer vision to analyze crop health. Farm at Hand focuses on crop yields and enables farmers to track the progress of tasks like planting, spraying, and harvesting. The app utilizes computer vision algorithms to highlight individual fields, and the user can set seasonal plans for the different fields or view reports for a specific field. Xarvio is a scouting app that allows users to take images of crops and identify potential threats to crop health, such as diseases, pests, and weeds. This app can be particularly useful in identifying diseases that are region-specific and not well-documented [36].

- **Pest Control:** Artificial intelligence has also been used to address the issue of the spread of plant pests, which can cause significant damage to crops. Technicians have developed semi-automated methodologies for monitoring and identifying pests using artificial neural networks, with the multilayer perceptron method proving to be effective. In addition, mobile applications that use machine learning and image classification software have been developed to assist in pest identification. For example, Wadhawani AI has developed algorithms that can detect pests that attack cotton crops. These innovations offer hope for better control and management of plant pests, helping to protect crops and secure the global food supply [36].

- **Yield Prediction:** The use of artificial intelligence and internet of things technologies in agriculture has enabled farmers to increase crop productivity and make informed decisions about crop cost estimation and marketing strategies. Crop yield prediction, based on empirical statistical models and data analytics techniques, helps to forecast crop yields on an operational basis and predict unknown patterns or insights. Two models used for crop yield estimation are Support Vector Machines (SVM) and Linear square support vector machine (LS-SVM). The implementation of descriptive analytics in agriculture has led to profitable farming and enabled farmers to equip healthy and well-nutritional foods without any loss of grains [36].

2.3 Challenges and Opportunities in Saharan Agriculture

Agriculture in the Sahara desert presents several challenges and opportunities. The region is characterized by harsh climatic conditions, including high temperatures, low rainfall, and sandy soils, which make agriculture difficult. However, with innovative approaches and technologies, it is possible to overcome these challenges and create opportunities for sustainable agricultural production in the region.

2.3.1 Challenges

There are various challenges associated with agriculture in the Sahara desert, We mention some of them:

- **Water Scarcity:** Agriculture in the Sahara desert faces the challenge of water scarcity, which is the most important factor limiting crop growth. Water resources in the region are limited and unevenly distributed, making it difficult for farmers to irrigate their crops adequately [12].
- **Extreme temperatures:** The Sahara desert is known for its extreme temperatures, with daytime temperatures reaching up to 50°C, and nighttime temperatures dropping to as low as 5°C. These temperature extremes can stress crops and lead to reduced yields [10].
- **Soil desiccation:** The soil in the Sahara desert is generally low in nutrients and organic matter, which can limit crop growth and productivity. Moreover, wind and water erosion can reduce soil quality and make it more difficult to maintain fertile soil [1].
- **Lack of infrastructure:** The Sahara desert is a remote region with limited infrastructure, including roads, electricity, and communication networks. This lack of infrastructure can make it difficult to transport goods to and from the region and limit access to markets.

2.3.2 Opportunities:

Despite these challenges, there are opportunities to develop sustainable agriculture in the Sahara desert,represented in:

- **Renewable energy:** The Sahara desert has abundant solar and wind resources, which can be used to power irrigation systems and other agricultural infrastructure. The use of renewable energy can help reduce the reliance on fossil fuels and promote sustainable agriculture [9].
- **High-value crops:** The arid conditions of the Sahara desert can be favorable for growing certain high-value crops such as date palms, olives, and almonds [33]. These crops can provide a good source of income for farmers in the region.
- **Diversification:** Agriculture in the Sahara desert can benefit from diversification of crops and livestock. Farmers can explore growing different types of crops, such as drought-resistant varieties, to reduce their dependence on a single crop and increase resilience.
- **Technological advancements:** With advancements in technology, farmers in the Sahara desert can benefit from improved irrigation systems, efficient water management techniques, and precision farming practices. These technologies can help maximize crop yields and minimize water use [35].

2.4 Previous Studies on AI in Saharan Agriculture

Recent studies have demonstrated the potential of artificial intelligence (AI) to enhance agricultural practices, including in desert regions. Some examples of research on AI in desert agriculture include:

- **"Machine Learning for Smart Farming: A Focus on Desert Agriculture":** This paper by Anandhavalli et al, in College o f Computer Science King Khalid University Abha, Kingdom o f Saudi Arabia discusses the use of smart farming, an integrated system of sensors, communication networks, artificial intelligence, and other advanced machinery connected via the internet of things (IoT), to improve food production. The paper highlights the increasing demand for food production due to population growth, scarceness of natural resources, and climate change, and the need for timely and informative farming data for solid strategies and operative decisions. The paper also focuses on the challenges faced by desert agriculture in the Kingdom of Saudi Arabia (KSA), including adverse weather conditions, scarcity of water, and short rainfall, and the need for modern resource-efficient agricultural technologies. The paper reviews recent trends in smart farming, including IoT and machine learning applications in desert agriculture initiatives, and modern resource-efficient agricultural technologies in KSA. The paper concludes by discussing the challenges and recommendations for the applicability of smart farming in KSA desert agriculture [31].

- **"Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries":** This paper by Rubby Aworka et al discusses the challenges faced by African agriculture, including climate variability, population growth, and competition for natural resources. The Food and Agriculture Organization has proposed climate-smart agriculture as a solution, which incorporates digital technologies such as machine learning to optimize resource use and increase crop yields. The authors of this paper propose a decision system that uses machine learning models to predict crop yield in East African countries. Three advanced predicting algorithms are proposed based on machine learning: Crop Random Forest, Crop Gradient Boosting Machine, and Crop Support Vector Machine. The paper presents data analysis results and experimental results, and concludes with future directions for this work [39].

- **"Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa":** The article Mupangwa et al discusses the use of crop simulation models and machine learning (ML) approaches in agricultural research, with a focus on predicting maize yields in different conservation agriculture (CA) cropping systems under highland and lowland conditions of Eastern and Southern Africa (ESA). The study aims to explore the use of ML techniques in predicting maize yields, evaluate the accuracy and precision of ML algorithms in predicting yields, and determine which algorithm(s) predict(s) maize yields better than others. The article highlights the advantages of ML approaches in agriculture, including the ability to learn from big data sets and fill the gap when required data sets for biophysical modelling are not available. Various ML techniques available for use in agriculture, including regression, fuzzy cognitive map learning, artificial neural networks, CART, KNN, random forest, and SVM, are discussed, and their performance in predicting maize yields is evaluated [32].

- **"Deep Learning for Monitoring Agricultural Drought in South Asia Using Remote Sensing Data":** The article by Foyez Ahmed and Jiahua Zhang et al is about monitoring drought using a deep learning method, specifically a deep forwarded neural network, to accurately and near-real-time identify the onset and severity of droughts in South Asia. The article describes different types of droughts and the challenges associated with monitoring droughts using traditional meteorological and remote sensing monitoring approaches. The article explains that deep learning is more flexible and robust in drought characterization and forecasting than traditional models and provides a detailed account of how the study used the soil moisture deficit index (SMDI) to monitor agricultural drought over South Asia using a deep neural network. The article also provides a literature review of previous studies that have used different machine learning models to build drought monitoring and forecasting models [37].

Chapter 3

Methodology

3 Methodology

Through this chapter, we provide a comprehensive overview of the methodology employed in our research. By discussing the research design, data collection and analysis procedures, AI techniques and tools, and the implementation framework, we ensure transparency and rigor in our study. The methodology is the cornerstone of our research, laying the groundwork for the subsequent chapters.

3.1 Research Design

In this section, we will delve into a comprehensive elucidation of our research design phase, illustrated in Figure 1. Our primary objective is to create a system that utilizes machine learning techniques for the purpose of classifying a collection of plant leaf images. We will now outline the process in detail.

- We begin by collecting the data for the project, which consists of a set of pictures of plant leaves.
- This data then goes through a pre-processing stage, which involves a series of steps taken to prepare the data for analysis or machine learning. This includes cleaning, transforming, and organizing the data to make it more suitable for use in a particular analysis or model. Pre-processing is a critical step because it greatly affects the accuracy and effectiveness of the analysis or model.
- After pre-processing, we move on to the processing stage, where we use Convolutional Neural Networks (CNN) to analyze the data and generate desirable outputs. To achieve this, we utilize the open-source TensorFlow machine learning framework developed by Google [15] to generate the CNN model in Python. This framework provides an efficient and scalable way to build, train, and evaluate machine learning models.
- Finally, we evaluate the model to determine how well it works and whether it is suitable for the intended purpose. The evaluation stage is crucial as it helps us identify any potential issues with the model and allows us to fine-tune the parameters to achieve better results. By following this research design, we hope to create a robust and accurate model that can analyze plant leaves effectively.

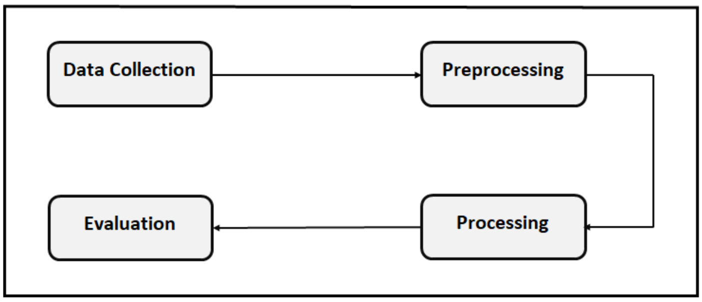

Figure 1: Stages of developing a system based on artificial intelligence.

3.2 Data Collection and Analysis

Collecting data is a crucial step when training deep learning models. Deep learning algorithms have a strong appetite for data and rely on large quantities of labeled data to effectively learn and provide accurate predictions. The quality and diversity of the collected data greatly impact the performance of deep learning models. Data collection and analysis are vital processes in various fields like agriculture, healthcare, and research [44].
In this context, we will explore how and where data were collected to develop a deep learning model aimed at detecting physiological leaf-coil disease in tomato plants. The collection process involved capturing images of both healthy and diseased tomato plant leaves over a three-day period using a range of different devices. The data acquisition took place on a farm called Mazraat Ayachi, situated in the municipality of Magrane in the state of El Oued [58]. The dataset obtained comprises images of tomato plant leaves exhibiting normal, healthy conditions, as well as images of leaves afflicted by tomato leaf curvature disease, which are visually depicted in Figure 2.

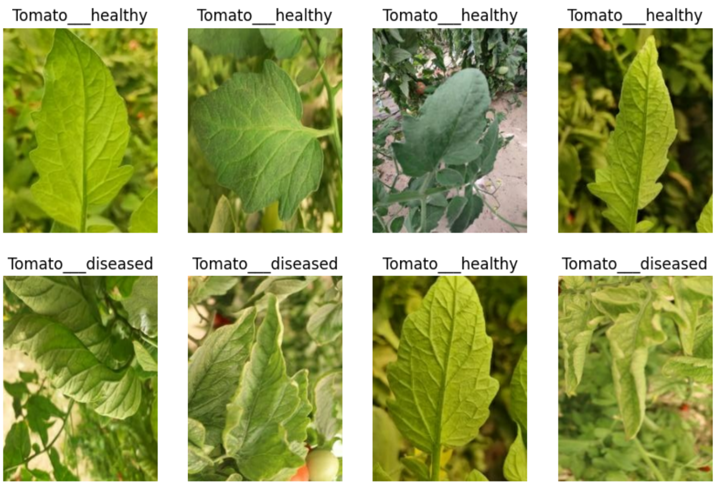

Figure 2: Sample Set from Training Data.

3.2.1 Research Data Exploration

We will provide an overview of the data collected:

- **Tomato plant:** Tomatoes, a popular vegetable crop, are rich in vitamins A and C, potassium, and lycopene, a powerful antioxidant with anti-cancer properties. Native to South America, tomato plants require good soil, full sun exposure, and constant watering to thrive. Pest and disease control is also necessary for a good crop. With their versatility and nutritional value, tomatoes are among the most consumed fruits in the world and are used in a wide variety of culinary dishes[7], Its shape is shown in Figure 3.

Figure 3: The fruits of the tomato plant.

- **Healthy Tomato Plant Leaves:** Before analyzing the images of diseased tomato plant leaves, it is essential to have a clear understanding of what healthy tomato plant leaves look like.Healthy tomato plant leaves are the ones that are green in color, not yellow or brown, with a smooth texture and an oval shape. They are typically 10-25 cm long and 6-12 cm wide. They are firmly attached to the stem, not droopy or wilted. The edges of the leaves are smooth and not jagged. The veins on the leaves are well-defined and spaced evenly apart. as shown in Figure4. Healthy tomato plant leaves should not have any spots or discoloration, or any signs of disease or insect damage [59].

Figure 4: Healthy Tomato Plant Leaves.

- **Physiological Leaf Roll Disease:** Physiological leaf roll is a common condition in tomato plants, it is characterized by the upward and inward rolling of leaf edges, as shown in Figure 5. usually in response to environmental stress such as heat or drought. It is a natural response of the plant to conserve water by reducing the amount of leaf surface exposed to the sun and air. This condition does not typically harm the plant and can often resolve on its own once the environmental stressor is removed. However, if the leaf roll is severe or accompanied by other symptoms such as wilting or yellowing of the leaves, it may indicate a more serious problem such as disease or nutrient deficiency. In such cases, it is important to identify and address the underlying issue to prevent further damage to the plant [6].

Figure 5: Infection of tomato leaves with physiological leaf roll disease.

3.3 AI Techniques and Tools

Artificial intelligence (AI) includes a set of technologies and tools that enable machines to perform tasks that would normally require human intelligence. These technologies include machine learning and deep learning. It also provides tools such as the Python libraries, OpenCV, TensorFlow, and Keras to give developers the tools they need to build and deploy AI applications.

3.3.1 Machine learning

Machine learning is a subfield of artificial intelligence that involves the development of algorithms and statistical models that enable computer systems to learn and improve from experience without being explicitly programmed. It is based on the idea that computers can be trained to recognize patterns in data, and then use those patterns to make decisions or predictions about new data. The goal of machine learning is to develop systems that can automatically learn from data and improve their performance over time, without the need for human intervention. It has applications in a wide range of fields, including image recognition, natural language processing, and predictive analytics [8].

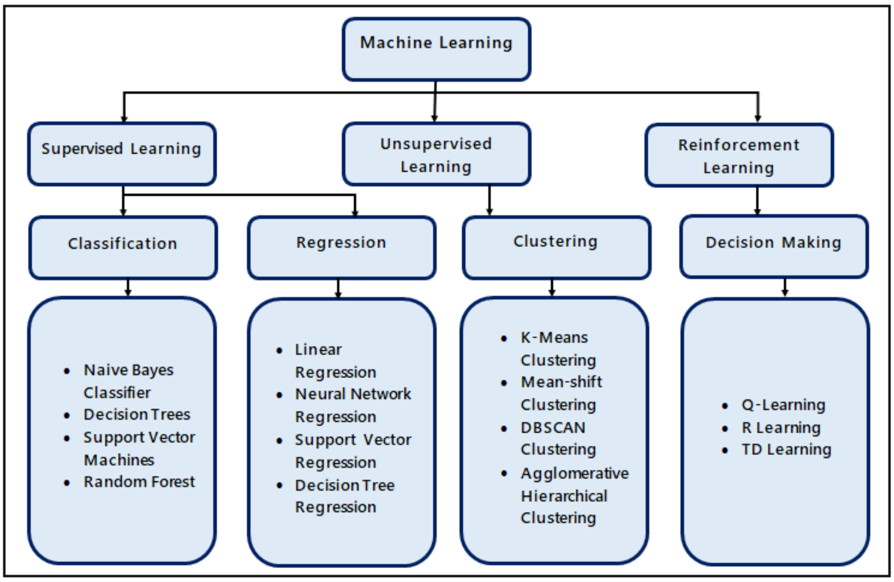

Figure 6: Types of machine learning and its algorithms.

1. **Types of machine learning**
   There are three main types of machine learning as depicted in Figure 6: supervised learning, unsupervised learning, and reinforcement learning [51].

   1.1 **Supervised Learning:** Supervised learning is a type of machine learning where the algorithm learns from labeled data. This means that the desired output for each input is already known, and the algorithm learns to map inputs to outputs. The algorithm uses this labeled data to find a pattern that it can use to predict outputs for new inputs. There are two main types of supervised learning: regression and classification. In regression, the algorithm predicts a continuous value, such as the price of a house or the temperature in a city. In classification, the algorithm predicts a discrete value, such as whether an email is spam or not [19].

   1.2 **Unsupervised Learning:** Unsupervised learning involves the machine receiving inputs x1, x2,... without any supervised target outputs or rewards from its environment. It may seem perplexing to imagine how the machine could learn without any feedback, but a formal framework for unsupervised learning can be developed. The goal of the machine is to build representations of the input that can be used for decision making, predicting future inputs, or efficiently communicating with another machine. Essentially, unsupervised learning finds patterns in the data beyond what would be considered pure unstructured noise. Clustering and dimensionality reduction are two classic examples of unsupervised learning [5].

   1.3 **Reinforcement Learning:** Reinforcement learning is a type of machine learning where the algorithm learns from trial and error. The algorithm interacts with an environment and receives feedback in the form of rewards or punishments. The goal is to learn a policy that maximizes the expected reward over time. Reinforcement learning is often used for tasks such as game playing, robotics, and autonomous driving. The algorithm learns to make decisions based on the feedback it receives from the environment [19].

2. **Machine learning algorithms**

There are many types of machine learning algorithms, each with its strengths and weaknesses. Some popular algorithms include decision trees, support vector machines, neural networks, and random forests.

2.1 **Decision tree:** A decision tree is a supervised learning algorithm used for both classification and regression tasks. It is a type of model that uses a tree-like structure to make decisions about the class or value of a target variable based on input features. The tree structure consists of nodes, branches, and leaves. Each node represents a feature or attribute, and each branch represents a possible value or decision based on that feature. The leaves of the tree represent the predicted class or value of the target variable as depicted in Figure 7 [52]. The decision tree algorithm works by recursively partitioning the data based on the features that best separate the classes or values of the target variable. At each split, the algorithm chooses the feature that results in the highest information gain [19]

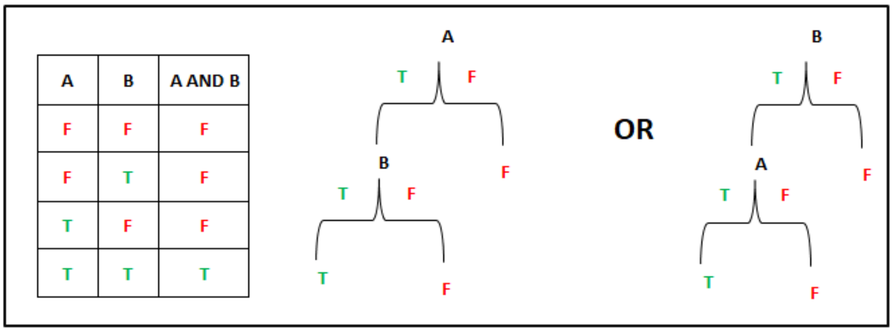

Figure 7: An example of a decision tree for an AND operation.

2.2 **Support Vector Machine (SVM):** Support Vector Machine (SVM) is a type of supervised machine learning algorithm used for classification and regression analysis. SVMs are based on the idea of finding a hyperplane that best separates data points of different classes. In a binary classification problem, the SVM algorithm tries to find the hyperplane that maximizes the margin between the two classes. The margin is the distance between the hyperplane and the closest data points from each class, and the SVM tries to find the hyperplane that maximizes this distance as depicted in Figure 8 [56]. In cases where the data is not linearly separable, the SVM algorithm can use a kernel function to map the data to a higher-dimensional space where it can be separated by a hyperplane [21].

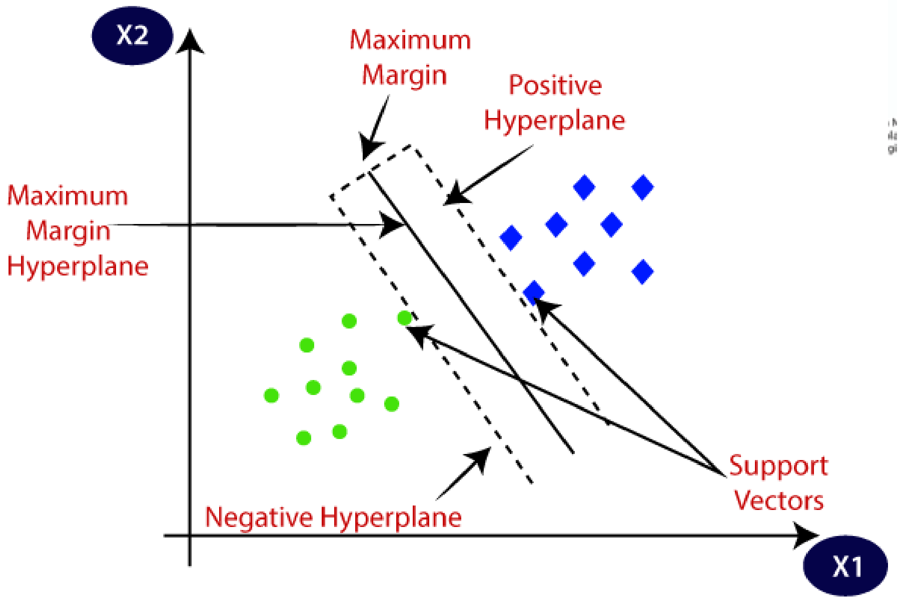

Figure 8: Support Vector Machine Algorithm.

2.3 **Neural Networks:** A neural network is a type of machine learning algorithm modeled after the structure and function of the human brain. It consists of layers of interconnected nodes, called neurons, that process information and make decisions based on that information. Each neuron in a neural network receives input from other neurons, and applies a mathematical function to that input in order to produce an output. The output from one neuron serves as the input to the next neuron in the network, and so on, until the final output of the network is produced, as depicted in Figure 9 [53] . Neural networks can be trained to recognize patterns and make predictions based on input data. During training, the weights and biases of the neurons in the network are adjusted in order to minimize the difference between the network's predicted output and the desired output. This process is repeated many times, using a large amount of training data, until the network's predictions become sufficiently accurate [30].

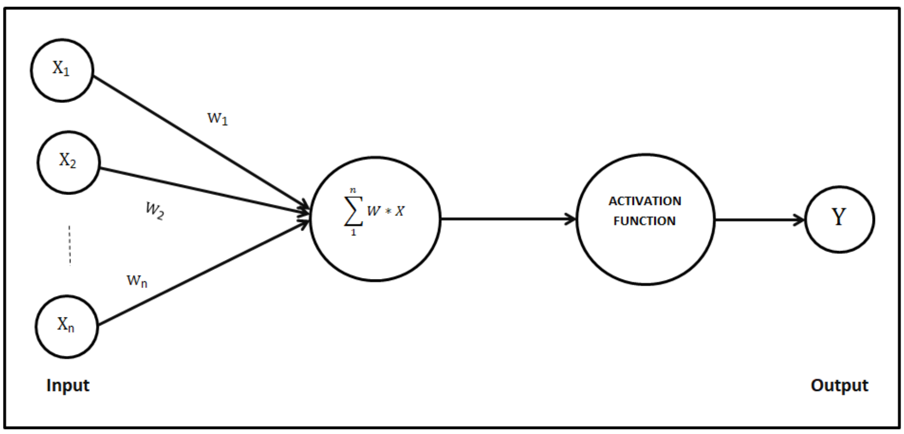

Figure 9: Simple neural network architecture.

3. **Applications of machine learning**

Machine learning has a wide range of applications in fields such as natural language processing, computer vision, speech recognition, recommendation systems, and many others. Some examples of machine learning applications include:

- Crop yield prediction.
- Recognizing faces in images.
- Generating captions for images.
- Translating text from one language to another.
- Recommending products to customers based on their purchase history.

Machine learning is a powerful tool that can be used to solve complex problems and make predictions in a wide range of industries.

3.3.2 Deep learning

Deep learning is a subfield of machine learning that uses neural networks with multiple layers to learn representations of data. It is called "deep" learning because the neural network has many layers, often ranging from tens to hundreds or even thousands of layers. Deep learning has revolutionized the field of artificial intelligence by enabling machines to perform complex tasks that were previously impossible or required a great deal of human effort. Deep learning models have achieved state-of-the-art performance in many areas such as image recognition, speech recognition, natural language processing, and game playing. Deep learning algorithms work by iteratively adjusting the weights and biases of the network based on the input data and the desired output. This process, allows the network to learn from the data and improve its performance over time. One of the advantages of deep learning is that it can automatically learn features from raw data, without the need for human engineering. This is in contrast to traditional machine learning approaches, which require manual feature engineering .As illustrated in Figure(10) [50].Deep learning has also benefited from advances in computing power and the availability of large amounts of data, which have made it possible to train very large and complex models [24].

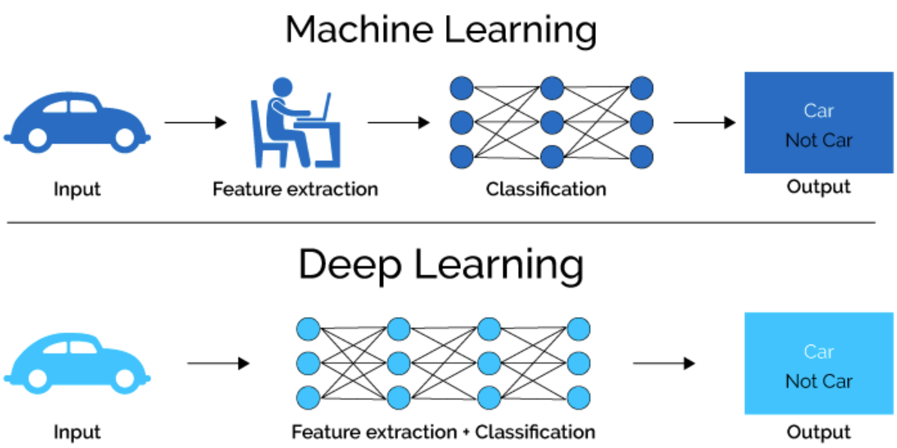

Figure 10: Comparison between machine learning and deep learning classification .

It's important to acknowledge that deep learning comprises a vast array of algorithms, each serving a specific purpose in the learning process, with their individual strengths and weaknesses. we have chosen to focus our explanation on Convolutional Neural Networks (CNNs). This selection does not diminish the value of other algorithms but rather serves to simplify our discussion and align with the primary focus of our research.

1. **Convolutional Neural Network (CNN):**

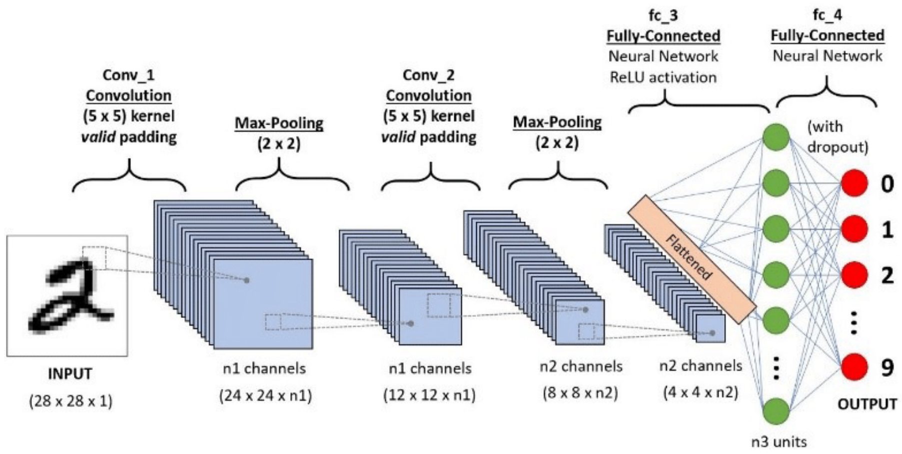

Figure 11: CNN Architecture.

A CNN is a type of artificial neural network commonly used in image recognition and computer vision tasks. It is inspired by the way the human visual system processes information, with layers of neurons that gradually learn to recognize more complex patterns in the data. At a high level, a CNN works by taking an input image and passing it through a series of convolutional layers, pooling layers, and fully connected layers. In the convolutional layers, the network applies a set of filters to the input image, looking for patterns or features that are relevant to the task at hand (e.g. edges, corners, shapes). The output of the convolutional layers is a set of "feature maps," which are fed into the pooling layers to reduce the spatial resolution of the feature maps and make the network more robust to variations in the input image. Finally, the fully connected layers take the output of the pooling layers and use it to make a prediction about the image, As illustrated in Figure(11) (e.g. "this image contains a dog") [49]. One of the key benefits of CNNs is that they are able to learn features automatically from the data, rather than relying on hand-crafted feature extractors. This makes them highly effective for tasks such as object recognition, where the relevant features may be highly complex and difficult to define manually. Overall.CNNs have achieved state-of-the-art performance in a wide range of computer vision tasks, including image classification, object detection, and image segmentation. They have also been used in other fields, such as natural language processing and speech recognition.and they continue to be an active area of research in the machine learning community [22]. CNN has a distinct operation for each layer, which is briefly summarized as follows:

1.1 **Convolutional Layers:** A convolutional layer is a key building block in a convolutional neural network (CNN). The purpose of the convolutional layer is to extract meaningful features from the input data (usually an image). This is done by convolving the input data with a set of filters, also known as kernels. Each filter scans the input data and detects a specific pattern or feature, such as an edge or a corner. By applying multiple filters, the convolutional layer is able to extract a variety of features from the input data[17], As illustrated in Figure(12) [45].

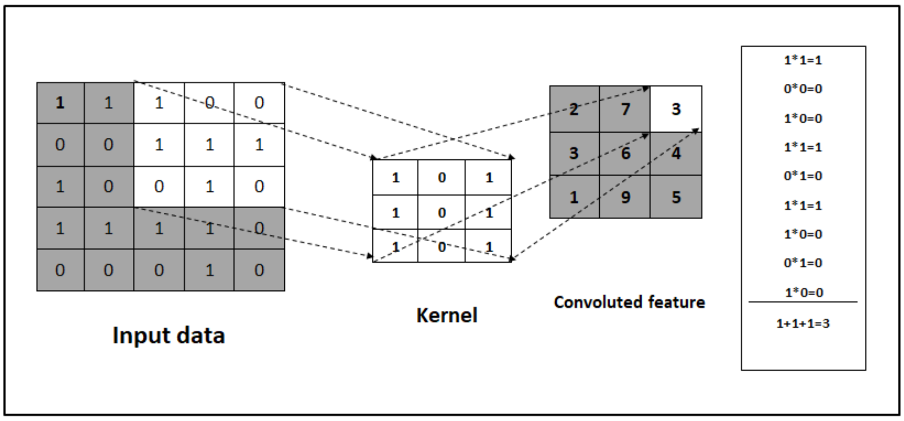

Figure 12: Feature Map computation by Convolution Layer.

1.2 **Pooling Layers:** Pooling layers are typically used in conjunction with convolutional layers in a CNN. The purpose of the pooling layer is to downsample the feature maps produced by the convolutional layer, in order to reduce the dimensionality of the data and make the network more computationally efficient. Pooling is typically done by taking the maximum or average value over a local region of the feature map, As illustrated in Figure(12) [48]. This has the effect of retaining the most important information from the feature map while discarding some of the spatial information [17].

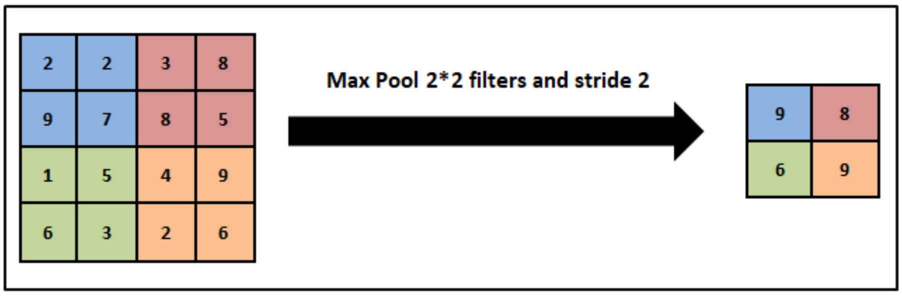

Figure 13: Max Pooling operation with 2x2 Filter and Stride value 2.

1.3 **Fully Connected Layers:** A fully connected layer is a type of layer in a neural network where every neuron in the layer is connected to every neuron in the previous layer. In a CNN, the fully connected layers are typically used at the end of the network to make a prediction about the input data. The output of the last pooling layer is flattened into a one-dimensional vector and fed into one or more fully connected layers, As illustrated in Figure(12) [55]. These layers typically use a non-linear activation function, such as ReLU or sigmoid, to introduce non-linearity into the network. Finally, the output of the last fully connected layer is passed through a softmax function to produce a probability distribution over the different classes in the classification task [17].

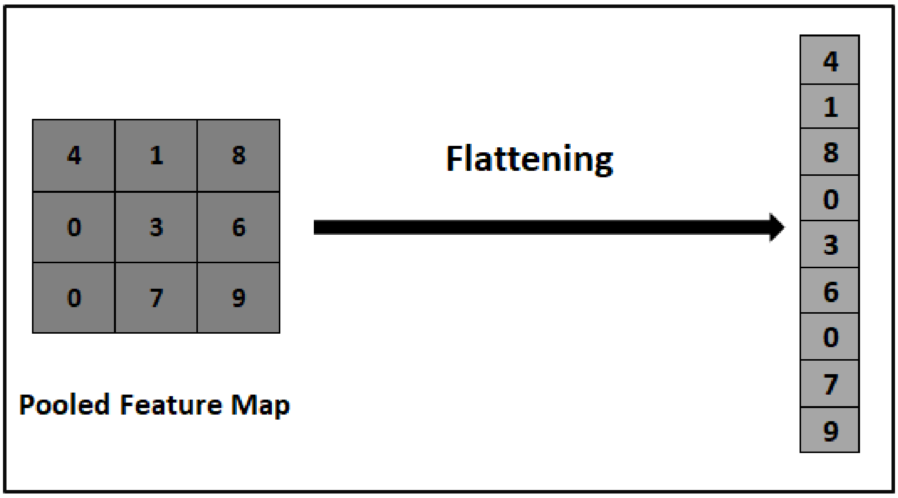

Figure 14: Flatten operation converting feature matrix into 1-D vector input.

3.3.3 Python

Python is a widely-used high-level programming language with a simple but powerful syntax and efficient data structures. Created by Guido van Rossum in the late 1980s [2], Python is popular for its readability, ease of use, and interpretive nature, allowing for faster development cycles. Its standard library includes modules for a variety of tasks, and its community support has led to the development of numerous libraries and frameworks for extending Python's functionality. Python is particularly well-suited for artificial intelligence and machine learning applications, with popular libraries such as TensorFlow, PyTorch, and Keras. Its flexibility and ease of use also make it popular in other areas of AI, such as natural language processing, robotics, and computer vision [28].

3.3.4 TensorFlow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a popular machine learning framework developed by the Google Brain team and is widely used for building deep learning models [34]. TensorFlow provides an extensive set of tools and libraries for building and training machine learning models. These tools include a wide range of pre-built models and algorithms, as well as tools for data preprocessing, model evaluation, and deployment. TensorFlow also provides support for distributed computing, allowing users to scale up their computations across multiple machines to speed up the training process.it is highly scalable, which makes it suitable for both small and large-scale machine learning projects.[14] The core of TensorFlow is a computational graph, which is a series of operations arranged in a directed acyclic graph. This graph represents the mathematical operations involved in the model, where the nodes represent mathematical operations and the edges represent the data that flows between them. and the TensorFlow runtime executes these operations efficiently on different devices [57]. Overall, TensorFlow is a powerful tool for building and training deep learning models, and its popularity and community support make it an excellent choice for machine learning practitioners.

3.3.5 OpenCV

OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. It is open source and free for both academic and commercial use. OpenCV was originally developed by Intel in 1999 and is now maintained by the OpenCV community. The OpenCV library has more than 2500 optimized computer vision algorithms, including image and video loading and saving, image filtering, feature detection, object detection and recognition, camera calibration, and many more. It supports various programming languages such as C++, Python, Java, and MATLAB. OpenCV has a large and active user community, and it is widely used in industries such as robotics, automotive, healthcare, and security, as well as in academic research. It can be used to solve a wide range of computer vision problems, from basic tasks such as face detection and object tracking to more advanced applications such as 3D reconstruction and augmented reality [47].

3.3.6 Keras

Keras is an open-source deep learning library written in Python that allows developers to easily build and train neural networks. It was developed by François Chollet and was released in 2015. Keras provides a high-level API that abstracts away many of the implementation details of building and training neural networks, making it easier for developers to focus on the design and architecture of their models [46]. Keras can run on top of various backend engines such as TensorFlow, Theano, and CNTK. It supports a wide range of neural network architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and combinations of both. Keras also provides a range of tools for data preprocessing and augmentation, model evaluation, and visualization [18]. Keras has become a popular choice for building deep learning models due to its ease of use and versatility. Its high-level API allows developers to quickly prototype and iterate on their models, while its modular design makes it easy to customize and extend to meet specific requirements [46].

3.4 Implementation Framework

Implementing AI in Saharan agriculture requires a comprehensive framework that takes into account the specific needs and challenges of the region. Here are some components that could be included in such a framework:

1. **Data collection and analysis:** The first step in implementing AI in Saharan agriculture is to collect data on the region's climate, soil, and other relevant factors. This data can then be analyzed to identify patterns and trends that can inform agricultural practices.

2. **AI algorithms and models:** Once the data has been collected and analyzed, AI algorithms and models can be developed to provide insights and recommendations for farmers. These algorithms and models can take into account factors such as weather patterns, soil conditions, and crop yields to help farmers make better decisions.

3. **IoT devices:** The use of IoT devices such as sensors and drones can provide real-time data on weather conditions, soil moisture levels, and crop health. This data can be used to inform AI algorithms and models, as well as provide farmers with actionable information.

4. **Mobile applications:** Mobile applications can be developed to provide farmers with access to AI-powered insights and recommendations. These applications can

also provide farmers with real-time alerts and notifications based on data collected by IoT devices.

5. **Training and education:** It is important to provide training and education to farmers on how to use AI-powered tools and technologies. This can include workshops, webinars, and other forms of training to ensure that farmers are able to make the most of these tools.

6. **Stakeholder collaboration:** Collaboration between farmers, agricultural experts, researchers, and policymakers is essential for the successful implementation of AI in Saharan agriculture. This collaboration can help ensure that the needs of all stakeholders are taken into account and that the implementation of AI is aligned with broader agricultural goals.

Overall, the framework for implementing AI in Saharan agriculture should be tailored to the specific needs and challenges of the region. By taking into account the unique characteristics of the region, such a framework can help improve agricultural practices and increase yields in a sustainable manner.

Chapter 4

Case Study: Implementing AI in Saharan Agriculture

4 Case Study: Implementing AI in Saharan Agriculture

This chapter offers a comprehensive examination of the application of artificial intelligence (AI) in the agricultural sector of the Saharan region. It encompasses an in-depth study of the designated area and provides detailed information on data collection and preprocessing for the AI implementation. Additionally, it outlines the specific AI techniques employed in the case study. The chapter further presents the outcomes derived from the AI implementation, subjecting them to thorough analysis. Furthermore, it explores the implications of these results and interprets them within the context of the research questions at hand.

4.1 Description of the Study Area

We can say that agriculture is a vital activity that allowed us to build prosperous societies and feed ourselves. It has spread throughout the world and is abundant in areas with favorable weather conditions, fertile lands, and access to water. Despite the challenges of harsh environments such as the Sahara, communities have adapted and developed specialized farming techniques. Our project focuses on Saharan agriculture in Africa, specifically in Algeria, the state of El Oued, where its location appears on the map. Figure 15 and aims to use AI to develop innovative solutions that will improve the productivity, sustainability, and resilience of desert agriculture.

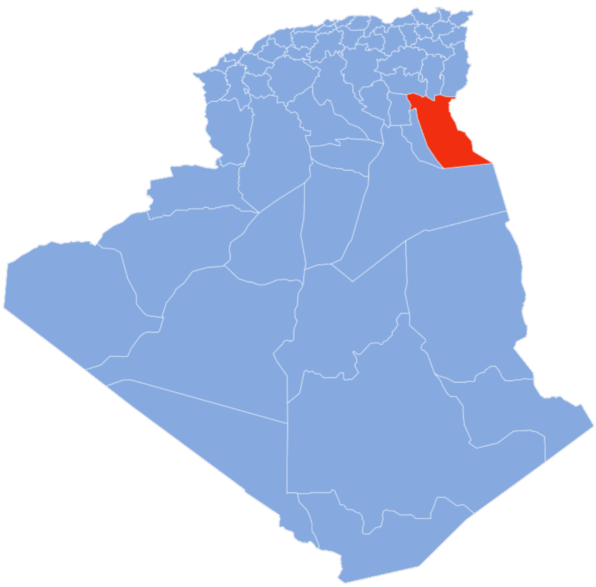

Figure 15: Map of Oued Souf Algeria.

4.2 Data Collection and Preprocessing

The title provides details of the data collection and related preprocessing.

4.2.1 Data Collection

Data collection forms the bedrock of successful AI implementation, pivotal in developing accurate and reliable models. High-quality data is the fuel that drives AI algorithms, enabling them to learn patterns, make predictions, and provide valuable insights. By meticulously collecting relevant and representative data, organizations can lay the foundation for robust AI systems that deliver enhanced performance. Data collection allows for acquiring real-world information, empowering AI models to learn from diverse scenarios and adapt to changing circumstances. It provides the raw material from which AI algorithms can derive meaningful insights, uncover hidden patterns, and make informed decisions. Ultimately, data collection is the crucial first step toward building AI systems that can drive innovation, optimize processes, and deliver transformative outcomes across various domains and industries [44].

- **The most important information and tools for data collection**
  1. **Place:** Al-Ayashi farm located in Al-Mana'a area, Al-Mogran municipality, El Oued state, Algeria, about 15 km from the municipality headquarters.
  2. **The geographical location of the farm:** https://goo.gl/maps/pqpYipz6soGGgQqF8 [58]
  3. **The date of data collection:**
     - **the first day:** 2023-03-19.
     - **the second day:** 2023-03-21.
     - **the third day:** 2023-03-22.
  4. **Imaging devices:**
     - Camera Canon
     - Redmi Note 9s
     - Samsung Galaxy A50
  5. **The number of images collected:** 2095 images were collected.
     - 1146 photos of healthy leaves
     - 946 pictures of diseased leaves
  6. **Image dimensions of each device:**
     - Camera Canon (1728 x 2592)
     - Redmi Note 9s(3000 x 4000)
     - Samsung Galaxy A50(3024 x 4032)
  7. **Image file size:** 7.6733 GB.

4.2.2 Pre-processing

Pre-processing is crucial in implementing AI as it prepares raw data for analysis and modeling, ensuring accuracy, reliability, and optimal performance of AI algorithms. By cleaning, transforming, and organizing data, pre-processing eliminates noise, handles missing values, and deals with outliers, resulting in a high-quality dataset. Pre-processing techniques such as data normalization, standardization, and dimensionality reduction enhance the efficiency and effectiveness of AI models. In essence, pre-processing maximizes the value of data, enabling AI systems to make accurate decisions, derive actionable insights, and drive impactful outcomes in various domains.In the pre-processing stage, the implementation encompassed three fundamental operations, namely data cleaning, partitioning, and dimension reduction.

1. **Data cleaning process:** During the data cleaning process, Get rid of images that do not carry information useful to the model during training, along with images exhibiting substantial noise, in order to enhance the quality and relevance of the dataset for training purposes as depicted in Figure 16.

Figure 16: Data cleaning process.

Following the completion of the data cleaning process, the study relied on a reduced set of 1,700 images instead of the initially available 2,095 images.

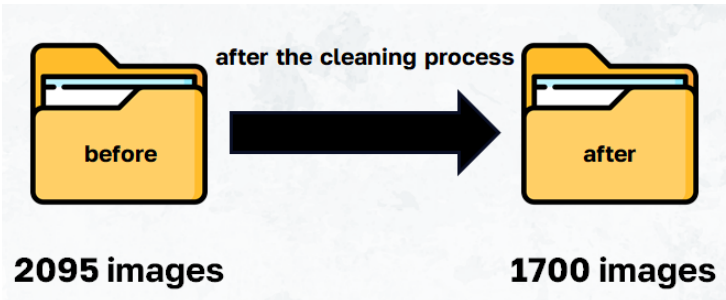

Figure 17: Quantity of photos pre- and post-cleaning procedure.

2. **Data segmentation process:** The data segmentation process holds particular significance when creating a deep learning model. By segmenting the dataset into distinct subsets or categories, it enables the model to focus on specific patterns and features within each segment.The data segmentation process involved the creat of three primary files: Training, Testing, and Verification. The dataset was partitioned into two distinct categories, namely healthy leaves and diseased leaves, within each of these files. This segmentation was crucial to ensure a balanced representation of both healthy and diseased instances in the training, testing, and verification sets. By separating the data into these categories, the model was trained on a diverse range of leaf samples, allowing it to learn and distinguish between healthy and diseased leaf patterns effectively. The testing and verification sets served to evaluate the model's performance.After the image cleaning and partitioning process, we obtained the results shown in Table 1, and Figure 18 shows the partitioning by percentage.

Table 1: Data Partition Numbers.

Partition	Number of Data
Tomato_healthy of train	650 images
Tomato_diseased of train	650 images
Tomato_healthy of test	180 images
Tomato_diseased of test	180 images
Tomato_healthy of valid	20 images
Tomato_diseased of valid	20 images

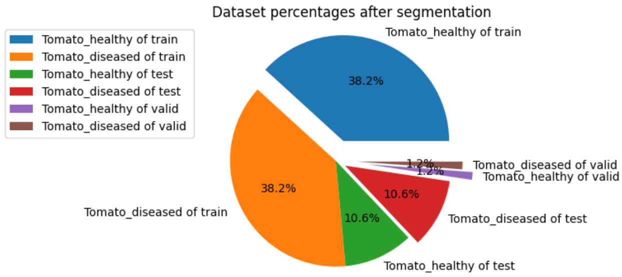

Figure 18: Dataset percentages after segmentation.

3. **Data Dimensionality Reduction:** Reducing the dimensions of image data holds immense importance when training CNN deep learning models. Images are typically high-dimensional data with multiple channels and pixels, making them computationally intensive to process. Dimension reduction techniques, such as resizing, cropping, or applying convolutional operations, are employed to reduce the spatial dimensions of the images while retaining the most salient visual features. This reduction not only enhances computational efficiency but also addresses the challenge of overfitting by reducing the model's parameter space. Additionally, reducing image dimensions can help alleviate the limitations of limited computational resources and memory constraints. Despite the dimension reduction, the key visual characteristics and patterns relevant to the task at hand are preserved, enabling the CNN model to effectively learn and recognize complex image representations. In summary, dimension reduction of image data is crucial for optimizing the performance, efficiency, and scalability of CNN deep learning models while retaining essential visual information.In our research, we employed the Python programming language to address the task of dimension reduction. To accomplish this, we developed a dedicated function capable of transforming and reducing the dimensions of the images as depicted in Figure 19. It is important to note that all the images utilized in our study were standardized to possess the same dimensions.

```python
from PIL import Image
import glob
import cv2 as cv

i=0
path1 = glob.glob("C:/MyFolder/CodeData/OriginalData/Tomato___healthy/*")
for img in path1:
    image = Image.open(img)
    print(f"Original size : {image.size}") # 3000x4000
    i=i+1
    image_resized = image.resize((256, 192))
    image_resized.save("C:\MyFolder\CodeData\OriginalData\image"+str(i)+".jpg")
    image = cv.imread("C:\MyFolder\CodeData\OriginalData\image"+str(i)+".jpg")
    image_rotate_90_clockwise = cv.rotate(image, cv.ROTATE_90_CLOCKWISE)
    cv.imwrite("C:\MyFolder\CodeData\OriginalData\test\image"+str(i)+".jpg", image_rotate_90_clockwise)
print("Okey")
```

Python

Figure 19: Code to resize and rotate images.

The first part of the figure 19 represents to import the necessary modules for working with images in Python. The PIL module (Python Imaging Library) provides functionality for opening, manipulating, and saving many different image file formats. The glob module is used for searching files. Lastly, cv2 is the module from OpenCV, which is a library used for computer vision tasks, including image processing and analysis.
In the second part as shown in Figure 19, the code takes a set of original images, resizes them, saves the resized versions, and then rotates the resized images by 90 degrees clockwise, saving the rotated images to a different directory.
Following the completion of the pre-processing stage, the data files exhibit the depicted format as presented in Figure 20

```
Tomato_data
    ├──train
    │   ├──Tomato___diseased
    │   │   ├──image1.jpg
    │   │   ...
    │   │
    │   └──Tomato___healthy
    │        ├──image1.jpg
    │       ...
    ├──test
    │    ├──Tomato___diseased
    │    │   ├──image651.jpg
    │    │   ...
    │    │
    │    └──Tomato___healthy
    │        ├──image651.jpg
    │         ...
    └──valid
        ├──Tomato___diseased
        │   ├──image831.jpg
        │   ...
        │
        └──Tomato___healthy
            ├──image831.jpg
            ...
```

Figure 20: Image file path.

4.3 AI Techniques Used

In this research, we will explore the AI techniques employed in a specific task, focusing on the utilization of Convolutional Neural Networks (CNN) with the popular deep learning frameworks, Tensorflow and Keras. CNNs have proven to be highly effective in various computer vision tasks, including image classification, object detection, and image segmentation. By applying several CNN models with different layers, we aim to enhance their capabilities for our specific application and validate the data collected for the study. Figure 21 shows an example of the layers selected and used in one of the models that were used in our research.

```python
import time
import numpy as np
import matplotlib.pyplot as plt
import tensorflow as tf

model = tf.keras.Sequential([
    data_augmentation,
    tf.keras.layers.Conv2D(60, 3, activation='relu'),
    tf.keras.layers.MaxPool2D(pool_size=2),
    tf.keras.layers.Conv2D(60, 3, activation='relu'),
    tf.keras.layers.MaxPool2D(pool_size=2),
    tf.keras.layers.Conv2D(60, 3, activation='relu'),
    tf.keras.layers.MaxPool2D(pool_size=2),
    tf.keras.layers.GlobalAveragePooling2D(),
    tf.keras.layers.Dense(total_classes, activation='sigmoid', name='output_layers')

])
```
Python

Figure 21: One of the CNN models used in the research.

To begin, we import the necessary libraries, including TensorFlow, NumPy, Time and Matplotlib as shown in Figure 21. TensorFlow is an open-source deep learning library that provides a flexible platform for building and deploying machine learning models. NumPy is a fundamental package for scientific computing in Python, enabling efficient numerical operations, Matplotlib is a powerful data visualization library and The time module in Python provides functions for working with time-related operations.After we import the libraries we prepare the image datasets, then we create the model and train it.

4.3.1 Models Architecture

We constructed our CNN models using the Keras API, which is a high-level neural networks API written in Python and designed to simplify the process of building deep learning models. The models architecture consists of a series of convolutional layers, followed by pooling layers, and concludes with a global average pooling layer and a fully connected layer.
The sequential models is built using the tf.keras.Sequential class, which allows us to stack layers in a linear manner.

1. **Data Augmentation:** Data augmentation is an essential technique used to increase the diversity of training data and reduce overfitting. It helps improve the model's generalization ability by applying random transformations to the input images, such as rotation, scaling, and flipping.

2. **Convolutional Layers:** Convolutional layers are the core building blocks of a CNN. They extract spatial hierarchies of features from the input data. The models

employed in this research uses many sets of convolutional layers. Each set comprises a Conv2D layer, followed by a MaxPooling2D layer for downsampling, and sometimes a Dropout layer for regularization. The activation function used is ReLU (Rectified Linear Unit) as shown in the figure 21

3. **Global Average Pooling:** After the convolutional layers, a GlobalAveragePooling2D layer is added. This layer reduces the spatial dimensions of the features while preserving important information. It computes the average value of each feature map, resulting in a global representation of the input volume.

4. **Flattening:** The Flatten layer reshapes the output from the previous layer into a one-dimensional vector, preparing it for the subsequent fully connected layers.

5. **Fully Connected Layers:** The flattened output is connected to a series of dense (fully connected) layers, which enable higher-level feature representation and classification. The activation function used in these layers is ReLU.sometimes Dropout layers are inserted between the dense layers.

6. **Output Layer:** The final dense layer consists of neurons equal to the total number of classes in the classification task. The activation function used is sigmoid, allowing the model to output probabilities for each class independently.

The following are the sources of the CNN models used in this research:

- **Model 1:** This model is designed for the detection of tomato leaf disease. they were designed by Antonio Guerrero-Ibañez et al., taken from an article entitled 'Monitoring Tomato Leaf Disease through Convolutional Neural Networks' [41].

- **Model 2 and 3:** Those models is designed for the detection of tomato leaf disease. they were designed by Hasan Ulutaş et al., taken from an article entitled 'Design of Efficient Methods for the Detection of Tomato Leaf Disease Utilizing Proposed Ensemble CNN Model' [43].

- **Model 4:** This model was designed to detect potato leaf disease. It was taken from kaggle and this is the link to the form on the site [54].

Parameter setting and information on layers for these models, as shown in the tables 2, 3.

Table 2: Parameter setting.

Training details	Parameters	Mode 1, 2, 3, 4
Input Shape	input size	(256, 192, 3)
Epoch	Number of epoch	100
Batch size	Number of batch size	32
Training compile	Optimizer	Adam
	metrics	accuracy
	loss function	Binary Cross-Entropy
Early Stopping	monitor	val_loss
	patience	15
Activation	Conv2D	relu
	Dense	sigmoid

Table 3: Information on the layers structure of the proposed models.

Layers	parameters	Model 1	Model 2	Model 3	Model 4
Convolution	Number of convolution layers	4	4	5	3
Pooling	Number of pooling layers	4	4	5	3
Dropout	Number of dropout layers	1	3	0	0
Flatten	Number of flatten layers	0	1	1	0
Fully connected	Number of fully connected layers	0	1	1	0
Global Average Pooling	Number of GAP layers	1	0	0	1
Convolution	Number of filters	128, 64, 32, 16	32, 64, 128, 256	16, 32, 64, 64,32	60, 60, 60
Convolution	Filter size	3×3	3×3	3×3	3×3
Pooling	Filter size	2×2	2×2	2×2	2×2
Dropout	Dropout rate	0.2	0.3, 0.4, 0.5	—	—
Fully connected	Number of neurons	—	128	512	—

4.3.2 Working Environment

In our working environment, we utilize a device of type mentioned in Table 4, along with its corresponding characteristics. Additionally, we leverage the power of Python 3 as the programming language, and we find the code editor Visual Studio Code to be highly beneficial in our tasks.

Table 4: Configuration of the hardware.

Name	Parameter
Memory	16 GB
Processor	Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz
Server model	HP 8th Gen
OS	Windows 10
Language	Python 3.11.0
Code editor	Visual Studio Code

4.4 Results and Analysis

After the deep learning models is created and trained on the dataset, we will delve into the learning outcomes of the Convolutional Neural Network (CNN) models. In this analysis, we will meticulously examine various metrics including training accuracy, loss, validation accuracy, and validation loss. Through this comprehensive evaluation, we aim to provide valuable insights and feedback on the performance of these models.

4.4.1 Evaluation Metrics

Before delving into the analysis, it is important to provide a brief definition of evaluation metrics such as loss, accuracy, recall, and precision. These metrics serve as quantitative indicators of the model's effectiveness in solving the given task. Understanding these evaluation metrics is crucial in assessing the capabilities of the CNN models and interpreting their results effectively.

1. **Accuracy:** Accuracy is a measure of how many predictions your model got correct compared to the total number of predictions. It is calculated by dividing the number of correct predictions (true positives and true negatives) by the total number of predictions [38].The formula for accuracy is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2. **Recall:** Recall measures the proportion of true positive instances correctly identified by the model over the total number of actual positive instances[38].The formula for recall is:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

3. **Precision:** Precision measures the proportion of true positive instances correctly identified by the model over the total number of instances predicted as positive [38].The formula for precision is:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Where:

TP : True Positives (Number of correctly predicted positive instances)
TN : True Negatives (Number of correctly predicted negative instances)
FP : False Positives (Number of incorrectly predicted positive instances)
FN : False Negatives (Number of incorrectly predicted negative instances)

4. **Loss:** Loss is a measure of how well the model's predictions match the true values. The choice of the loss function depends on the specific problem and algorithm used. One commonly used loss function for binary classification is binary cross-entropy loss [27]. The formula for Binary Cross-Entropy Loss is:

$$\text{Binary Cross-Entropy Loss} = -(y \cdot \log(y_{hat}) + (1 - y) \cdot \log(1 - y_{hat})) \tag{4}$$

Where:

$$y : \text{True label (0 or 1)}$$
$$y_{hat} : \text{Predicted probability of the positive class (between 0 and 1)}$$

4.4.2 Analyze the results of the models

The focus of this title revolves around showcasing the outcomes of training carried out on four distinct models, accompanied by detailed notes and an all-encompassing analysis of the obtained results. These results. Graphical curves are utilized to visually represent the performance of the models, along with presenting the evaluation results found in Table 4. These outcomes are depicted in terms of training time, the number of sessions, the peak training accuracy, and other relevant.

First the accuracy results are given in Table 5 with the models used.

Table 5: Performance metrics of CNN models.

Parameters	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Epochs to Train the Model	69	64	52	64
Training Time (in minutes)	143.5	77.16	27.72	78.24
Training Accuracy (max)	97.62%	99.4%	99.61%	97%
Validation Accuracy (max)	98.61%	98.01%	98.58%	98.33%
Training Recall (max)	97.46%	99.4%	99.61%	97.08%
Training Precision (max)	97.38%	99.4%	99.61%	97.08%
Training Loss (min)	7.8%	1.73%	1.43%	9.96%
Validation Loss (min)	5.09%	5.6%	4.4%	6.06%
evaluation Accuracy (validation data)	100%	100%	100%	97.5%

Next, we will examine the curves representing the results for loss, val_loss, accuracy, and val_accuracy during the training of the CNN models. Each curve will be analyzed individually.

- **Model 1**

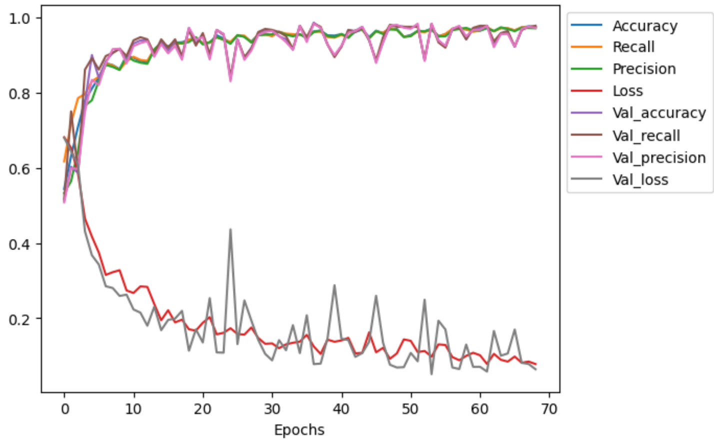

Figure 22: Training Results Model 1.

Based on the results presented in Figure 22, the CNN model demonstrated a satisfactory performance. Here are the key observations and conclusions:

- Throughout the training period, the accuracy, precision, and recall curves consistently exhibited a well-balanced predictive ability.
- these metrics gradually improved and reached a peak of approximately 97.62%, indicating the model's effective learning and enhanced classification capabilities. The gradual decrease in loss from 68.13% to 7.8% further confirmed the model's proximity to the true values.
- The evaluation on unseen data, represented by validation accuracy, loss, and accuracy/recall metrics, provided valuable insights. The validation accuracy, precision, and recall achieved a high value of around 98%, similar to the training accuracy. The validation loss decreased from 67.97% to 5.09%, demonstrating the model's generalization to unseen data without overfitting the training set.
- It is observed that the training did not reach 100 epochs. This is because the EarlyStopping callback was used with a patience of 15. The callback monitors the validation loss and restores the best weights when the loss stops improving for 15 consecutive epochs. It helps in preventing overfitting and saves computation time by stopping the training early when no significant improvement is observed.

- **Model 2**

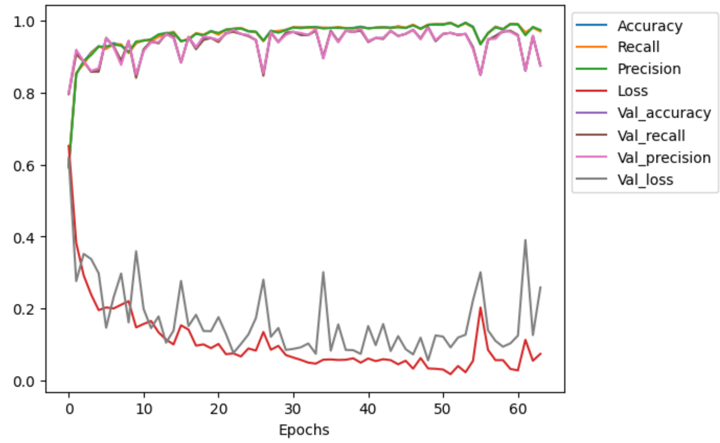

Figure 23: Training Results Model 2.

According to the findings illustrated in Figure 23, the CNN model exhibited a satisfactory level of performance. The following are the main observations and conclusions derived from the results:

- Throughout the entire duration of the training period, it can be observed that the accuracy, precision, and recall curves consistently demonstrated a commendable and well-balanced predictive ability. These curves, serving as reliable indicators of the model's performance, maintained a stable and desirable level throughout the training process.
- It can be concluded that the model consistently improved its accuracy during the training process. Initially starting at approximately 59.15%, the accuracy rapidly increased in the early stages and continued to steadily rise over time, ultimately reaching a peak of around 99.4%. This signifies that the model effectively learned and performed well on the training data. The validation accuracy followed a similar trend to the training accuracy. Beginning at approximately 79.83%, it displayed progressive enhancement over time and reached a peak of 98.01%. This indicates that the model exhibits strong generalization capabilities and effectively handles unseen data.
- Throughout the training process, the training loss consistently decreased, eventually settling around 1.73%. Similarly, the validation loss followed a comparable pattern, albeit with minor fluctuations, steadily decreasing and reaching a low value. These observations suggest that the model reduced its errors and made improved predictions.
- Overall, the presented findings affirm that the CNN model was successfully trained. It demonstrated commendable learning and generalization abilities, as indicated by the increased accuracy and reduced loss values. The model showcased promising performance on both the training and validation sets, highlighting its potential effectiveness.

- **Model 3**

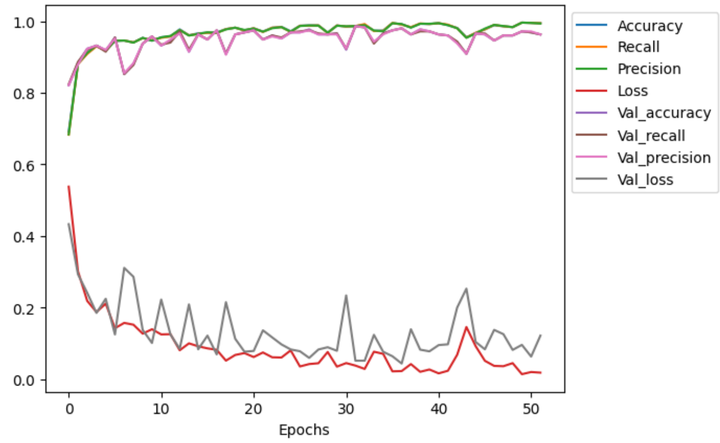

Figure 24: Training Results Model 3.

Based on the results presented in Figure 24, it can be concluded that the CNN model showcased a satisfactory level of performance. The following observations and conclusions can be drawn from the obtained results:

- During the entire training period, a notable and well-balanced predictive ability was consistently observed in the accuracy, precision, and recall curves. These curves, which serve as reliable metrics for evaluating the model's performance, exhibited a stable and desirable level throughout the training process.
- It can be said that the model exhibits consistent improvement in accuracy on the training set, starting from 69.16% and peaking at 99.61%. This indicates that the model effectively learns and generates accurate predictions for the training data. Additionally, the validation accuracy demonstrates the model's performance on unseen data, with an initial value of approximately 82.1% and a steady increase to a maximum of around 98.58%. These results suggest that the model generalizes well and can make accurate predictions when presented with new, unseen examples.
- The loss metric, representing the error during training, decreases consistently throughout the training process. It begins at 53.8% and reaches a minimum of 1.43%, indicating that the model effectively adjusts its parameters to minimize the prediction error. Similarly, the validation loss exhibits a decreasing trend with fluctuations, ultimately reaching a minimum of approximately 4.4%. This further supports the model's ability to perform well on unseen data.
- Overall, these results strongly indicate that the CNN model learns efficiently and generalizes effectively to previously unseen data. The combination of high accuracy and low loss in both the training and validation sets demonstrates the model's ability to accurately classify images.

- **Model 4**

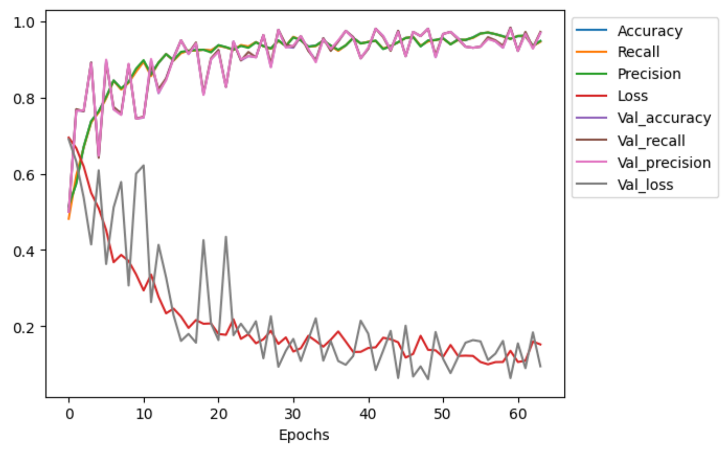

Figure 25: Training Results Model 4.

The findings depicted in Figure 25 lead to the conclusion that the CNN model demonstrated a satisfactory level of performance. Several noteworthy observations and conclusions can be drawn from the obtained results:

- Throughout the entire training period, the accuracy, precision, and recall curves consistently showcased a remarkable and well-balanced predictive ability. These curves, serving as reliable metrics to assess the model's performance, remained stable and desirable across the entire training process.
- The results presented demonstrate a consistent improvement in the model's accuracy over time. Initially, the accuracy starts at approximately 51.3% and gradually rises to a peak of around 97%. Similarly, the validation accuracy follows a similar trend, starting around 69.5% and steadily increasing to about 98.33%. The noteworthy aspect is that the validation accuracy closely matches the training accuracy, suggesting that the model does not suffer from overfitting and can effectively generalize to unseen data.
- Examining the loss metrics, the training loss begins at a relatively high value of approximately 69% and consistently decreases throughout the model's training process, eventually reaching a minimal value of about 6.06%. The validation loss exhibits a similar pattern, although with some fluctuations, ultimately reaching a low value. This behavior indicates that the model successfully minimizes the loss function and acquires the ability to make accurate predictions.
- In general, the model exhibits a positive performance trend, with increasing accuracy and decreasing loss over time. The model's high accuracy, along with comparable validation accuracy and reduction in loss, indicates its capacity to effectively learn from the data and generalize well.

4.4.3 Compare CNN models

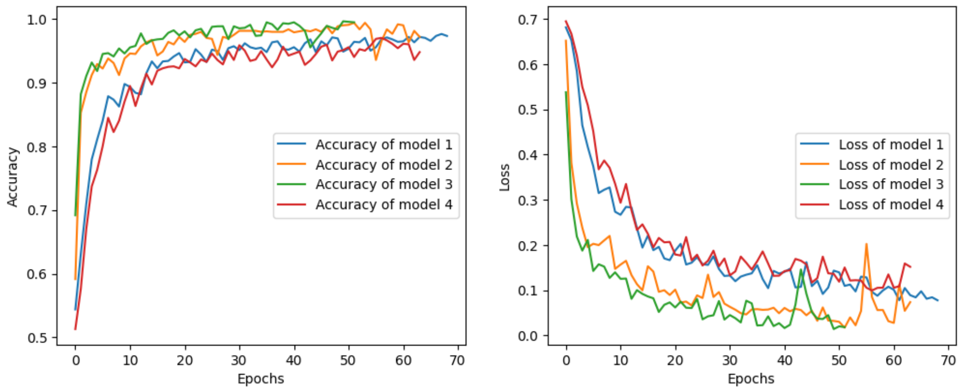

Figure 26: Comparison of models CNN (Training Accuracy and loss).

Based on the information presented in Figure 26, we can make the following observations and draw conclusions:

- All four models exhibit a downward trend in loss, indicating effective learning and a good fit to the dataset. Models 2 and 3 consistently achieve lower loss values compared to Models 1 and 4, with Model 2 reaching around 1.73% and Model 3 reaching around 1.43%. This suggests that Models 2 and 3 perform better in minimizing the discrepancy between expected and actual values. Additionally, Models 2 and 3 reach the minimum loss value more quickly than Models 1 and 4. Models 1 and 4 display similar loss patterns, with Model 1 reaching approximately 7.8% and Model 4 reaching around 9.96%, albeit at a relatively slower rate of improvement compared to Models 2 and 3.

- Model 3 appears to achieve the lowest loss value among the four models in the shortest amount of time.

- All four models demonstrate an improvement in accuracy throughout the training process. Models 2 and 3 consistently perform well, reaching a peak accuracy of over 99.4% in fewer training periods. This indicates that these models are effectively learning and improving their ability to correctly classify the training data. Model 1 achieves a slightly lower but still commendable accuracy of 97.62%. Initially, Model 4 lags behind the others, but it gradually catches up and achieves a respectable accuracy of 97%.

- Once again, it is worth noting that Model 3 attains the highest accuracy value in a shorter period compared to the other three models.

- It is important to acknowledge that the training periods for these models varied due to the use of early stopping. While this difference affected the number of epochs and training duration, it did not necessarily impact the final accuracy achieved by each model.

In summary, the analysis of the CNN models reveals that all four models exhibit a downward trend in loss and improvement in accuracy throughout the training process , indicating effective learning and a good fit to the dataset. Models 2 and 3 consistently outperform Models 1 and 4 in terms of minimizing loss and achieving higher accuracy.

Model 3 stands out as the top performer, achieving the lowest loss value and highest accuracy in a shorter training period compared to the other models. Overall, the results demonstrate the effectiveness of the CNN models in learning and accurately classifying the training data, with slight variations in performance and convergence speed between the different models.

4.4.4 Prediction Results of Trained Model

Figure 27 represents the prediction results obtained from one of the trained models. These results provide valuable insights into the model's performance and its ability to accurately predict outcomes.

Figure 27: Model prediction results 4.

4.5 Discussion and Interpretation

In this research, we focused on the application of a deep learning technique in the context of desert agriculture. The results we obtained were highly promising, providing evidence that artificial intelligence techniques can be successfully utilized in this challenging environment. However, it is important to acknowledge and address certain obstacles that we encountered during the study, such as the scarcity of data and the inherent difficulties associated with data collection in desert regions.

The lack of data poses a significant challenge in implementing artificial intelligence techniques in desert agriculture. Desert environments are characterized by their unique and harsh conditions, making data collection a complex task. However Nevertheless, despite these challenges, we were able to surmount them by undertaking the data collection process ourselves and leveraging technologies to standardize and preprocess the collected data. This allowed us to successfully apply deep learning techniques and develop a high-quality model that effectively described the patterns and dynamics within the dataset.

Building upon our achievements, it is evident that there is tremendous potential for future progress in this field. As technologies advance and more data becomes available, the scope of research in desert agriculture can be expanded to encompass a wider range of crops and plant species. By collecting additional data and incorporating it into our models, we can refine and enhance the accuracy and robustness of the predictions and analyses.

In summary, In spite of the obstacles posed by the scarcity of available data and the harsh conditions of the desert environment, our research demonstrates the practicality and potential of employing deep learning techniques in desert agriculture.And the results obtained confirm this

Conclusion and Future Work

5 Conclusion and Future Work

To conclude, agriculture plays a crucial role in ensuring food security and supporting economic development. However, the agricultural sector faces various challenges and diseases, necessitating the implementation of protective measures. The advancement of artificial intelligence (AI) technologies has shown significant progress in leveraging AI in agriculture, particularly in developed regions like Europe, Asia, and North America. However, desert countries encounter substantial obstacles and limitations in adopting AI technologies for their agricultural practices. The objective of this research was to address these challenges and contribute to the field of desert agriculture.

Throughout this study, we explored fundamental concepts of artificial intelligence, its key techniques, and its applications, with a specific focus on its relevance to agriculture. We also outlined the methodology employed, which involved collecting images of healthy and diseased tomato plant leaves. These images were used to train a CNN AI model, resulting in promising outcomes.

This research has made noteworthy contributions and implications for desert agriculture. By harnessing AI technologies, we demonstrated the potential to enhance disease detection and prevention in crops, leading to increased yields and improved agricultural productivity in desert regions. However, it is crucial to acknowledge the limitations of this study, such as limited data availability and the use of basic computers for model training.

Based on the findings of this research, we offer several recommendations to advance the utilization of AI in desert agriculture. Firstly, there should be a concerted effort to recruit a significant number of individuals for data collection, ensuring a comprehensive and diverse dataset for training AI models. Secondly, it is crucial to use robust and long-lasting computing devices to achieve more accurate and efficient results within shorter time frames.

Looking ahead, we envision several future directions for research in the field of artificial intelligence and its applications in desert agriculture. It is imperative to explore and invest in additional areas that can benefit from AI technologies, such as automated irrigation systems and precision farming techniques. Moreover, studying the integration of AI with other emerging technologies, such as the Internet of Things (IoT) and robotics, could unlock further possibilities and advancements in desert agriculture.

To summarize, this research has shed light on the potential of artificial intelligence in desert agriculture. By implementing AI technologies, we can establish sustainable and efficient farming practices in desert regions, leading to food security and economic prosperity for these areas.

References

[1] J Deckers. "Soil fertility and environmental problems in different ecological zones of the developing countries in Sub-Saharan Africa". In: (1993), pp. 37–52.

[2] Guido Van Rossum and Fred L Drake Jr. Python tutorial. Vol. 620. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

[3] Floyd P Horn and Roger G Breeze. "Agriculture and food security". In: 894.1 (1999), pp. 9–17.

[4] Francis J. Pierce and Peter Nowak. "Aspects of Precision Agriculture". In: ed. by Donald L. Sparks. Vol. 67. Advances in Agronomy. Academic Press, 1999, pp. 1–85. DOI: https://doi.org/10.1016/S0065-2113(08)60513-1. URL: https://www.sciencedirect.com/science/article/pii/S0065211308605131.

[5] Zoubin Ghahramani. "Unsupervised learning". In: (2004), pp. 72–112.

[6] Mark Lawrence Gleason and Brooke A Edmunds. Tomato diseases and disorders. Iowa State University, University Extension Ames, IA, 2005.

[7] Seisuke Kimura and Neelima Sinha. "Tomato (Solanum lycopersicum): a model fruit-bearing crop". In: 2008.11 (2008), pdb–emo105.

[8] Hua Wang, Cuiqin Ma, and Lijuan Zhou. "A brief review of machine learning and its application". In: IEEE. 2009, pp. 1–4.

[9] Amine Boudghene Stambouli. "Promotion of renewable energies in Algeria: Strategies and perspectives". In: 15.2 (2011), pp. 1169–1181.

[10] Elodie Blanc. "The impact of climate change on crop yields in Sub-Saharan Africa". In: (2012).

[11] Mark Riedl and Vadim Bulitko. "Interactive narrative: A novel application of artificial intelligence for computer games". In: vol. 26. 1. 2012, pp. 2160–2165.

[12] Pauline Chivenge et al. "The potential role of neglected and underutilised crop species as future crops under water scarce conditions in Sub-Saharan Africa". In: 12.6 (2015), pp. 5685–5711.

[13] Martín Abadi. "TensorFlow: learning functions at scale". In: 2016, pp. 1–1.

[14] Martín Abadi et al. "Tensorflow: a system for large-scale machine learning." In: vol. 16. 2016. Savannah, GA, USA. 2016, pp. 265–283.

[15] Peter Goldsborough. "A tour of tensorflow". In: (2016).

[16] AI Strong. "Applications of artificial intelligence & associated technologies". In: 5.6 (2016).

[17] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network". In: Ieee. 2017, pp. 1–6.

[18] Nikhil Ketkar and Nikhil Ketkar. "Introduction to keras". In: (2017), pp. 97–111.

[19] Vladimir Nasteski. "An overview of the supervised machine learning methods". In: 4 (2017), pp. 51–62.

[20] Mahind Rupali and Patil Amit. "A review paper on general concepts of artificial intelligence and machine learning". In: 4.4 (2017), pp. 79–82.

[21] Shujun Huang et al. "Applications of support vector machine (SVM) learning in cancer genomics". In: 15.1 (2018), pp. 41–51.

[22] Andreas Kamilaris and Francesc X Prenafeta-Boldú. “A review of the use of convolutional neural networks in agriculture”. In: 156.3 (2018), pp. 312–322.

[23] Abebe Shimeles, Audrey Verdier-Chouchane, and Amadou Boly. “Introduction: understanding the challenges of the agricultural sector in Sub-Saharan Africa”. In: (2018), pp. 1–12.

[24] Pramila P Shinde and Seema Shah. “A review of machine learning and deep learning applications”. In: IEEE. 2018, pp. 1–6.

[25] Kun-Hsing Yu, Andrew L Beam, and Isaac S Kohane. “Artificial intelligence in healthcare”. In: 2.10 (2018), pp. 719–731.

[26] Ngozi Clara Eli-Chukwu. “Applications of artificial intelligence in agriculture: A review”. In: 9.4 (2019), pp. 4377–4383.

[27] Xuefeng Jiang et al. “Capsnet, cnn, fcn: Comparative performance evaluation for image classification”. In: 9.6 (2019), pp. 840–848.

[28] Igor Stančin and Alan Jović. “An overview and comparison of free Python libraries for data mining and big data analysis”. In: IEEE. 2019, pp. 977–982.

[29] Muhammad Shoaib Farooq et al. “Role of IoT technology in agriculture: A systematic literature review”. In: 9.2 (2020), p. 319.

[30] Batta Mahesh. “Machine learning algorithms-a review”. In: 9 (2020), pp. 381–386.

[31] Anandhavalli Muniasamy. “Machine Learning for Smart Farming: A Focus on Desert Agriculture”. In: 2020, pp. 1–5. DOI: 10.1109/ICCIT-144147971.2020.9213759.

[32] W Mupangwa et al. “Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa”. In: 2 (2020), pp. 1–14.

[33] Iva Prgomet et al. “Physiological and biochemical performance of almond trees under deficit irrigation”. In: 261 (2020), p. 108990.

[34] SC Huang and TH Le. “Introduction to TensorFlow 2”. In: (2021), pp. 1–26.

[35] Nawab Khan et al. “Current progress and future prospects of agriculture technology: Gateway to sustainable agriculture”. In: 13.9 (2021), p. 4883.

[36] Yusuf Mustafa Khan and Sandeep Kumar. “Applications of artificial intelligence in agriculture”. In: (2021).

[37] Foyez Ahmed Prodhan et al. “Deep learning for monitoring agricultural drought in South Asia using remote sensing data”. In: 13.9 (2021), p. 1715.

[38] Alper Taner, Yeşim Benal Öztekin, and Hüseyin Duran. “Performance analysis of deep learning CNN models for variety classification in hazelnut”. In: 13.12 (2021), p. 6527.

[39] Rubby Aworka et al. “Agricultural decision system based on advanced machine learning models for yield prediction: case of east african countries”. In: 2 (2022), p. 100048.

[40] A Yeshe, P Gourkhede, and P Vaidya. “Blue River Technology: Futuristic Approach of Precision Farming”. In: (2022).

[41] Antonio Guerrero-Ibañez and Angelica Reyes-Muñoz. “Monitoring Tomato Leaf Disease through Convolutional Neural Networks”. In: 12.1 (2023), p. 229.

[42] Ipsita Samal et al. “Plantix app: A success story of artificial intelligence in plant protection”. In: (2023), p. 24.

[43] Hasan Ulutaş and Veysel Aslantaş. "Design of Efficient Methods for the Detection of Tomato Leaf Disease Utilizing Proposed Ensemble CNN Model". In: 12.4 (2023), p. 827.

[44] Steven Euijong Whang et al. "Data collection and quality challenges in deep learning: A data-centric ai perspective". In: (2023), pp. 1–23.

[45] A Summary of Neural Network Layers. Available from: https://medium.com/machine-learning-for-li/different-convolutional-layers-43dc146f4d0e. Accessed May 1, 2023.

[46] About Keras . Available from. https://keras.io/about/. Accessed May 04, 2023.

[47] About OpenCV . Available from. https://opencv.org/about/. Accessed May 04, 2023.

[48] Convolutional Neural Networks (CNN): Step 3 - Flattening. Available from. https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-step-3-flattening. Accessed May 1, 2023.

[49] Deep Learning model-based Multimedia forgery detection. Available from: https://www.researchgate.net/figure/Architecture-of-CNN-20-The-architecture-of-a-CNN-is-as-follows-Input-layer-The_fig2_346036530. Accessed May 2, 2023.

[50] Deep Learning vs. Machine Learning – What's The Difference? Available from. https://levity.ai/blog/difference-machine-learning-deep-learning. Accessed April 15, 2023.

[51] Everything you need to know about Machine Learning. Available from. https://www.analyticsvidhya.com/blog/2021/03/everything-you-need-to-know-about-machine-learning/. Accessed May 1, 2023.

[52] Machine Learning, Machine Learning Algorithms, Decision Tree. Available from. https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/ml-decision-tree/tutorial/. Accessed May 1, 2023.

[53] Neural Networks: parameters and hyperparameters.Available from. https://towardsdatascience.com/neural-networks-parameters-hyperparameters-and-optimization-strategies-3f0842fac0a5. Accessed May 1, 2023.

[54] Potato-Disease(Test-data 99.67% Valid-data 98%) . Available from. https://www.kaggle.com/code/hsakash/potato-disease-test-data-99-67-valid-data-98/log. Accessed March 04, 2023.

[55] SIBI (Sistem Isyarat Bahasa Indonesia). Available from. https://www.researchgate.net/figure/Illustration-of-max-pooling-with-filter-size-2x2-and-stride-2_fig4_338845633. Accessed May 1, 2023.

[56] Support Vector Machine Algorithm. Available from. https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm. Accessed May 1, 2023.

[57] TensorFlow 1.0 vs 2.0, Part 1: Computational Graphs. Available from. https://rb.gy/jugvw. Accessed April 20, 2023.

[58] The geographical location of the farm . Available from. https://shorturl.at/ehCJ6. Accessed May 25, 2023.

[59] Tomato plant study . Available from. https://shorturl.at/cepX8. Accessed May 20, 2023.