: N° d'ordre
: N° de série

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Eduction 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:

- **Salah Eddine OMANE**
- **Abderrahmane LARGOT**

Title

An intelligent approach for early detection of potato diseases

Examination Commitee :

M.	Mohib Eddine KHEBBACHE	MAA	Président
MISS.	Chourouk GUETTAS	MAA	Examiner
M.	Adel BERHOUM	MAA	Supervisor
M.	Mohammed Charaf Eddine MEFTAH	PR	Co-Supervisor

Academic year: 2023-2024

شكر و تقــــدير

الحمد لله أولا و آخرا على توفيقه لإتمام هذا العمل .

أتقدم بالشكر الجزيل الى من تشرفت بإشرافه

الدكتور: عادل برهوم

على جميع توجيهاته ونصائحه القيمة

كما أشكر أعضاء لجنة المناقشة على تفضلهم بقبول مناقشة هذه المذكرة

و أتقدم بالشكر الجزيل إلى كل طاقم كلية العلوم الدقيقة بجامعة الشهيد

حمه لخضر بالوادي

وإلى كل من ساعدنا من قريب أو بعيد بالقليل أو بالكثير في إنجاز المذكرة

شكرا لكم جميعا

اهداء

إلى من علَّمتني العطاء ، وغمرتني بحنانها وكرمها .

إلى نبع المحبة والإيثار والكرم

أمي العزيزة

إلى من لم يدخر جهدا عنا، ولا عطاء لنا. من أرهقته الحياة لأجلنا

أبي العزيز

إلى أخوتي كل باسمه وكل بمقامه من علَّموني أن الحياة من دون ترابط وحب وتعاون لا تساوي شيئًا.

إلى خطيبتي ومن ملأت حياتي بالتحدي، وتخطِّي الصعاب.

الى كل من يعرفني من بعيد أو قريب

إلى جميع من تلقّيتُ منهم النصح والدعم

إن إنهائي عملي لم يكن ليتم لولا دعمكم، وأتمنّى أن ينال رضاكم.

عمان صلاح الدين

ملخص

سعى القطاع الزراعي بشكل دائم إلى تعزيز وتطوير أنظمته. كما هو الحال مع القطاعات الأخرى، اعتمد القطاع الزراعي في الآونة الأخيرة، مثل كافة القطاعات، على تقنيات الذكاء الاصطناعي لمعالجة البيانات الزراعية، لتحقيق إنتاج محاصيل عالي الجودة. وفي مناطق متنوعة مثل أوروبا، وأمريكا الشمالية، وشرق آسيا، تم استخدام تقنيات التعلم العميق للكشف عن أمراض النباتات، وتحديد أسبابها، وحتى التنبؤ بإنتاجية المحاصيل في مواسم محددة.

يهتم هذا البحث بتطبيق هذه التقنيات في البيئة الصحراوية لأنها مختلفة تماما من حيث نوعية التربة غير الخصبة والجفاف وارتفاع ملوحة المياه ودرجات الحرارة القصوى وغيرها. ونقترح استخدام نماذج تم تطويرها سابقا تعتمد على الشبكات العصبية التلافيفية لحل المشكلة. مشكلة درجة تطور مرض أوراق البطاطس. ينصب تركيزنا على إنشاء تطبيق لتحديد وتصنيف أمراض أوراق البطاطس باستخدام مجموعة بيانات تم جمعها خصيصًا من بيئة صحراوية محلية.

الكلمات الرئيسية : الزراعة الصحراوية، أمراض أوراق البطاطس، تقنيات الذكاء الاصطناعي، التعلم العميق، شبكات CNN .

Abstract

The agricultural sector constantly seeks to strengthen and develop its systems. As with other sectors, the agricultural sector has recently, like all sectors, relied on artificial intelligence technologies to process agricultural data, to achieve high-quality crop production. In regions as diverse as Europe, North America, and East Asia, deep learning techniques have been used to detect plant diseases, determine their causes, and even predict crop yields in specific seasons. This research is concerned with applying these techniques in the desert environment because they are completely different in terms of infertile soil quality, drought, high water salinity, extreme temperatures, etc. We propose to use previously developed models based on convolutional neural networks to solve the problem of the degree of progression of potato leaf disease. Our focus is to create an application to identify and classify potato leaf diseases using a dataset specially collected from a local desert environment. abstract environment. Desert agriculture, Potato leaf diseases, Artificial intelligence technologies, Deep Learning, CNNs

key words : Desert agriculture ,Potato leaf diseases ,Artificial intelligence
deep learning.CNNs

Table of Contents

List of Tables

List of Figures

Chapter 1

General introduction

1.1 Background & Motivation

Farming is an essential activity it has played a crucial role in the development and growth of human civilization,agriculture has spread throughout the world and has become increasingly popular due to its significance in providing food security for people [1]. 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. These imaging devices have become crucial tools for farmers, allowing them to collect data. 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.
AI algorithms can analyze large amounts of data in a short amount of time, enabling farmers to make informed decisions quickly.

1.2 Research problems & challenges

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:

Q_1: Is it possible to provide a program that helps the farmer to detect diseases early?

Q_2: Can we access a program that gives high quality results?

Q_3: Is it possible to provide a program that accepts all kinds of images in all their forms?

Q_4: 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:

Obj_1: The system must recognize all plant varieties.

Obj_2: The system accepts all types of images in any format.

Obj_3: 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 artificial intelligence system for plant disease detection. However, there are several limitations that may affect the study results. First, there are two summer and winter potato crops and a lack of data available for study, which may make it difficult to train the AI system accurately. In addition, collecting data from remote areas can be difficult, limiting the amount of data that can be collected. The imaging device used to collect data may not be the most powerful, which may 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 only certain species. Finally, the computer used to process the images may not be powerful enough to handle the amount of data required for accurate analysis. These limitations should be taken into account when interpreting study results

1.6 Research structure

The structure of this research is as follows:

- **Chapter 1: 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. and comprehensive overview of artificial intelligence techniques and tools.And give a comprehensive overview of them Use and explain its concepts.

- **Chapter 2: The proposed approach and methodology** This chapter presents an extensive study on the application of artificial intelligence in the agricultural sector in the Sahara region. It specifically identifies the artificial intelligence techniques used in the case study.
- **Chapter 3: Results and discussion**This chapter presents Results obtained from implementing artificial intelligence and a comprehensive procedure Analyze these results and provide a comprehensive picture of the environment for implementing this work
- **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

2.1 General Introduction

In this chapter, we wanted to emphasize and deepen our consideration of artificial intelligence and its applications, focusing on its many uses in agriculture, whatever it may be, especially desert ones. By understanding the applications, challenges and opportunities, as well as building on previous research, we aim to pave the way for new approaches and strategies that harness the power of artificial intelligence to overcome the obstacles of desert agriculture.

2.2 Research Design

In this section, we will delve into a comprehensive elucidation of our research design phase, illustrated in Figure **??**. 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.

2.3 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 [2].

2.4 Artificial intelligence applications

Artificial Intelligence is a term in the form of neural networks and specialized systems, which has wide-ranging uses. It is used in almost all human fields. What proves her high heels is her high accuracy in work. Expert robot systems already exist Taking on shop-level jobs in large industries, which has greatly marginalized the human factor. In addition, for example, financial brokerage companies use artificial intelligence to analyze data, conduct investment analyzes and buy or sell stocks without any human intervention. We also provide below some examples of artificial intelligence applications [2].

- **Precision agriculture:** Optimises agricultural production through the analysis of climate data, crop monitoring, efficient irrigation water management and early detection of plant diseases. In this way, measures against pests can be implemented before it is too late, for example[3].

- **Machine translation:** Allows the translation of text or voice from one language to another automatically and quickly. It is used in online translators, mobile applications, content translation or in global human-to-human communication. Although its quality and accuracy have improved, nuances, localisms and cultural context require the concurrence of human translators[4].

- **Industrial automation and robotics:** Used to convert certain industrial processes into automated activities, thereby improving manufacturing efficiency. Robots and autonomous systems use AI algorithms to execute complex tasks efficiently and safely. They are used, for example, in manufacturing and assembly tasks. On the other hand, automation is also common in preventive maintenance of plant and machinery, process control, logistics and inventory management.[5].

- **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 [6].

- **Video Games** artificial intelligence (AI) is used to generate responsive, adaptive or intelligent behaviors primarily in non-playable characters (NPCs) similar to human-like intelligence. Artificial intelligence has been an integral part of video games since their inception in the 1950s.[1] AI in video games is a distinct subfield and differs from academic AI. It serves to improve the game-player experience rather than machine learning or decision making. During the golden age of arcade video games the idea of AI opponents was largely popularized in the form of graduated difficulty levels, distinct movement patterns, and in-game events dependent on the player's input. Modern games often implement existing techniques such as pathfinding and decision trees to guide the actions of NPCs. AI is often used in mechanisms which are not immediately visible to the user, such as data mining and procedural-content generation[7]. There are many applications for AI in manufacturing as industrial IoT and smart factories generate large amounts of data daily. AI in manufacturing is the use of machine learning (ML) solutions and deep learning neural networks to optimize manufacturing processes with improved data analysis and decision-making.

- **Manufacturing** A commonly cited AI use case in manufacturing is predictive maintenance. By applying AI to manufacturing data, companies can better predict and prevent machine failure. This in turn reduces expensive downtime in manufacturing processes. AI in manufacturing has many other potential uses and benefits, such as improved demand forecasting and reduced waste of raw materials. AI and manufacturing have a natural relationship since industrial manufacturing settings already require people and machines to work closely together[8].

2.4.1 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[7].

- **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 [9].

- **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 [10]. 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 [11]. These advancements in agricultural robotics not only address the labor shortage but also improve crop yields and reduce costs [9].

- **Automated Crop Irrigation System:** AI-based software could analyze soil moisture levels and predict upcoming rainfall, adjusting irrigation accordingly to give crops the water they need. It conserves water and ensures that crops are neither overwatered nor underwatered, promoting healthier growth and higher yields. Such intelligent irrigation systems greatly improve water use efficiency, a critical factor in sustainable agricultural practices. Imagine a vineyard with an AI-based irrigation system. It constantly monitors soil moisture and predicts weather patterns. The system can then automatically adjust the amount of water delivered to each vine. Each plant gets exactly the amount it needs for optimal growth. [12].

- **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 [7].

2.4.2 Challenges and Opportunities in Saharan Agriculture

While all businesses and projects face challenges and difficulties, desert agriculture must also have its share of challenges and obstacles, as we realize that the region is characterized by harsh climatic conditions, including sharply high and scarce temperatures and lack of rainfall, or as we say. The semi-dryness that the soil suffers from. But through science and the development of studies, there are innovative technologies that can overcome these challenges and create opportunities for sustainable agricultural production in the region.

Challenges:

While we are talking about desert agriculture, we must mention some of the challenges and difficulties. As shown in the table below:

Table 2.1: Challenges for developing agriculture in the Sahara Desert.

Challenges	**the explanation**
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 [13].
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 [14].
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 [15].

Opportunities:

Despite these challenges, there are opportunities to develop sustainable agriculture in the Sahara Desert, as shown in the table below:

Table 2.2: Opportunities for developing agriculture in the Sahara Desert.

Opportunities	**the explanation**
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.[16]
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 [17].
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.[18].
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 [19]. These crops can provide a good source of income for farmers in the region.

2.4.3 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:

Table 2.3: Previous Studies on AI in Saharan Agriculture.

scientific researcher	**Research Method**	**the Objectif of research**
Anandhavalli Muniasamy [20]	Machine Learning for Smart Farming	Its goal is to produce food due to population growth, scarcity of natural resources, climate change, and the need for timely and useful agricultural data to develop strong strategies and practical decisions.,
The article Mupangwa [21]	Evaluating machine learning algorithms for predicting maize yield.	To focus on corn production forecasting Different agricultural systems for grid-based and reduced conservation agriculture. Focus on studying to explore Using learning techniques to accurately predict corn products and determine the algorithm Corn is produced better than others. .
The article by Foyez Ahmed and Jiahua Zhang [22]	Deep Learning for Monitoring Agricultural.	To determine the start accurately and in almost real time Drought severity in South Asia. Explaining the challenges associated with monitoring drought conditions using traditional meteorological devices Logical monitoring and remote sensing methods..
Rubby Aworka [23]	Machine Learning models for yield prediction.	To improve performance Resource Utilization and Increasing Crop Yield They proposed using Machine Learning models to predict crop productivity in East Africa countries..

2.5 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.

2.5.1 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 [24].

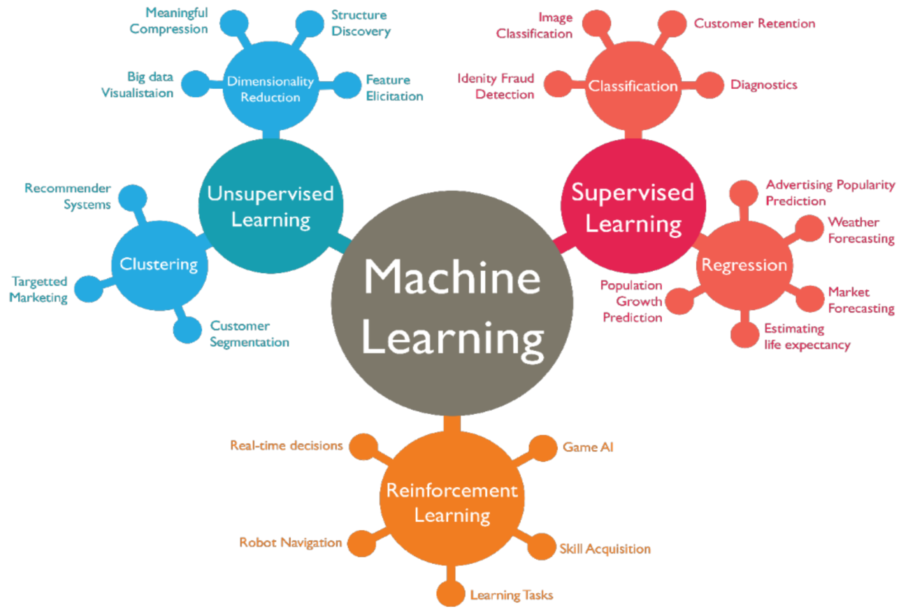

Figure 2.1: Types of machine learning and its algorithms.

2.5.2 Types of machine learning

There are three main types of machine learning as depicted in Figure 2.1: supervised learning, unsupervised learning, and reinforcement learning [25].

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 [26].

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 [27].

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 [28].

2.5.3 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.

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 2.2 [29]. 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 [29]

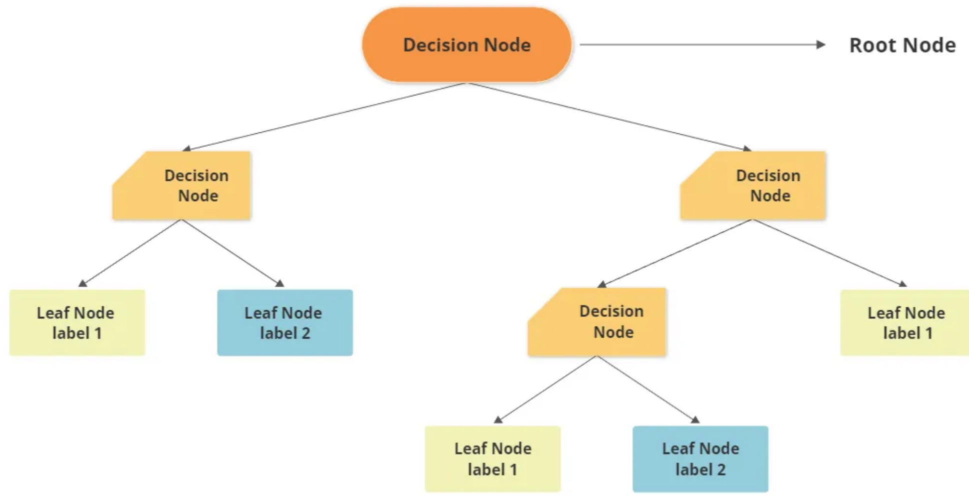

Figure 2.2: The general form of a decision tree.

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 2.3 [30]. 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 [31].

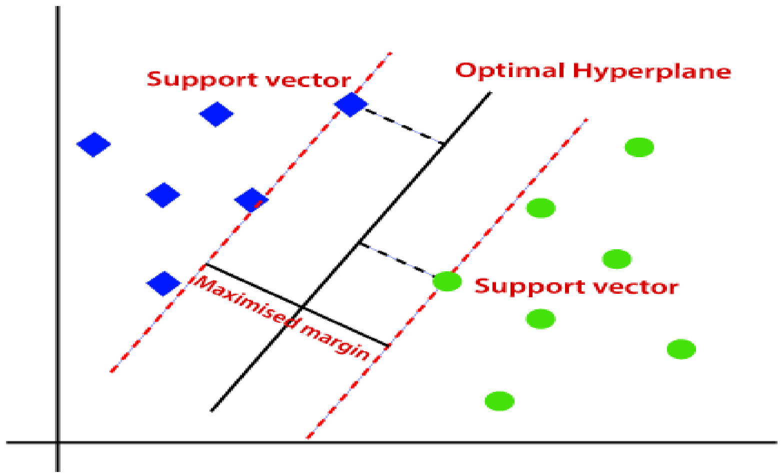

Figure 2.3: Support Vector Machine (SVM) Algorithm.

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 2.4 [32] . 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 [33].

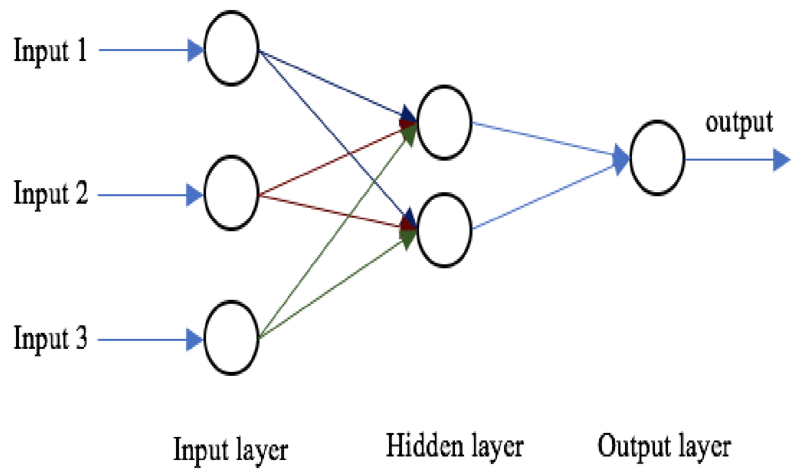

Figure 2.4: Simple neural network architecture.

2.5.4 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.

2.5.5 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(2.5) [34].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[35].

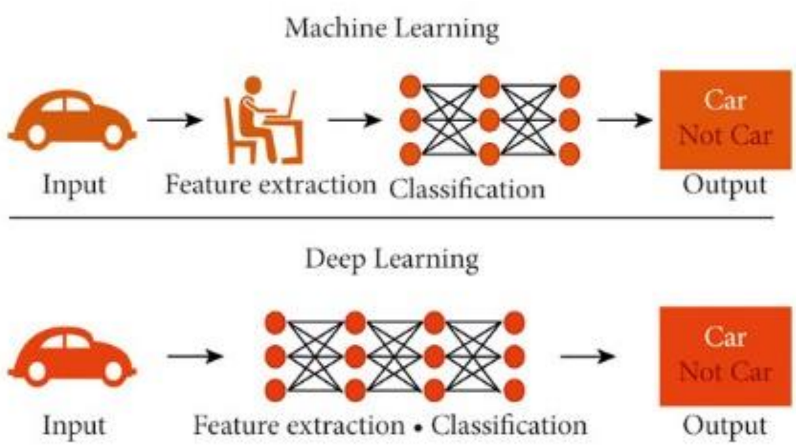

Figure 2.5: Comparison between machine learning and deep learning classification .

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):**

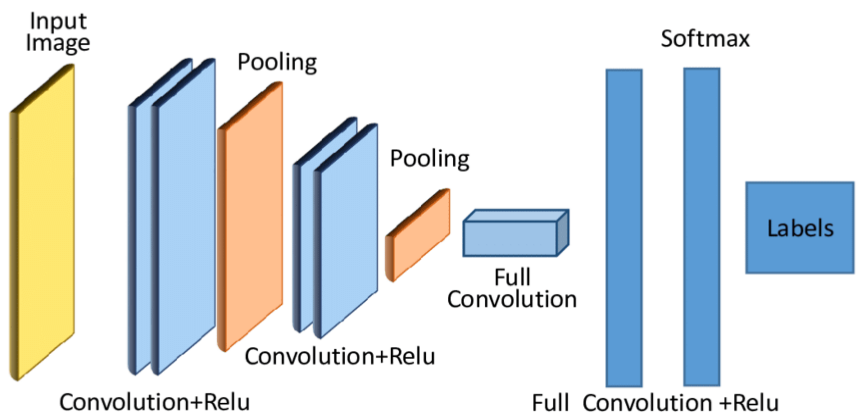

Figure 2.6: Convolutional Neural Network Architecture.

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(2.6) (e.g. "this image contains a dog") [36]. 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 [37]. CNN has a distinct operation for each layer, which is briefly summarized as follows:

(a) **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[38], As illustrated in Figure(2.7) [?].

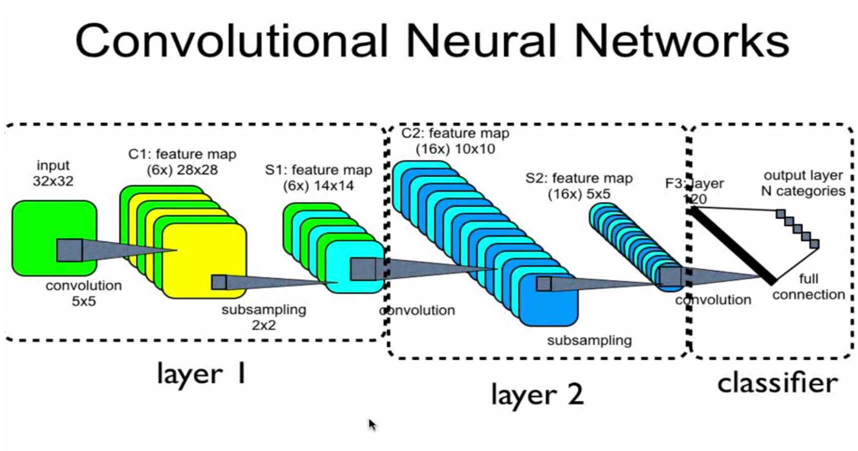

Figure 2.7: Feature Map computation by Convolution Layer.

(b) **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(2.7) [39]. This has the effect of retaining the most important information from the feature map while discarding some of the spatial information [40].

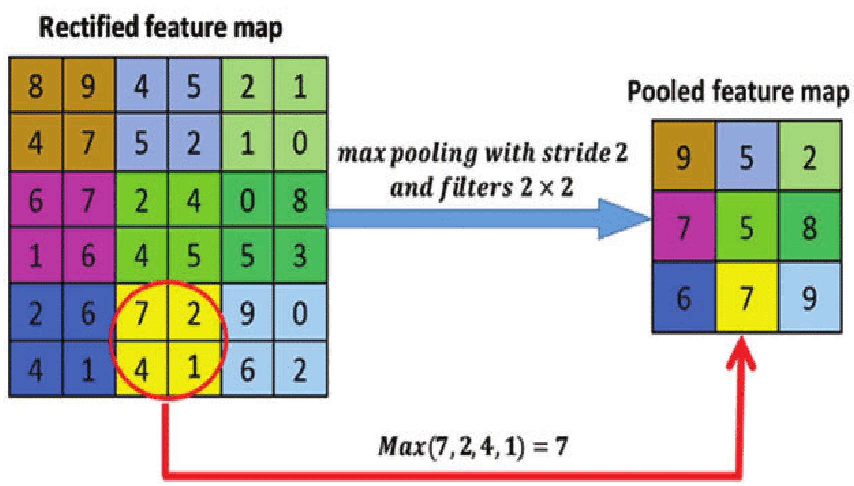

Figure 2.8: Max Pooling operation with 2x2 Filter and Stride value 2.

(c) **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(2.7) [41]. 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 [42].

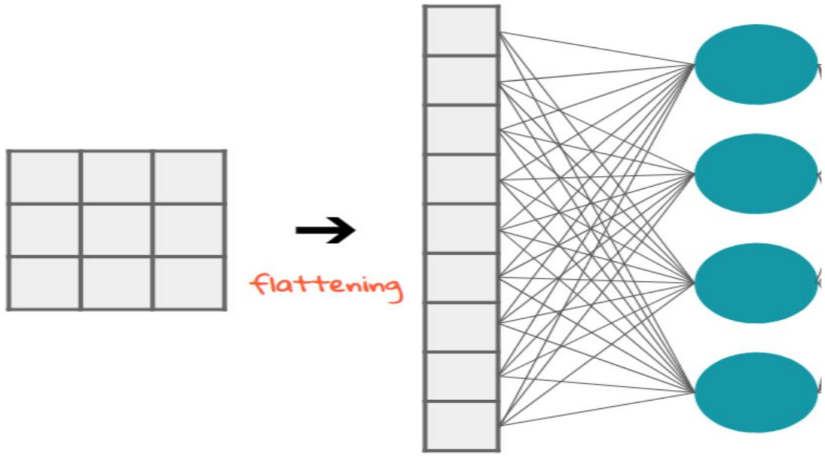

Figure 2.9: Flatten operation converting feature matrix into 1-D vector input.

- 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 [43], 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 [44].

- 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 [45]. 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.[46] 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 [47]. 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.

- 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 [48].

- 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 [49]. 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 [49]. 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 [50].

2.5.6 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.

2.6 Conclusion

It is clear from this chapter that the intersection of artificial intelligence with modern agricultural problems generates an intelligent approach capable of exploiting data sets that arise from tracking agricultural and agricultural crops and even livestock. An approach that provides a qualified learning and testing model to give accurate results for validation metrics. One of the basic techniques of artificial intelligence is for prediction and classification. It can be interfaced with supervised machine learning classifiers by applying it to a set of data that can produce a model that can be applied to similar agricultural crops to give accurate classification results or detect diseases in agricultural products. Below we present and detail the hypothesis of the general structure of this smart curriculum to give an available picture of the basic levels of the curriculum structure and how to apply it.

Chapter 3

The proposed approach and methodology

Case Study: Implementing artificial intelligence in potato cultivation in desert areas

3.1 Introduction

This chapter is followed by the chapter that includes how, methods and what is the use of artificial intelligence in the agricultural sector in the desert region. It includes a thorough and careful study of the area designated below, where there is detailed information on the contributed data and digital processing of artificial intelligence. In addition, identifying the traditional smart techniques used in studying this case. We must also present the results from the implementation of artificial intelligence, and they must be subject to accurate analysis. Furthermore, it explores and explains the implications of these findings them in the context of the research questions posed.

3.2 Description of the Study Area

We can say that agriculture is a vital activity that has allowed us to build thriving communities and feed ourselves. It has spread in all regions with fertile climate and lands that produce abundant produce of excellent quality. Despite the challenges of harsh environments such as the Sahara Desert, science must have a solution and techniques. In this project, we focus on desert agriculture in Africa, specifically in Algeria, Oued Province, where its location is shown on the map. Figure **??** It aims to use artificial intelligence to develop innovative solutions that will improve the productivity and sustainability of desert agriculture.

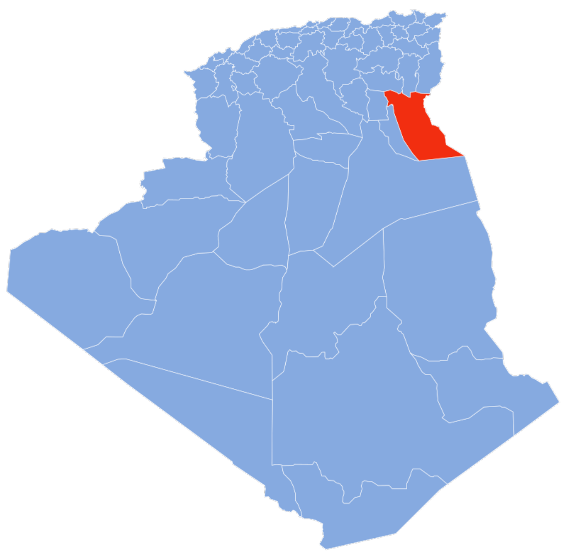

Figure 3.1: Map of Oued Souf Algeria.

3.2.1 Data Collection and Proposed method and Preprocessing

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

3.2.2 Data Collection

Collecting the required amount is the basis for the success of artificial intelligence, due to the accuracy of personal data results. High-quality data is the engine that drives AI algorithms, enabling them to identify all the looks and trends, make predictions, and drive value insights. By collecting relevant and accurately 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 trans-formative outcomes across various domains and industries [51].

- **The most important information and tools for data collection**

  1. **Place:** Five farms, depending on the location, are spread across the municipalities of Hassani-Abd-ELKrim and Debila, Oued Souf, Algeria
  2. **The date of data collection:**
     - **MARS 2024**
  3. **Imaging devices:**
     - poco X3 pro
     - Realme X50
     - Honor x9a
  4. **The number of images collected:**

Table 3.1: Data Partition Numbers.

Samples	Number of Data
Potato_healthy of	444 images
Late infection potatoes	294 images
Early infection potatoes	777 images
total	1515 images

5. **Image file size:** 38 MO.

3.2.3 Proposed method

The general structure of the proposed approach is illustrated in Figure 3.2. The initial step is to prepare a dataset to feed the deep learning algorithm as input. The plant leaf disease dataset includes different disease images and their labels as directories. Pre-processing is performed to improve the quality of extracting the needed segments or features from the input images. Different operations can be performed for images like rescaling and resizing. Pre-processing operations are important since they can affect the model's classification performance directly. Afterward, the dataset is divided into training, validation, and testing sets. The next step is the training process in which the inputs are fed to the CNN architecture and trained to evaluate a set of weights that will result in a prediction with the trained labels. After that, the performance of the model is evaluated according to various criteria such as accuracy, precision, recall, and f-score. Finally, the class label of the given test image is obtained as an output from the deep learning model which is evaluated with the achieved probability scores for each label.

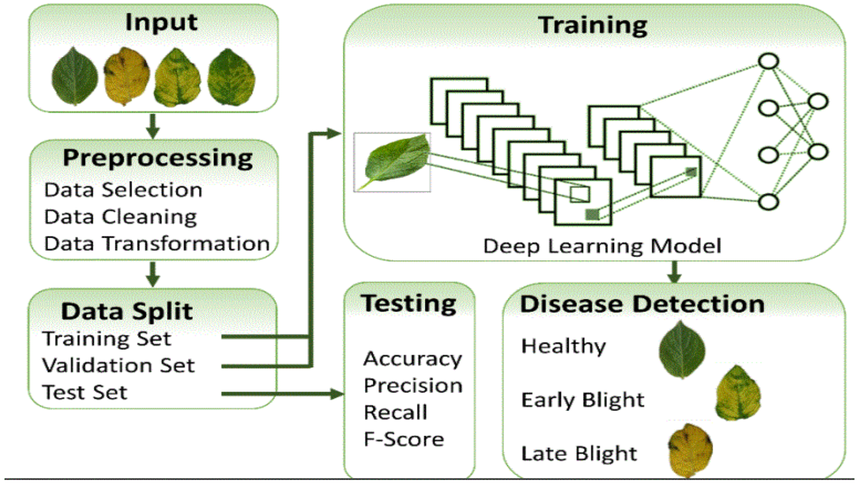

Figure 3.2: The proposed method: an overview of the general structure.

3.2.4 Pre-processing

The potato-disease dataset is divided into three classes. These classes include Late_infection_potatoes (294 images), Healthy class (444 images) and Early_infection_potatoes (777 images). Hence, the dataset was divided into 19.4objective assessment of the final model fit on the training dataset. Noise in the picture can first be reduced by removing any unnecessary parts. If there is too much disturbance in the image, it won't be used. Pictures collected from various sources and of varying sizes must be resized to 256x256 pixels for the dataset's standardized input photos. The batch size is of 32 .

1. **Data partitioning 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. within each of these files. This segmentation was crucial to ensure a balanced representation of both healthy and Late and early 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.

Table 3.2: Data Partition Numbers.

Partition	**percentage of Data**
Potatoes of train	80 %
Potatoes of test	10 %
potatoes of valid	10 %

2. **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 3.3. It is important to note that all the images utilized in our study were standardized to possess the same dimensions.

```python
Exécuter la cellule | Exécuter ci-dessous | Cellule de débogage
#%%
from PIL import Image
import glob
import cv2 as cv

i=0
path1 = glob.glob("d:/Desert_agriculture/Potatoes_healthy/*")
for img in path1:
    image = Image.open(img)
    print(f"Original size : {image.size}") # 3000x4000
    i=i+1
    image_resized = image.resize((256, 256))
    image_resized.save("d:/Desert_agriculture_2/Potatoes_healthy\image"+str(i)+".jpg")
    image = cv.imread("d:/Desert_agriculture_2/Potatoes_healthy\image"+str(i)+".jpg")
    image_rotate_90_clockwise = cv.rotate(image, cv.ROTATE_90_CLOCKWISE)
    cv.imwrite("d:/Desert_agriculture_2/Potatoes_healthy\image"+str(i)+".jpg", image_rotate_90_clockwise)
```

Figure 3.3: Code to resize and rotate images.

The first part of the figure 3.3 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 3.3, 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 3.4

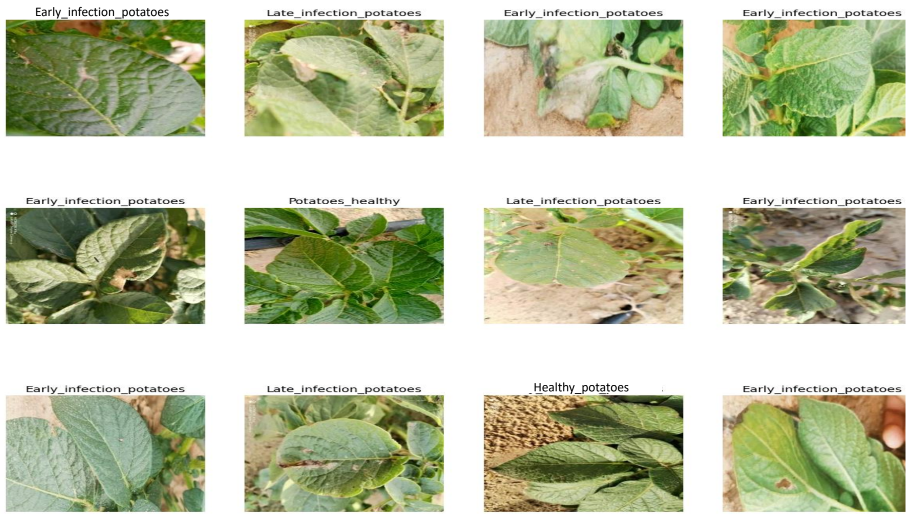

Figure 3.4: Image file path.

3.3 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 3.5 shows an example of the layers selected and used in one of the models that were used in our research.

Figure 3.6 represents the architecture of the proposed model based on the CNN structure. CNN is a type of deep neural network, mostly used to analyze images and videos. The convolution layer uses filters that perform convolution operations which is the dot product of two matrices. One of the matrices is the kernel which is a set of parameters, whereas another matrix is the inputhat is converted to an array. The pooling operation is simple in that there is a two-dimensional filter over each feature by sliding on them and creating an array with a smaller portion that contains the features extracted from the Conv2D layer. Pooling layers are useful in reducing the number of parameters to be learned and simultaneously decrease the

```
# Creating Convolution layer
input_shape = (Batch_Size, Image_Size, Image_Size, Channels)
model = models.Sequential([
    resize_and_rescale,
    data_augmentation,
    layers.Conv2D(filters = 16, kernel_size = (3,3), activation = 'relu', input_shape = input_shape),
    layers.MaxPool2D((2,2)),
    layers.Conv2D(64, (3,3), activation = 'relu'),
    layers.MaxPool2D((2,2)),
    layers.Conv2D(128, (3,3), activation = 'relu'),
    layers.MaxPool2D((2,2)),
    layers.Conv2D(64, (3,3), activation = 'relu'),
    layers.MaxPool2D((2,2)),
    layers.Conv2D(128, (3,3), activation = 'relu'),
    layers.MaxPool2D((2,2)),
    layers.Conv2D(64, (3,3), activation = 'relu'),
    layers.MaxPool2D((2,2)),
    layers.Flatten(),
    layers.Dense(128, activation = 'relu'),
    layers.Dense(64, activation = 'softmax'),

])

model.build(input_shape = input_shape)
```

Figure 3.5: the CNN model used in this research.

computational complexity of the network. The fully connected input layer, called flatten, takes the output from the last layer and performs a flattening operation to turn it into a single vector. After the flattening operation is completed, the output is fed to the neural network and applies weights to predict the class label. In the end, it gives the final probabilities for each label.

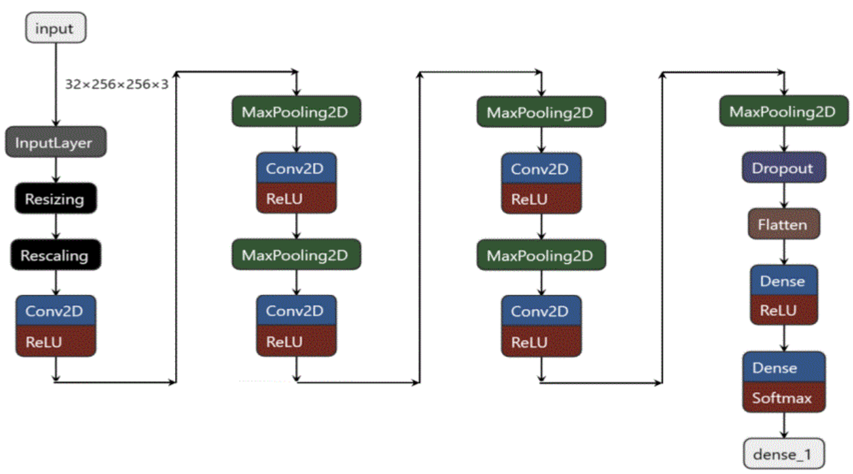

Figure 3.6: Proposed CNN architecture for potato plant leaf disease detection and classification.

1. **Data Augmentation:** In contrast to the shallow networks used in machine-learning, deep learning (Deep Network) needs a lot of data. Machine learning and deep learning frequently encounter issues with a lack of training data and an imbalance in the quantity of data for each class . Data augmentation is

the approach taken to solve this issue. Data augmentation is a technique for modifying data without changing its initial meaning. 1515 datasets are still insufficient for this research to achieve optimal performance, necessitating the use of data augmentation. The augmentation factors in this study are generated by the automatic application of simple geometric transformations like translations, rotations, scale changes, shearing, vertical and horizontal flips. But for this dataset, only the rescaling and resizing of images (1/255), horizontal flip, vertical flip and random rotation of 0.2 degrees is used as the leaves do not have high number of features to be considered.

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 3.5

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.

Table 3.3: Parameter setting.

Training details	**Parameters**	**sequentiel_2**
Input Shape	input size	(256, 192, 3)
Epoch	Number of epoch	80
Batch size	Number of batch size	57
Training compile	Optimizer	Adam
	metrics	accuracy
	loss function	SparseCategoricalCrossentropyy
Early Stopping	monitor	val_loss
	patience	15
Activation	Conv2D	relu
	Dense	sigmoid

Table 3.4: Depiction of layers and its parameters involved during training the modeL.

Layer (type)	**Output Shape**	**Param #**
sequential_1 (Sequential)	(32, 256, 256, 3)	0
sequential_2 (Sequential)	(32, 256, 256, 3)	0
conv2d_1 (Conv2D)	(32, 254, 254, 16)	448
max_pooling2d (MaxPooling2D)	(32, 127, 127, 16)	0
conv2d_1 (Conv2D)	(32, 125, 125, 64)	9.280
conv2d_2 (Conv2D)	(32, 60, 60, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(32, 30, 30, 128)	0
conv2d_3 (Conv2D)	(32, 28, 28, 64)	73,792
max_pooling2d_3 (MaxPooling2D)	(32, 14, 14, 64)	0
conv2d_4 (Conv2D)	((32, 12, 12, 128)	73,856
max_pooling2d_4 (MaxPooling2D)	(32, 6, 6, 128)	0
conv2d_5 (Conv2D)	(32, 4, 4, 64)	73,792
max_pooling2d_5 (MaxPooling2D)	(32, 2, 2, 64)	0
flatten (Flatten)	(32, 256)	0
dense_1 (Dense)	(32, 128)	32,896
dense_2 (Dense)	(32, 64)	8,256

3.3.1 Working Environment

In our working environment, we utilize a device of type mentioned in Table 4.1, along with its corresponding characteristics. Additionally, we leverage the power of Python 3 as the programming language.

3.3.2 Evaluation Metrics

After the model is created, with the usage of testing and validation datasets, performance metrics are calculated to identify the effectiveness and reliability of the proposed model. The classification metrics could be accuracy, precision, recall and Loss.

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 the proportion of correct results (either true positive or true negative) in a testing set. It is calculated using the following equation:

   $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.3.1}$$

   Where; the parameters TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives) are the metrics to calculate specificity, sensitivity, and accuracy in performance measurement. TP is used to present the number of correctly predicted positive classes. TN is the result of correctly predicted negative classes. FP is the number of cases predicted positive when they are negative. Finally, FN represents the number of cases predicted as negative when the results should be positive. Precision is the accuracy of positive predictions with the proportion of correctly predicted positive classes out of the total classes positively predicted. The calculation formula is as follows:

   $$Accuracy = \frac{TP}{TP + FP} \tag{3.3.2}$$

2. **Recall:**

   Recall (sensitivity) represents correctly predicted positive classes divided by the total actual positive classes. It is calculated using the following equation: [**?**].The formula for recall is:

$$Recall = \frac{TP}{TP + FN} \tag{3.3.3}$$

3. **Precision:**

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

$$Precision = \frac{TP}{TP + FP} \tag{3.3.4}$$

   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 [**?**]. The formula for Binary Cross-Entropy Loss is:

$$\text{Binary Cross-Entropy Loss} = -(y \cdot \log(y_{hat}) + (1 - y) \cdot \log(1 - y_{hat})) \tag{3.3.5}$$

   Where:

   y : True label (0 or 1)

   y_{hat} : Predicted probability of the positive class (between 0 and 1)

3.4 Conclusion

In this chapter, the general structure of the memorandum is detailed. The details focused on the basic levels:

1. The level of data collection from potato fields in desert farms.

2. The data set is divided and classified according to the proposed treatment category (healthy, infected, and diseased).

Finally, collecting the results of the proposed measures, as the study showed the possibility of improving the results of the measures during the training and testing process several times. Indicate the percentage of improvement in the metrics results, discuss and detail the approach applied.

Chapter 4

Results and discussion

4.1 Environment and Development Tools

The table below shows the characteristics of the work environment

Table 4.1: The Work Environement .

Name	Parameter
Memory	4 GB
Processor	Intel(R) Pentium(R) CPU J3710 @ 1.60GHz 1.60 GHz
Server model	acer
OS	Windows 10 Professionnel
Language	Python 3.11.0
Code editor	Visual Studio Code

4.1.1 Kaggle

Kaggle is an interactive web platform that offers machine learning competitions in data science. The platform provides free datasets, notebooks and tutorials that data scientists need to complete their machine learning projects 4.1

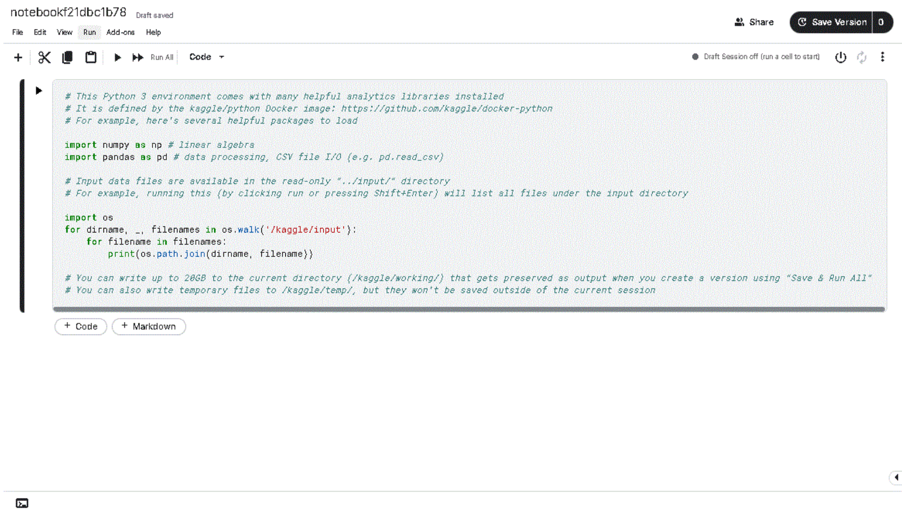

Figure 4.1: Interface Work Of Kaggle .

Some of the most popular languages for data science are Python,R and Julia. According to Kaggle State of Data Science Survey 2021,Python-based tools continue to dominate the machine learning frameworks giving the language a monopoly in this industry

4.1.2 Python

It is an interpreted, multi-paradigm, multi-format programming language. Prefer essential programming with structure, function, and direction. There is a durable dynamic type, automatic memory management over wide ranges, and an exception management system; It is also similar to Perl, Ruby, Scheme, Smalltalk, and Tcl. In the figure below 4.2 we show some types of codes used in this work.

```python
# import libraries
import tensorflow as tf
from tensorflow.keras import models, layers
import matplotlib.pyplot as plt
import numpy as np
import pathlib
import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '2' # to disable all debugging logs
```

```python
# importing dataset
dir = os.listdir('../input/agriculture-el-oued/Agriculture-2')
for filenames in dir:
    print(filenames)
```

Figure 4.2: Many python Code .

4.1.3 Datasets

A dataset is a structured collection of data, organized and stored for analysis or processing purposes.For example, in our work, we accessed many farms in order to collect the necessary number of data to reach at least high-quality results.4.3

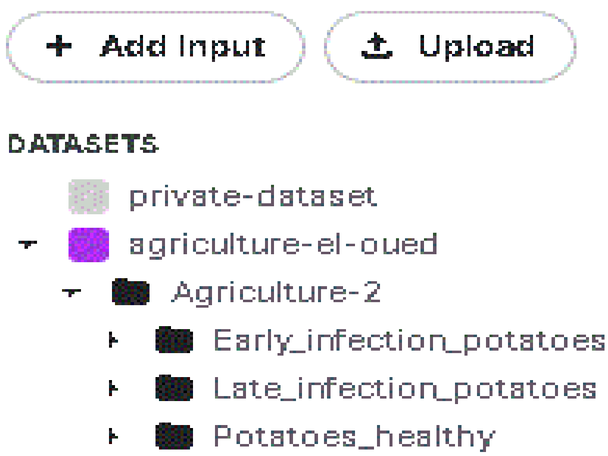

Figure 4.3: Datasets Folder In Kaggle .

4.2 Results

4.2.1 Training Process

The CNN model was finally created, and then the Adam optimizer was used to build the model. It is a very effective optimization method used to train deep neural networks. It incorporates the benefits of two optimization techniques, RMSProp (root mean square propagation) and AdaGrad (adaptive gradient algorithm). A five-layer convolutional neural network architecture model was used in the dataset learning exercise. The study period specifications called for 80 and 41 epochs, a batch size of 32, and a learning rate of 0.01 to enhance model performance. An algorithm is proposed that searches for values in an image dataset as part of the learning process in order to recognize new images.

To calculate the loss and accuracy values, the results of the epochs are recorded. The accuracy value indicates how accurately the system classifies objects, as the loss value must be close to zero. Loss is an indication that the model has a bad value. The results of the accuracy and loss values are shown in the images below. Where values close to one and others close to zero were obtained. During training, early and later epochs show more volatility values. After the 80 qnd 40 epoch, the accuracy of the training dataset, which includes the original images, reached very close to 100 percent **??**

```
Epoch1/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 156s2s/step - accuracy: 0.3981 - loss: 1.7752 - val_acc
uracy: 0.4688 - val_loss: 1.0552
Epoch2/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 154s2s/step - accuracy: 0.5107 - loss: 1.0390 - val_acc
uracy: 0.4965 - val_loss: 1.0398
Epoch3/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 151s2s/step - accuracy: 0.5387 - loss: 1.0033 - val_acc
uracy: 0.5827 - val_loss: 0.8762
Epoch4/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 203s2s/step - accuracy: 0.6237 - loss: 0.8751 - val_acc
uracy: 0.6389 - val_loss: 0.8603
Epoch5/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 201s2s/step - accuracy: 0.6284 - loss: 0.8812 - val_acc
uracy: 0.7230 - val_loss: 0.7386
Epoch6/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 223s2s/step - accuracy: 0.6421 - loss: 0.8369 - val_acc
uracy: 0.5972 - val_loss: 0.9006
Epoch7/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 156s2s/step - accuracy: 0.6269 - loss: 0.8417 - val_acc
uracy: 0.6007 - val_loss: 0.8161
Epoch8/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 199s2s/step - accuracy: 0.6573 - loss: 0.7965 - val_acc
uracy: 0.6771 - val_loss: 0.7516
Epoch9/41
76/76 ━━━━━━━━━━━━━━━━━━━━ 228s2s/step - accuracy: 0.6704 - loss: 0.7686 - val_acc
uracy: 0.6295 - val_loss: 0.7971
Epoch10/41
```

Figure 4.4: Exemple Of Accuracy and Loss recorded while Training the Dataset epoche 41 .

4.2.2 Testing Process

Following the training procedure using the generated dataset, we retrieve the remaining part of the dataset i.e., testing dataset to be tested with the model. The testing accuracy is recorded as 80% and testing loss as 20%. Refer, to the fig. 6 for results. Also, the model trained and tested is capable of signifying the confidence up to what extent it is able to classify or predict the disease. When assessing the model's accuracy with a leaf image object, it is crucial to keep in mind that image noise will result in subpar classification results. By cutting out everything, but the desired leaf object and ensuring that there is only one leaf in each frame, noise can be removed from a picture.

4.2.3 Analyze the result of the model

The focus of this title is to present the results of the training conducted on this model, by implementing it over two epochs. To know the differences, the accuracy and loss of this model are accompanied by detailed and comprehensive notes Analyze the results obtained. Graphic curves are used to display ally representation of models performance,

Next, each graph will be analyzed individually

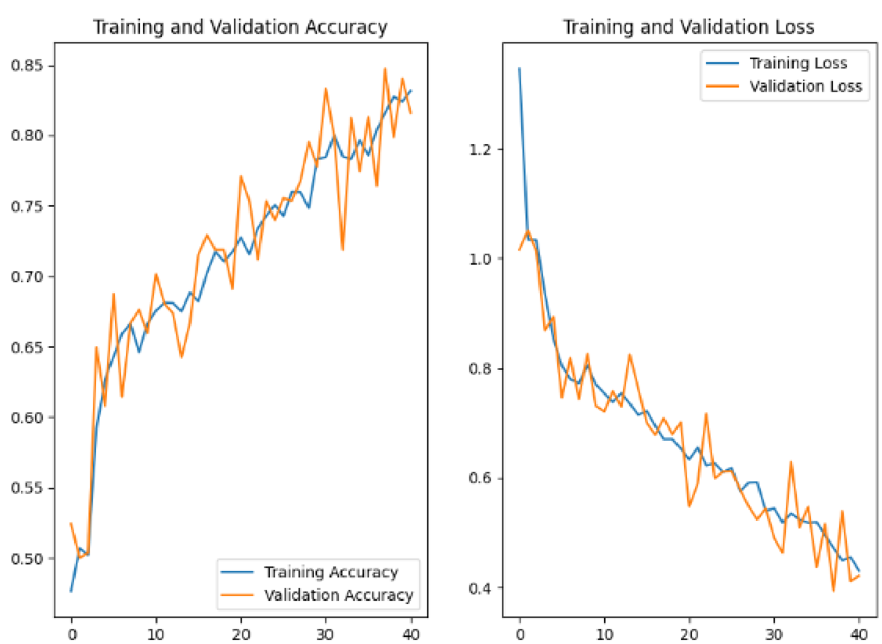

Figure 4.5: Fit the model (Accuracy and Loos extraction.

Based on the results in Figure 3.6, the Training Results of CNN model showed satisfactory performance compared to Figure 4.5. The following are the most important observations and conclusions:

- Throughout the training period, the accuracy, precision, consistently exhibited a well-balanced predictive ability.
- these metrics gradually improved and reached a peak of approximately 80%, indicating the model's effective learning and enhanced classification capabilities. The gradual decrease in loss further confirmed the model's proximity to the true values.
- The evaluation on unseen data, represented by validation accuracy, loss, and accuracy provided valuable insights. The validation accuracy, precision, and recall achieved a high value, similar to the training accuracy. The validation loss decreased, demonstrating the model's generalization to unseen data without overfitting the training set.

4.3 Discussion and Interpretation

This research focused and scrutinized the application of deep learning technology in the context of desert agriculture. The success in its use was observed through promising results, as we confirmed and verified again that there is no obstacle to the practice of agriculture, as different environments differ according to their climate. However, we must admit that there are obstacles and many difficulties and challenges to obtaining a small and modest amount of data.

Chapter 5

Conclusion and Future Work

5.1 CONCLUSION

In this research, we suggested a deep learning strategy for categorising potato leaf diseases using a CNN (Convolutional Neural Network). Pre-processing the leaf images, training a CNN model on the pre-processed data, and assessing the model's performance on a test set are all steps in the suggested approach. With an overall accuracy of 99.18accurate in classifying two different potato leaf diseases, including Early Blight and Late Blight.The suggested method may offer a trustworthy and effective means of diagnosing potato disease, which is essential for ensuring food security and minimising financial losses in agriculture. This approach can be further extended to other plant species, enabling early detection of diseases, and facilitating prompt interventions to prevent further spread. Overall, the findings of this research show the efficacy of deep learning-based methods for identifying plant diseases and their potential to revolutionise the agricultural technology industry.The proposed approach can assist farmers and experts in the early detection of diseases and timely implementation of control measures, thereby improving crop yields, reducing losses, and contributing to global food security.Furthermore, the proposed deep learning-based approach offers several advantages over traditional methods of plant disease diagnosis. Unlike manual inspection, which is timeconsuming and prone to human error, deep learning-based approaches can provide accurate and consistent results with minimal human intervention. Additionally, this approach can process large amounts of data in a short amount of time, making it ideal for real-time monitoring and detection of plant diseases.While the proposed approach achieved high accuracy in identifying potato leaf diseases, further research is necessary to investigate its performance on other plant species and under different environmental conditions. Additionally, the proposed method can be further improved by incorporating additional data sources and developing more efficient models on these datasets. In conclusion, this study demonstrates the potential of deep learning-based approaches for plant disease diagnosis and management.

5.2 FUTURE WORK

Early detection of potato leaf diseases is made possible by the suggested technique, which can have a significant impact on crop protection and global food security. With further research and development, this approach can be extended to other crops and environmental conditions.

Bibliography

[1] Fernando P Carvalho. Agriculture, pesticides, food security and food safety. *Environmental science & policy*, 9(7-8):685–692, 2006.

[2] Djamel Saba, Youcef Sahli, Rachid Maouedj, Abdelkader Hadidi, and Miloud Ben Medjahed. Towards artificial intelligence: concepts, applications, and innovations. *Enabling AI Applications in Data Science*, pages 103–146, 2021.

[3] Ishita Bhakta, Santanu Phadikar, and Koushik Majumder. State-of-the-art technologies in precision agriculture: a systematic review. *Journal of the Science of Food and Agriculture*, 99(11):4878–4888, 2019.

[4] Xinjie Deng and Zhonggen Yu. A systematic review of machine-translation-assisted language learning for sustainable education. *Sustainability*, 14(13):7598, 2022.

[5] Ashwani K Gupta and Satish K Arora. *Industrial automation and robotics.* Laxmi publications, 2011.

[6] Alice Segato, Aldo Marzullo, Francesco Calimeri, and Elena De Momi. Artificial intelligence for brain diseases: A systematic review. *APL bioengineering*, 4(4), 2020.

[7] Geogios N Yannakakis. Game ai revisited. In *Proceedings of the 9th conference on Computing Frontiers*, pages 285–292, 2012.

[8] Zeki Murat Çınar, Abubakar Abdussalam Nuhu, Qasim Zeeshan, Orhan Korhan, Mohammed Asmael, and Babak Safaei. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability*, 12(19):8211, 2020.

[9] Tom Duckett, Simon Pearson, Simon Blackmore, Bruce Grieve, Wen-Hua Chen, Grzegorz Cielniak, Jason Cleaversmith, Jian Dai, Steve Davis, Charles Fox, et al. Agricultural robotics: the future of robotic agriculture. *arXiv preprint arXiv:1806.06762*, 2018.

[10] Steven A Fennimore, David C Slaughter, Mark C Siemens, Ramon G Leon, and Mazin N Saber. Technology for automation of weed control in specialty crops. *Weed Technology*, 30(4):823–837, 2016.

[11] SK Roy, SK Mondal, PP Pal, KS Das, FH Rahman, and A Haldar. Icar-atari kolkata annual report 2021. 2022.

[12] David Tilman. *Plant strategies and the dynamics and structure of plant communities.* Princeton university press, 1988.

[13] A Bationo, SP Wani, CL Bielders, PLG Vlek, and AU Mokwunye. Crop residue and fertilizer management to improve soil organic carbon content, soil quality and productivity in the desert margins of west africa. In *Global climate change and tropical ecosystems*, pages 117–145. CRC Press, 2019.

[14] DR Sikka. Desert climate and its dynamics. *Current Science*, pages 35–46, 1997.

[15] Johan Rockström and Jennie Barron. Water productivity in rainfed systems: overview of challenges and analysis of opportunities in water scarcity prone savannahs. *Irrigation Science*, 25:299–311, 2007.

[16] Muhammad Iftikhar Hussain, Muhammad Farooq, Adele Muscolo, and Abdul Rehman. Crop diversification and saline water irrigation as potential strate-

gies to save freshwater resources and reclamation of marginal soils—a review. *Environmental Science and Pollution Research*, 27(23):28695–28729, 2020.

[17] Claudia Ringler, Dawit Kelemework Mekonnen, Hua Xie, Agbonlahor Mure Uhunamure, et al. Irrigation to transform agriculture and food systems in africa south of the sahara. *2020 Annual trends and outlook report: Sustaining Africa's agrifood system transformation: The role of public policies*, pages 57–70, 2020.

[18] Ugur Korkut Pata. Linking renewable energy, globalization, agriculture, co2 emissions and ecological footprint in bric countries: A sustainability perspective. *Renewable Energy*, 173:197–208, 2021.

[19] Loutfy I El-Juhany et al. Degradation of date palm trees and date production in arab countries: causes and potential rehabilitation. *Australian Journal of Basic and Applied Sciences*, 4(8):3998–4010, 2010.

[20] Anandhavalli Muniasamy. Machine learning for smart farming: a focus on desert agriculture. In *2020 International Conference on Computing and Information Technology (ICCIT-1441)*, pages 1–5. IEEE, 2020.

[21] W Mupangwa, L Chipindu, I Nyagumbo, S Mkuhlani, and G Sisito. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in eastern and southern africa. 2:1–14, 2020.

[22] Foyez Ahmed Prodhan, Jiahua Zhang, Fengmei Yao, Lamei Shi, Til Prasad Pangali Sharma, Da Zhang, Dan Cao, Minxuan Zheng, Naveed Ahmed, and Hasiba Pervin Mohana. Deep learning for monitoring agricultural drought in south asia using remote sensing data. 13(9):1715, 2021.

[23] Rubby Aworka, Lontsi Saadio Cedric, Wilfried Yves Hamilton Adoni, Jérémie Thouakesseh Zoueu, Franck Kalala Mutombo, Charles Lebon Mberi Kimpolo, Tarik Nahhal, and Moez Krichen. Agricultural decision system based on advanced machine learning models for yield prediction: case of east african countries. 2:100048, 2022.

[24] Iqbal H Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3):160, 2021.

[25] Mohamed Alloghani, Dhiya Al-Jumeily, Jamila Mustafina, Abir Hussain, and Ahmed J Aljaaf. A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised and unsupervised learning for data science*, pages 3–21, 2020.

[26] Pádraig Cunningham, Matthieu Cord, and Sarah Jane Delany. Supervised learning. In *Machine learning techniques for multimedia: case studies on organization and retrieval*, pages 21–49. Springer.

[27] Zhonghe Ren, Fengzhou Fang, Ning Yan, and You Wu. State of the art in defect detection based on machine vision. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 9(2):661–691, 2022.

[28] Chao Yu, Jiming Liu, Shamim Nemati, and Guosheng Yin. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)*, 55(1):1–36, 2021.

[29] Hubert Szczepaniuk and Edyta Karolina Szczepaniuk. Applications of artificial intelligence algorithms in the energy sector. *Energies*, 16(1):347, 2022.

[30] Derek A Pisner and David M Schnyer. Support vector machine. In *Machine learning*, pages 101–121. Elsevier, 2020.

[31] Himani Bhavsar and Mahesh H Panchal. A review on support vector machine for data classification. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 1(10):185–189, 2012.

[32] Akash Goel, Amit Kumar Goel, and Adesh Kumar. The role of artificial neural network and machine learning in utilizing spatial information. *Spatial Information Research*, 31(3):275–285, 2023.

[33] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016.

[34] Devarshi Shah, Jin Wang, and Q Peter He. Feature engineering in big data analytics for iot-enabled smart manufacturing–comparison between deep learning and statistical learning. *Computers & Chemical Engineering*, 141:106970, 2020.

[35] Xue-Wen Chen and Xiaotong Lin. Big data deep learning: challenges and perspectives. *IEEE access*, 2:514–525, 2014.

[36] Lingqiao Liu, Chunhua Shen, and Anton Van den Hengel. The treasure beneath convolutional layers: Cross-convolutional-layer pooling for image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4749–4757, 2015.

[37] Li Deng. Deep learning: from speech recognition to language and multimodal processing. *APSIPA Transactions on Signal and Information Processing*, 5:e1, 2016.

[38] Maram G Alaslani. Convolutional neural network based feature extraction for iris recognition. *International Journal of Computer Science & Information Technology (IJCSIT) Vol*, 10, 2018.

[39] Afia Zafar, Muhammad Aamir, Nazri Mohd Nawi, Ali Arshad, Saman Riaz, Abdulrahman Alruban, Ashit Kumar Dutta, and Sultan Almotairi. A comparison of pooling methods for convolutional neural networks. *Applied Sciences*, 12(17):8643, 2022.

[40] Nadeem Akhtar and U Ragavendran. Interpretation of intelligence in cnn-pooling processes: a methodological survey. *Neural computing and applications*, 32(3):879–898, 2020.

[41] Mohaiminul Islam, Guorong Chen, and Shangzhu Jin. An overview of neural network. *American Journal of Neural Networks and Applications*, 5(1):7–11, 2019.

[42] Qiuyu Zhu, Pengju Zhang, Zhengyong Wang, and Xin Ye. A new loss function for cnn classifier based on predefined evenly-distributed class centroids. *IEEE Access*, 8:10888–10895, 2019.

[43] Kuldeep Singh Kaswan, Jagjit Singh Dhatterwal, and B Balamurugan. *Python for Beginners*. Chapman and Hall/CRC, 2023.

[44] Sikender Mohsienuddin Mohammad. Artificial intelligence in information technology. *Available at SSRN 3625444*, 2020.

[45] Bo Pang, Erik Nijkamp, and Ying Nian Wu. Deep learning with tensorflow: A review. *Journal of Educational and Behavioral Statistics*, 45(2):227–248, 2020.

[46] Bastiaan Sjardin, Luca Massaron, and Alberto Boschetti. *Large scale machine learning with Python*. Packt Publishing Ltd, 2016.

[47] Moustafa Alzantot, Yingnan Wang, Zhengshuang Ren, and Mani B Srivastava. Rstensorflow: Gpu enabled tensorflow for deep learning on commodity android devices. In *Proceedings of the 1st International Workshop on Deep Learning for Mobile Systems and Applications*, pages 7–12, 2017.

[48] Ivan Vasilev, Daniel Slater, Gianmario Spacagna, Peter Roelants, and Valentino Zocca. *Python Deep Learning: Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow*. Packt Publishing Ltd, 2019.

[49] M Mazhar Rathore, Syed Attique Shah, Dhirendra Shukla, Elmahdi Bentafat, and Spiridon Bakiras. The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. *IEEE Access*, 9:32030–32052, 2021.

[50] Jeremy Howard and Sylvain Gugger. Fastai: a layered api for deep learning. *Information*, 11(2):108, 2020.

[51] David Sjödin, Vinit Parida, Maximilian Palmié, and Joakim Wincent. How ai capabilities enable business model innovation: Scaling ai through co-evolutionary processes and feedback loops. *Journal of Business Research*, 134:574–587, 2021.