



Republic of Tunisia
Higher Education and Scientific
Research Ministry
University of Carthage



NATIONAL SCHOOL OF SCIENCES AND ADVANCED TECHNOLOGIES OF BORJ CEDRIA

END-OF-YEAR
PROJECT REPORT

Application of Machine Learning for Plant diagnostician

Prepared by:

INTISSAR ISSAOUI 1TA2
ABIR YAHIA 1TA2

Defended on **28/05/2022** before the jury:

Mr. WALID NSIBI Examiner - ENSTAB
Mr. SALEM SAIDI Academic Supervisor - ENSTAB

Academic Year: 2022-2023

Réf: PSI2023_17

Dedications

*I would like to dedicate this report to our **parents** for their unwavering support through our academic persistence, for being our constant resource of inspiration, ethical values, perseverance and determination, for always supporting us morally and financially so we could reach this academic milestone. We would also like to dedicate this report to our supervisor **Mr. SALEM SAIDI** who have always been there for us, guiding and mentoring us with his expertise and knowledge. In addition, we dedicate this work to our **friends** for their support and faith in us, and for always being there to lift our spirits when times got tough.*

**INTISSAR ISSAOUI
ABIR YAHIA**

Acknowledgements

In this section, we would like to express our sincere gratitude to our families, our supervisor **Mr. Salem Saidi**, and everyone who contributed to the assistance and support provided to us. Your guidance has been invaluable, and we cannot thank you enough. We extend our utmost gratitude for the immense dedication and hard work you have invested in enlightening our path, enabling us to attain this esteemed achievement.

You have not only imparted valuable knowledge but also inspired confidence in our abilities through your guidance. The depth of gratitude we feel is beyond measure, as your generosity has positively impacted both our personal growth and professional success.

We would also like to thank **Mr. Walid Nsibi**, Assistant Professor at the National School of Advanced Sciences and Technologies of Borj Cedria (ENSTAB), for accepting to honor us by serving as a member of the jury for this synthesis project.

Contents

1 Computer vision and Problem statement	3
1 Computer vision.....	3
1.1 Object detection	3
1.2 Image segmentation.....	4
1.3 Image classification	4
1.4 Techniques of image classification	5
1.5 Application of image classification	7
2 Problem statement	9
2 Theoretic Background	10
1 Basic Concepts of Artificial Neural Networks	10
1.1 Structure of an Artificial Neural Networks	10
1.2 Perceptron	11
1.3 Activation functions	11
1.4 Loss Function	13
1.5 Type of Neural Nets Learning	14
2 Different types of Neural Networks in DeepLearning.....	15
2.1 Recurrent Neural Networks RNN	15
2.2 Graph Neural Networks GNN	15
2.3 Convolutional neural network CNN.....	16
3 Predefined models.....	17
3.1 AlexNet	17
3.2 MobileNet	17
3.3. VGG	18
3 Description and Implementation of the Plants Diagnostician	19
1 The chosen Model: VGG16.....	19
2 Software (Technology) stack	20
3 Implementation and experience results	21

3.1	Dataset	21
3.2	Data preparation	22
3.3	Experimental results	23
	Bibliography	26
	Netography	27

List of Figures

Figure 1.1 Object detection	4
Figure 1.2 Object segmentation	4
Figure 1.3 Image classification	5
Figure 1.4 Result of skin cancer classification	7
Figure 1.5 Classification of the printed tablets	8
Figure 1.6 Trafic monitoring classification	8
Figure 1.7 Facial recognition	9
Figure 2.1 Layers of an Artificial Neuron Network	11
Figure 2.2 Perceptron	11
Figure 2.3 Activation functions	13
Figure 2.4 RNN structure	15
Figure 2.5 Graph Neural Network structure	16
Figure 2.6 Typical architecture of CNN	16
Figure 2.7 MobileNet neural network structure	17
Figure 3.1 VGG16 model	20
Figure 3.2 Software stack	21
Figure 3.3 Trainig and Validation Accuracy	23
Figure 3.4 Trainig and Validation loss	24

General Introduction

Artificial intelligence is the concept that brought to humankind the opportunity to reach a much more comfortable life since it simulates the human's intelligence.

Actually Artificial Intelligence is developed for machines in order to solve complex problems in record time as long as we are living in the era of speed, here we highlight on its new capability of reasoning and making strategic and optimum plans to reach precised goals in different cases, so allow us to point out its overwhelming potential in several operations in object recognition whatever it was images or even voices, prediction, etc.. Mainly Artificial intelligence matches the human skills in a way to stand by their issues and give them the opportunity to decide prioritizing some tasks over some others and this depends on the situation. This intelligence is constructed by generalized learning, reasoning and problems solving besides it stands on building complex algorithms and mathematical functions.

Historically, the concept of Artificial intelligence was inspired from the biological human brain, in fact the first artificial neuron were invented in 1943 by two mathematicians, and neural scientists Warren McCullough and Walter Pitts, as in a scientific article entitled "A LOGIC CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY" McCullough and Pitts explained that "Neural events and the relation among them can be treated by means of propositional logic". That's to say artificial neural network was inspired from the biological neural network.

Machine learning, as a subpart of Artificial intelligence, is a science concerned with imparting human-like intelligence onto machines and creating machines able to sense, to reason, to act and to adapt. It uses algorithms to learn how to deal with tasks as long as it is fed with data as input. In real life we already use Machine Learning in email spam recognition, spell check music and video recommendation, etc.

Machine Learning has indulged radical changes in the world's technologies, so it would be profitable to let it in our homes as some advanced techniques to facilitates some specific tasks to let us gain time for performing some much more important responsibilities. Indeed, this is the reason beyond choosing machine learning in building an application dedicated to help people looking after their domestic plants and farmers to supervise their crops for an exceptional production, and this is guaranteed in order to reduce the expenses of bringing up some expert to deal with the main issues.

Technically, we will present our project "Application of Machine Learning for Plant Diagnostician" that would help us handling this situation by offering advice thanks to the fact that is related to an encyclopedic data base that involves a huge bench of images and information about plant pathology. Indeed in case of diseases it will afford us with the necessary amount of data about these diseases and it will offer us a prescription of required medicines and remedies.

For our rapport, it will be devised into three chapters. In the first chapter "Computer Vision and Problem statement" we will be going in depth of explaining computer vision and subtract our solution. On the second chapter, we will give a theoretic background about Deep learning and its various application . Finally, in the last chapter we will present the architecture of our model working with in our algorithm, and we will show the experience result of the algorithm.

Chapter1 : Computer vision and problem statement

Introduction

Computer vision refers to the interdisciplinary field of study that deals with enabling computers and machines to interpret visual information from images, videos or other sources. In this chapter we will present essential concepts of computer vision, the problem we are trying to solve and the proposed solution.

1 Computer vision

Computer vision [1] is a method utilized for analyzing and interpreting images or videos, it combines techniques from various branches of science and engineering, such as mathematics, physics and signal processing. It involves the use of mathematical algorithms and machine learning models to extract meaningful information from visual data. Computer vision is a rapidly growing and innovative field that has the potential to revolutionize many industries, enhance human capabilities, and improve daily life. It has contributed to advancements in fields like facial recognition, object detection and tracking, image restoration and enhancement, and augmented reality. In addition, computer vision is also used in industries such as healthcare for medical image analysis and in retail for product recognition and tracking.

1.1 Object detection

Object detection is an essential task in numerous applications, from self-driving cars to security systems.

The purposes of object detection is to accurately determine which objects are present in a given image or video frame, and where they are located like the figure 1.1 shows below.



Figure 1.1: Object detection

1.2 Image segmentation

Image segmentation is a technique used in computer vision to divide an image into multiple segments, each of them represents separate objects or region within the image. This technique has various applications in fields such as medical imaging, robotics, and autonomous vehicles. To highlight its use in the medical field let us represent as example the MRI (magnetic resonance imaging) scans as shown in picture below (See figure 1.2) in a picture of a brain scan, segmentation is used to identify and differentiate tissue types.

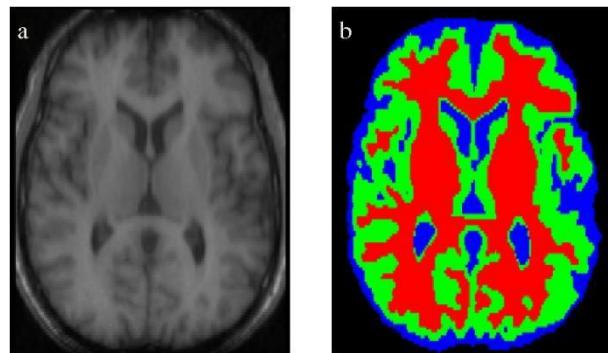


Figure 1.2: Object segmentation

1.3 Image classification

Image classification [2] is a major topic in machine learning and computer vision since it has numerous practical applications in healthcare including disease diagnosis and medical imaging processing research based on machine learning algorithms such as SVM, K-NN, K-Means, LSTM, DCNN, etc.

It works by using machine learning and deep learning algorithms to analyze images and identify specific features or patterns that can be used to classify the image into one of several predetermined categories. We mention here, three types of classification (see Figure 1.3):

- **Binary classification:** Binary classification aims to categorize the input images into one of two possible classes. This approach is often used for tasks such as detecting the presence or absence of a specific object, identifying whether an image contains something relevant to a specific topic, or distinguishing between two similar images with subtle differences.
- **Multi-class classification:** It is a type of classification problem where the goal is to classify instances into one of several possible classes. In contrast to binary classification, where the goal is to classify instances into one of two classes, multi-class classification involves predicting one of several possible labels.
- **Multi-label classification:** Multi-label classification refers to the task of predicting multiple labels or categories for a given instance, where each label represents a different class or category. In contrast to traditional classification problems where an instance belongs to a single class, multi-label classification allows for an instance to belong to more than one class.

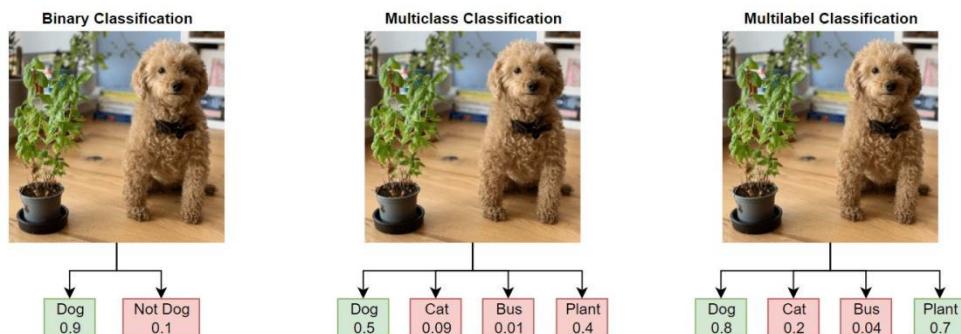


Figure 1.3: Image classification

1.4 Techniques of image classification

Image classification is a computer vision technique used to categorize images into one of several predefined classes or labels. There are many machine learning algorithms and deep learning algorithms that do the image classification.

1.4.1 Classification with classical method of machine learning

There are many method of classification with classical method of machine learning :

- **Naive Bayes:** It is a widely used as classification algorithm in the field of ML that applies Bayes' theorem. It assumes that all features are independent, and thus it can efficiently predict outcomes with great accuracy even when dealing with large data-sets. It has been successfully applied to various tasks such as text categorization, spam filtering, sentiment analysis, medical diagnosis and more.
- **Regression Trees:** Regression Trees are a ML technique used primarily for solving regression problems. A regression problem involves predicting continuous numeric values, such as stock prices or medical test results, from input features. Regression trees are constructed using a series of binary decisions on the input features that recursively split the data into smaller and smaller subsets until reaching leaf nodes containing predicted values.
- **SVM:** Support Vector Machines (SVMs) are a popular ML algorithm that is can be used in various applications, such as classification, regression, and outlier detection. The key concept behind SVMs is to find the hyperplane that maximizes the margin between two classes of data points. This means that SVMs aim to separate different classes of data with an optimal decision boundary by maximizing the distance from each class's closest point to the separating line.

1.4.2 Classification using deep neural network

Deep learning also have many methods of classification let us mention :

- **DC-NN:** Deep Convolutional Neural Networks (DCNNs) have revolutionized the field of image classification, achieving unprecedeted levels of accuracy and speed. DCNN's have the ability to learn hierarchical representations of features, enabling the network to recognize complex patterns within images. These features are extracted from images using convolutional layers, resulting in abundant and complex structures that can effectively classify objects, detect or segment them.
- **VGG:** The VGG architecture comprises of a complex and sophisticated convolutional neural network that is designed to operate with extreme depth, ensuring exceptional accuracy in image recognition tasks. This DL model consists of several layers, each performing intricate calculations on the input data, leading to progressive feature extraction and representation.

- **ResNet:** Short for Residual Network, is a deep learning architecture that revolutionized image recognition tasks by enabling the training of extremely deep neural networks. The main innovation behind ResNet is its use of skip connections or shortcuts between layers that allow signals from earlier parts of the network to bypass several layers and reach deeper into the network without attenuation. This unique feature significantly improved gradient flow within the model and allowed it to overcome performance degradation issues commonly associated with very deep neural networks.

1.5 Application of image classification

1.5.1 Medical

Medical image classification is the process of using machine learning and computer vision techniques to categorize images generated from medical imaging modalities such as X-rays, CTscans, MRIs and ultrasound scans. The goal of medical image classification is to assist healthcare professionals in making more accurate and informed diagnoses, leading to better patient outcomes. An example of an application is the classification of skin cancer with AI vision system, which can help dermatologists diagnose skin cancer more accurately and confidently. This technology has the potential to revolutionize dermatology by providing quick, reliable and automated diagnosis of skin cancer which can improve patient outcomes while reducing time and costs associated with traditional diagnosis methods. The figure 1.4 shows the result of skin cancer classification program.



Figure 1.4: Result of skin cancer classification

1.5.2 The pharmaceutical industry

In the pharmaceutical industry image classification has become a valuable tool for the automatic inspection of identifying and analyzing the conventional forms and volumes of medications that do not meet quality standards. As instance the image classification in this operation detects the unfilled syrup's bottles or even the imprinted pharmaceutical tablets as the following figure 1.5 shows. Then the system declares the issues to the controlling team in order to take the necessary measures.

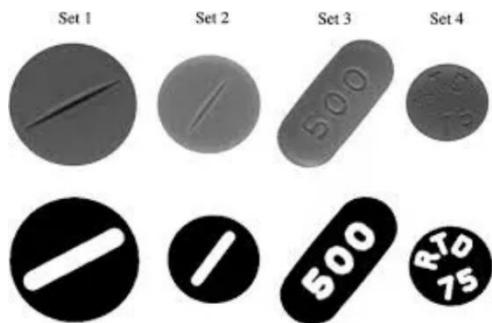


Figure 1.5: Cclassification of the printed tablets

1.5.3 Traffic monitoring and congestion detection

AI is utilized in avenue visitors management to help analyze real-time data from diverse approach of transportation, along with vehicles, buses and trains like shown in figure 1.6. The AI analyzes this statistics for patterns that might imply protection risks. This data is then used to indicate approaches to mitigate these risks and decrease the quantity of injuries that arise.

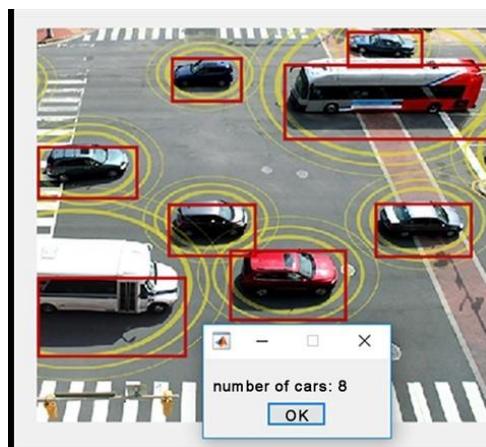


Figure 1.6: Trafic monitoring classification

1.5.4 Security face recognition technology

Security face recognition technology has become increasingly important in recent years as a means of identifying and verifying individuals in various settings. This advanced technology offers a high level of accuracy and convenience in access control, surveillance, and other security-related applications. Many organizations, including government agencies and businesses, have implemented security face recognition technology to enhance their security measures and provide a safer environment for employees and customers. With the ability of detecting and recognizing faces in real-time as shown in figure 1.7, security face recognition technology has proven to be an effective tool in preventing unauthorized access and detecting potential threats.



Figure 1.7: Facial recognition

2 Problem statement

Nowadays, in the agriculture sector, plant diseases have become a major concern for farmers because it lead to huge economic losses as well as for people who are passionate in growing plants in their gardens. Meanwhile Traditional methods of visual observation by plant protection experts for detection of plant diseases are expensive, time-consuming, and labor-intensive with a high error rate.

To address this issue, new methods are needed to enable early and accurate detection of plant diseases. One potential solution is developing our "plant diagnostician" system that can quickly and accurately detect the presence of pathogens, allowing farmers to take appropriate actions to reduce their impact.

Conclusion

In his chapter we have expounded the concept of computer vision and its useful techniques with highlighting on image classification by mentioning some of its useful applications in

several fields. We also over-viewed the scope of our project and the proposed solution to tackle it. In the next chapter we will provide a comprehensive explanation about many concepts that we are going to use in developing our "Plants Diagnostician".

Chapter2 : Theoretic Background

Introduction

Deep learning is a sub-field of machine learning that utilizes neural networks with multiple layers to learn hierarchical representations of data. It has gained immense popularity in recent years for its ability to recognize patterns and make predictions from complex data. In this chapter we will present a depth explanation of Deep learning, focusing on certain sub-classes of this fields.

1 Basic Concepts of Artificial Neural Networks

Artificial Neural Networks (ANN) are a subfield of artificial intelligence that attempt to mimic the structure and function of biological neural networks found in human brains. ANNs consist of connected units or nodes called artificial neurons that process information, with each neuron receiving inputs from several other neurons and producing an output. These outputs are then fed into other neurons or the final output layer of the ANN, producing a desired result.

1.1 Structure of an Artificial Neural Networks

ANN are a class of machine learning models that have gained popularity in recent years due to their ability to solve complex problems. ANN are inspired by the structure and function of the human brain, consisting of interconnected nodes (neurons) organized in layers. These networks, As shown in figure2.1 below, are composed of interconnected nodes, known as neurons, that process information and produce output. ANN is consist of vertical group of neurons, Each line represents a layer. There are three types of layers in a Neuron Network:

- The Input Layer is the first layer of an ANN and is responsible for receiving input data. This layer is usually not modified during training and simply passes the input data to the next layer.
- The Hidden layers are the layers of an ANN that lie in between the input and output layers.
- The output layer is the final layer of an ANN and produces the network's prediction.

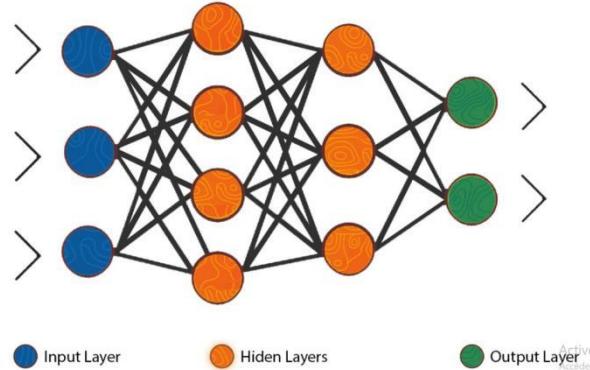


Figure 2.1: Layers of an Artificial Neuron Network

1.2 Perceptron

The perceptron is an artificial neuron that serves as the fundamental building block for various types of neural networks. [3] It is designed to take one or more inputs, perform a weighted sum of those inputs, and output a binary signal based on whether the resulting weighted sum is above or below a certain threshold, As shown in figure2.2. The perceptron is a linear classifier that can be trained using supervised learning to delineate between two classes of data points.

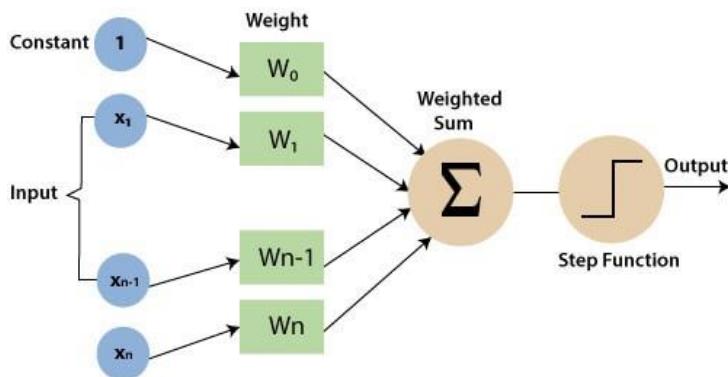


Figure 2.2: Perceptron

1.3 Activation functions

Artificial neurons are fundamental building blocks of artificial neural networks that are used to model and simulate the behavior of biological neurons. In an artificial neuron, the activation function plays a crucial role in sending an output signal or not based on the input it receives [5].

Activation functions can be Linear, Non-linear, Binary Step Function Moreover, choosing the right activation function can significantly impact the performance of a deep learning model. There are several activation functions commonly used in deep learning, including sigmoid, ReLU, tanh, and softmax:

- **Sigmoid Activation function:** The sigmoid activation function figure 2.3a, a widely used mathematical model in DL, is defined as a non-linear function that maps any input value to an output value between 0 and 1. This non-linearity makes it useful for binary classification problems, where the goal is to predict one of two classes. However, it can suffer from the vanishing gradient problem, where the gradient of the function becomes very small for large or small input values. The mathematical formula for the sigmoid function is:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.1)$$

- **ReLU Activation Function:** For instance, ReLU figure 2.3b is often employed as it has shown to be effective by being computationally efficient and addressing the vanishing gradient problem commonly associated with sigmoid activation functions. However, ReLU can suffer from the "dying ReLU" problem in which neurons become inactive and no longer contribute to the model's performance. The mathematical formula for the Relu function is:

$$f(x) = \max(0, x) \quad (2.2)$$

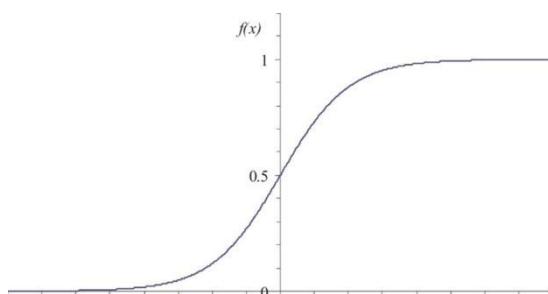
- **Tanh Activation Function:** The hyperbolic tangent (tanh) activation function (figure 2.3c), is a popular choice in neural networks. It is a mathematically convenient activation function due to its symmetric nature and range of values between -1 and 1, making it useful in situations where a normalized output is desired. However, similar to the sigmoid function, the tanh function can also suffer from the vanishing gradient problem. The tanh activation function can be described mathematically as:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.3)$$

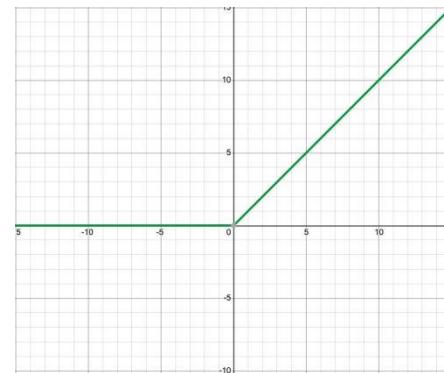
- **Softmax Activation Function:** The Softmax is a Machine Learning algorithm used for multi-class classification problems. The softmax function takes a vector of inputs and normalizes them into a probability distribution over K different classes. The output probabilities sum to 1, making it useful for multi-class classification problems.

$$f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (2.4)$$

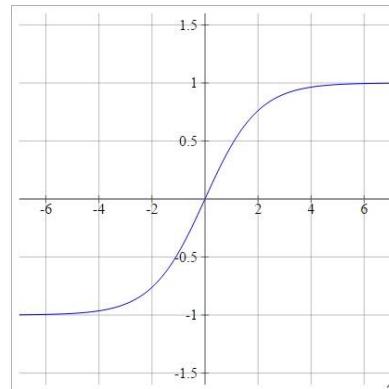
where z_i is the input to the function for the i^{th} class, K is the total number of classes, and $f(z_i)$ is the output probability for the i^{th} class.



(a) Sigmoid Activation function



(b) Relu Activation Function



(c) Tanh Activation Function

Figure 2.3: Activation functions

1.4 Loss Function

One of the key components in building a neural network model using deep learning is selecting an appropriate loss function. A loss function is a mathematical function that measures the difference between predicted and actual values of an output variable. The loss function guides

the learning process by quantifying the level of error between predicted and actual values, intending to minimize it [6].

1.4.1 Mean Squared Error Function

Mean Squared Error function is a commonly used loss function in deep learning, which calculates the average squared difference between predicted and actual values. The Mean Squared Error function finds frequent usage in regression problems, wherein the output constitutes a continuous variable. Typically, the predicted value of the target variable is derived by feeding the input through a neural network, and the network's final layer output is taken as the prediction.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.5)$$

where n is the total number of observations, y_i is the actual value of the target variable for the i -th observation, and \hat{y}_i is the predicted value of the target variable for the i -th observation

1.4.2 Cross-Entropy Function

The cross-entropy function 2.6 is a widely used loss function in deep learning that assesses the difference between the predicted probability distribution and the true probability distribution of the target variable.

$$H(p, q) = -\sum_{i=1}^n p_i \log q_i \quad (2.6)$$

where p represents the true probability distribution of the target variable and q represents the predicted probability distribution. n is the total number of possible outcomes.

1.5 Type of Neural Nets Learning

1.5.1 Supervised Learning

Supervised learning is a type of machine learning in which the algorithm learns from labeled data. The main goal of supervised learning is to enable the algorithm to make accurate predictions or decisions when given new, unseen data [7].

1.5.2 Unsupervised Learning

Unsupervised learning deals with unlabeled data and focuses on finding structure or patterns within that data. The labeled data used in supervised learning allows the algorithm to train on known inputs and outputs, which enables it to make predictions about new input [8].

2 Different types of Neural Networks in Deep Learning

2.1 Recurrent Neural Networks RNN

Recurrent Neural Networks (RNNs) are a type of neural network that are designed to handle sequential data by allowing feedback connections between the nodes in the network as it represented in figure 2.4 . This architecture allows RNNs to exhibit dynamic temporal behavior, making them ideal for processing time series data such as speech, music and video. One notable feature of RNNs is their ability to maintain an internal memory state, which can be used to store information about past inputs. This enables the network to make predictions based on previous context and helps it capture long-term dependencies within sequences. While traditional feedforward networks process input data one at a time without any regard for order or sequence, RNNs excel at tasks that involve analyzing patterns over time. Recently, they have been applied successfully across many fields including natural language processing , anomaly detection etc [9].

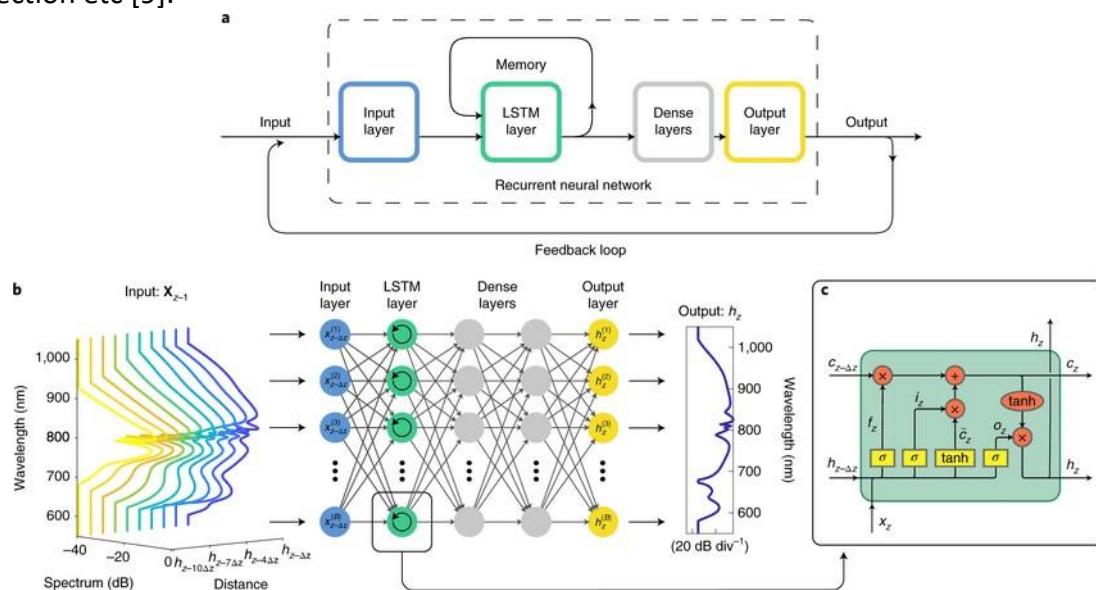


Figure 2.4: RNN structure

2.2 Graph Neural Networks GNN

Graph Neural Networks (GNN) have recently gained significant popularity due to their ability to work with complex data structures such as graphs. A graph is a mathematical structure consisting of nodes and edges that can be used to represent relationships between various entities as shown in the figure 2.5. GNNs are designed specifically for learning from these types of data, enabling them to perform tasks such as node classification, link prediction, and recommendation systems. One key advantage of GNNs over traditional neural networks is their ability to incorporate information about the entire network topology into their computations. This allows them to better understand how individual nodes interact within larger structures, resulting in more accurate predictions and higher overall performance on many different types

of problems.

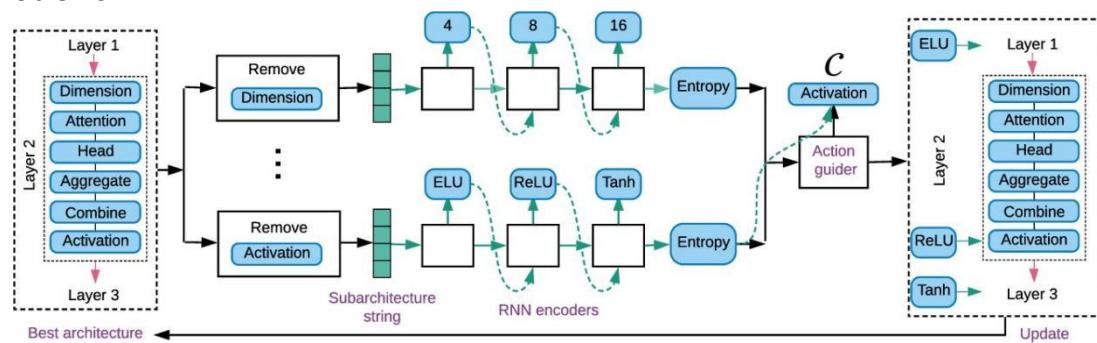


Figure 2.5: Graph Neural Network structure

2.3 Convolutional neural network CNN

A convolutional neural network (see figure 2.6) or CNN for short, is a kind of deep learning algorithm that has been widely adopted in the field of computer vision. This sophisticated technique involves multiple layers of interconnected neurons that are designed to automatically identify patterns and features within images. By applying mathematical operations known as convolutions across these layers, CNNs can effectively classify and recognize complex visual data with remarkable accuracy. With their ability to learn from vast amounts of training data, coupled with advancements in hardware technology such as GPUs and TPUs, CNNs have become one of the most powerful tools available for image analysis tasks ranging from object detection and segmentation to medical imaging diagnosis and others.

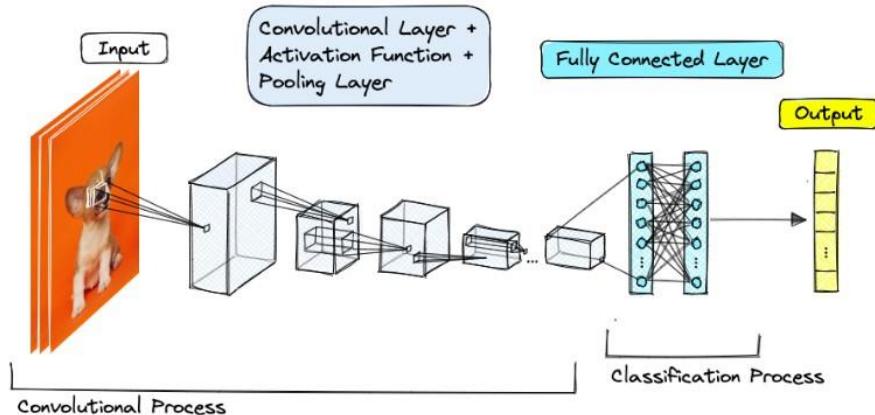


Figure 2.6: Typical architecture of CNN

3 Predefined models

3.1 AlexNet

AlexNet, introduced in 2012 by Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton, is a convolutional neural network comprising of five convolutional layers and three fully connected layers with approximately 60 million parameters. The impressive error rate of only 16 % led to its victory at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), making it the first deep learning architecture to do so , its architecture is shown below in figure ???. This accomplishment marked a significant milestone for computer vision which further inspired research into deep learning architectures. The use of rectified linear units (ReLU) contributed to its exceptional performance as compared to traditional activation functions like sigmoid or tanh since they allowed for faster training times. Another key feature that improved generalization performance was local response normalization (LRN) which reduced overfitting issues.

3.2 MobileNet

MobileNet is a pioneering neural network structure as displayed in figure 3.2 intended to meet the demands of mobile and embedded devices. Its distinctive features allow for excellent precision with substantially fewer computational resources compared to traditional convolutional neural networks (CNNs), making it highly effective. Despite its impressive efficiency, MobileNet does not sacrifice accuracy or functionality due to its incorporation of

advanced techniques such as depthwise separable convolutions and linear bottlenecks, which minimize computation while maintaining quality.

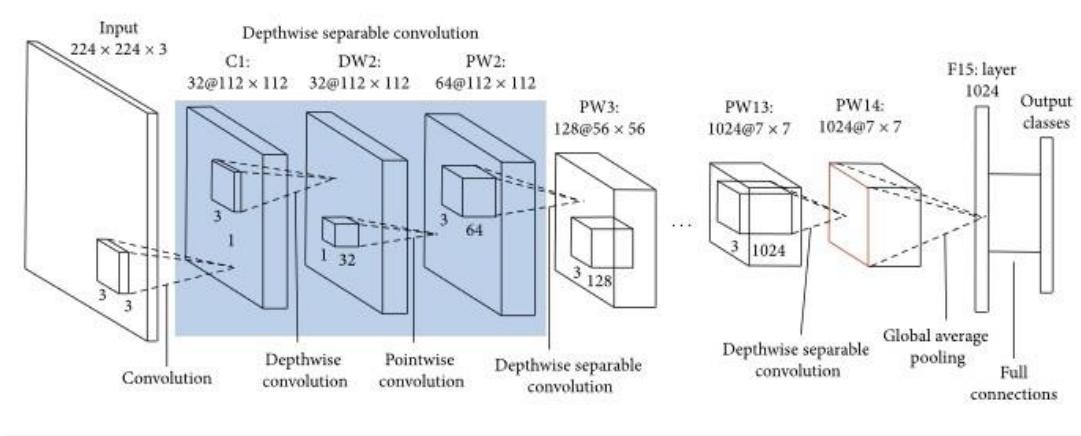


Figure 2.7: MobileNet neural network structure

3.3 VGG

Since we have defined VGG in the first chapter , The various types of VGG models that have been developed for different applications in the field of computer vision. These models are based on convolutional neural networks . The original VGG model was proposed by Simonyan and Zisserman in 2014 with a total of 19 layers. It is widely used for image classification tasks where it achieved state-of-the-art results on the ImageNet dataset. Later, variations of the VGG model were introduced which aimed to reduce computation time while maintaining high accuracy levels , we distinguish:

- **VGG16 algorithm:** Its CNN architecture consists of only 16 layers, but still performs comparably well to its predecessor. Another modification made to this architecture includes inserting batch normalization layers between each pair of convolutional filters. This helps improve training stability and allows for faster convergence during optimization processes.
- **VGG19 algorithm:** Its CNN architecture was developed by the Visual Geometry Group at the University of Oxford. This deep learning model has 19 layers as represented, and is highly regarded for its exceptional performance in image recognition tasks, such as classifying objects within an image. Despite being relatively simple compared to other architectures, VGG19's remarkable accuracy on large datasets demonstrates the power of deeper networks with increased capacity for feature extraction and classification.

conclusion

In this chapter, we have explained what a neural network is and we have introduced the most common examples of its predefined models so that we can get a better understanding of its functionality, especially to choose the most adequate model for building our "Plant Diagnostician" program. In Addition to this, we finally found that VGG16 fits the most, that's why we will provide a deeper explanation about it in the last but not least third chapter.

Chapter3 : Description and Implementation of the Plants Diagnostician

Introduction

In this chapter, we will outline the various steps involved in developing our program "Application of Machine Learning in Plants Diagnostician", thus we will delve deeper into the different perspectives of the chosen model the "VGG16", discussing some of the key libraries and frameworks used in its implementation. Then, we will display the experimented results so that we can be able to evaluate the accuracy of our model.

1 The chosen Model: VGG16

VGG16 is a deep convolutional neural network architecture figure 3.1 that was developed by the Visual Geometry Group (VGG) at the University of Oxford in 2014. It is composed of 16 layers, including 13 convolutional layers and 3 fully connected layers.

The input to the VGG16 network is a 224x224x3 RGB image. The first layer is a convolutional layer with 64 filters, each with a size of 3x3 and a stride of 1, followed by a ReLU activation function. This is followed by another convolutional layer with 64 filters, and then a max pooling layer with a pool size of 2x2 and a stride of 2.

The next two sets of layers are similar to the first, with two convolutional layers followed by a max pooling layer in each set. The first set has 128 filters, and the second has 256 filters.

The fourth set of layers has three convolutional layers, each followed by a max pooling layer. The first two layers have 512 filters, while the third layer has 512 filters with a size of 3x3 and a stride of 1.

VGG16 achieved state-of-the-art results in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014, where it significantly outperformed previous models by reducing the top-5 error rate from 0.26 to 0.74. Since then, the VGG16 architecture has become a popular choice for many computer vision tasks, including image classification, object detection, and image segmentation.

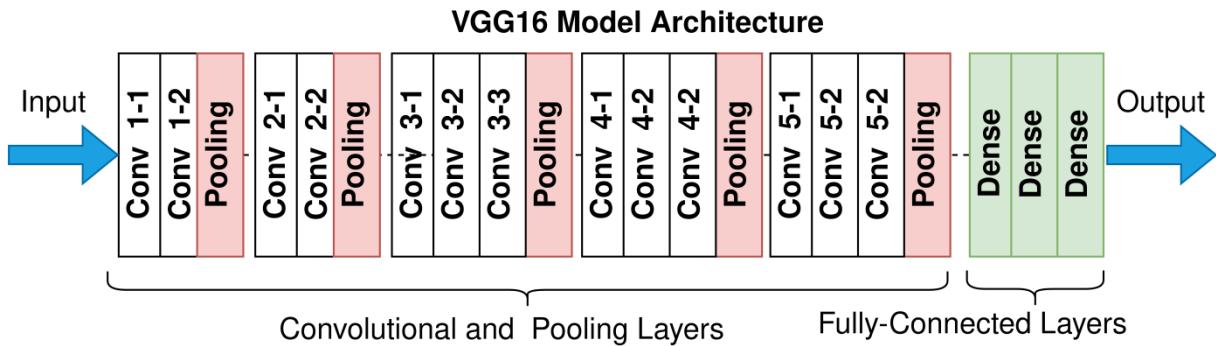


Figure 3.1: VGG16 model

- **Convolutional layer:** A convolutional layer performs a mathematical operation called convolution on its input, which is typically an image or a feature map produced by a previous layer. The convolution operation involves sliding a small matrix of weights, called a filter or kernel, over the input and computing the dot product between the filter and the portion of the input that is currently being covered by the filter. The result of this operation is a single value in the output feature map, which represents a particular feature or pattern that the filter is designed to detect.
- **Pooling layer:** Pooling layers are typically used to reduce the spatial size (i.e., the width and height) of the input feature maps produced by convolutional layers, while retaining the important features. This can help to reduce the computational complexity of the network and prevent overfitting.
- **Fully-connected layer:** A fully-connected layer, also known as a dense layer, is a type of neural network layer where each neuron in the layer is connected to every neuron in the previous layer. In other words, each neuron in a fully-connected layer receives inputs from all the neurons in the previous layer, and produces a single output value that is then passed on to the next layer.

2 Software (Technology) stack

Python has become one of the most widely-used programming languages in the field of computer vision, owing to its numerous libraries and frameworks for machine learning.

Particularly, Python is a popular language for image classification tasks, as it provides developers with a range of powerful tools for building and training neural networks, and this is the reason beyond the fact of choosing Python as programming language in our project .

Python provides a range of libraries and frameworks that are well-suited for image classification, including Open-CV, TensorFlow, Keras etc. These tools offer high-level APIs for building and training neural networks, as well as pre-trained models that can be used for a variety of image classification tasks. In this context let us mention more details about :

- **Open-CV** : (Open Source Computer Vision) is a commonly used Python library for computer vision and image processing tasks. It provides a wide range of functions and tools that make it easier to preprocess, extract features, and train models for image classification.
- **TensorFlow** : is a free and open-source framework developed by Google Brain Team that provides a wide range of tools and libraries for building machine learning models particularly for creating and training deep learning models including convolutional neural networks .TensorFlow also offers a variety of APIs for different languages, precisely Python language .
- **Keras**: was originally developed as a standalone library for deep learning, but it was later integrated with TensorFlow as its official high-level API. Keras provides a simple and intuitive interface for building deep learning models, making it easier for researchers and developers to experiment with different neural network architectures.[8]

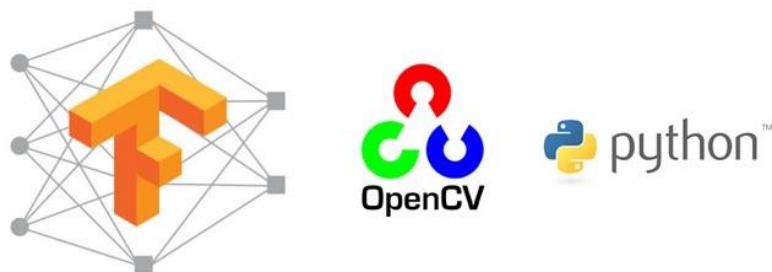


Figure 3.2: Software stack

3 Implementation and experience results

3.1 Dataset

Our dataset is composed of eight classes as shown in this table below 3.1. Indeed it includes four apple classes, two peach classes and finally two strawberry classes. Actually we have chosen these plants because of the morphology of their leaves which is wider enough to be scanned easily by this program. In fact each data class contains the healthy version and the affected version of leaves, and this is for better interpreting the issue.

Table 3.1: The dataset classes

N°	Plant	Disease	Files
1	Apple	Apple scab	630
2	Apple	Black rot	621
3	Apple	Cedar apple rust	275
4	Apple	Healthy	1645
5	Peach	Bacterial spot	2297
6	Peach	Healthy	360
7	Strawberry	healthy	456
8	Strawberry	Leaf scorch	1109

3.2 Data preparation

Data preparation refers to the systematic procedure of refining, arranging, and converting unprocessed data into a structured format that is conducive for analysis and more accurate through the prescribed machine learning model. [9] The process comprises various stages including :

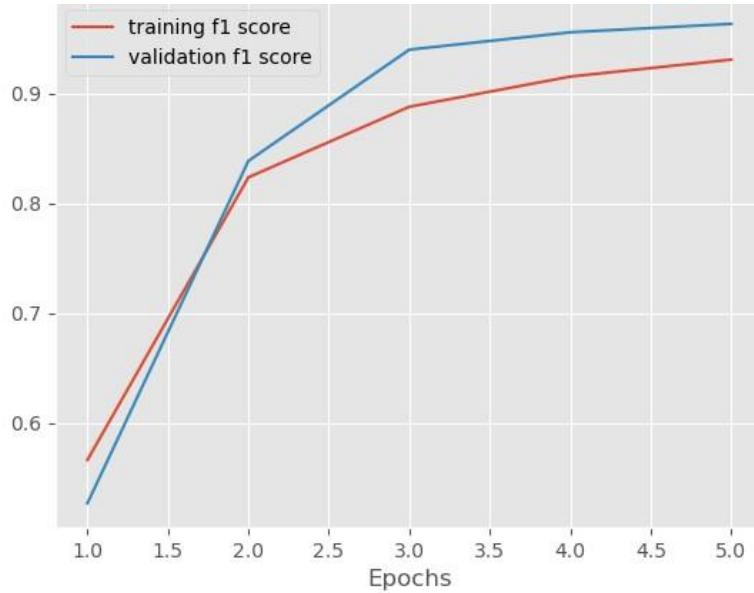
- **Data Splitting:** The goal of data splitting is to divide the available dataset into three separate subsets: a training set, a validation set, and a test set. This is done to ensure that the model is trained on one set of data, validated on a separate set, and lastly tested on a completely independent set. It is an important step in data preparation as it helps

to prevent over-fitting of the model to the training data, which can result in poor generalization to new, unseen data.

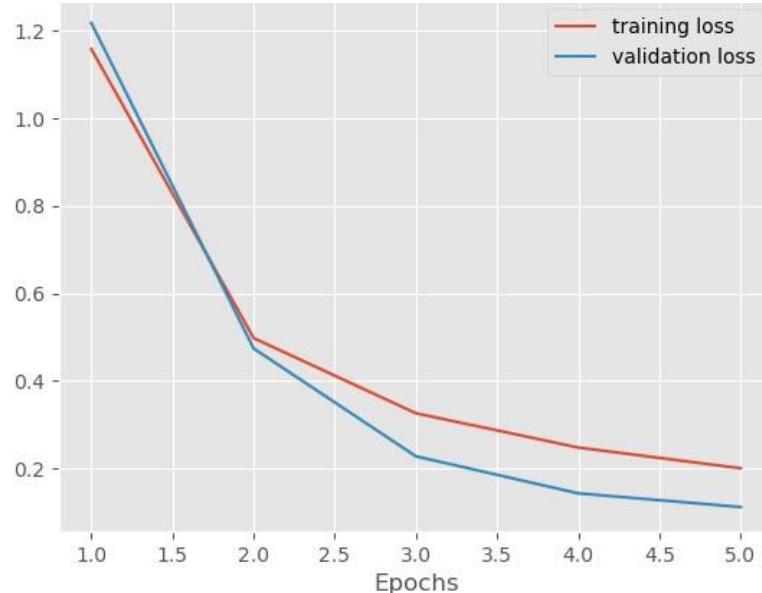
- **Data labeling:** is the act of assigning labels to data, it is simply categorizing data into different classes of data. Actually this is a crucial step in supervised machine learning, where labeled data is used to train models to recognize patterns and make accurate predictions.
- **data preprocessing:** Data preprocessing refers to the crucial preliminary steps of refining and enhancing raw data in preparation for its input into a machine learning algorithm. This indispensable process encompasses several procedures, including cleansing, normalizing, transforming and selecting or extracting features from the data. The aim of this critical step is to generate a well-polished and functional dataset that has been specifically tailored to optimize performance in conjunction with the particular machine learning algorithm being employed.
- **Data augmentation:** Data augmentation is a sub category of data preparation that aims to make the dataset bigger and more diverse by creating new versions of the existing data. It works by applying different modifications, like rotating or flipping an image, to the original data. It aims to make the machine learning model more accurate and reduces its probability of errors while processing novel data .

3.3 Experimental results

Accuracy is a common evaluation metric used in machine learning to measure how well a model is able to correctly classify or predict data. Our algorithm accuracy up to 90% as shown in figure3.3

**Figure 3.3:** Trainig and Validation Accuracy

Loss (or cost) is a measure of how well a model is able to fit the training data. The figure 3.4 show how the algorithm is about to reduce the loss during the 5 Epochs reaching 20%

**Figure 3.4:** Trainig and Validation loss

Conclusion

In this chapter we have presented our model of classification by giving a detailed explanation about its architecture meanwhile the crucial steps about preparing our dataset to be classified successfully in order to reach critical experimental results.

General Conclusion

In this report, we have presented the different perspectives on how we have built this project "Application of Machine Learning for Plant Diagnostician", from a theoretical background that was dedicated for more explanations about the tools, techniques and methods that were widely essential to comprehend the next strategic steps, to the final implementation and experimented results, and this was by diving deep into the science of artificial intelligence in order to realize the functioning of the neural network models and their diverse and inspirational architectures so that we can enlarge our outlook since we are supposed to choose the most efficient model for our case. in this context we selected the VGG16 model for our image classification task because of its efficient and effective performance. Furthermore, we mentioned that it is crucial to prepare the data before executing this process in order to obtain critical and accurate results as it was well explained in our case. Moreover, we are willing to enlarge the data set and make it encyclopedic, we also aim to modify the program by adding some more advanced tools to update it in order to process and classify a huge amount of data. Overall, the project aims to facilitate and overcome numerous obstacles in the agriculture sector. In Addition, this project is a promising step towards revolutionizing the agricultural industry asit demonstrates the potential in providing cost-effective solutions to farmers and policymakers by accurately diagnosing plant diseases and identifying stress factors affecting crop yields, which can lead to better-informed decisions and increased productivity.

Bibliography

[10]S.Sharma-International Journal of Engineering Applied Sciences and Technology In Neural Network, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, Published Online April 2020 in IJEAST. Pages 310-316

[11]AN.Sharkawy-Journal of Advances in Applied & Computational Mathematics-2020- pages 8-19.

Netography

[1] I.Mansouri-https://medium.com/@ilias_mansouri/part-2-overview-of-computer-vision-methods69c56843c567. 04/02/2023

[2] B.Luis-<https://www.aquiladata.fr/insights/classification-dimages-et-detection-dobjets-par-cnn/>.

10/02/2023

[3] S.Sharma-<https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>. 11/02/2023

[4] <https://ijeast.com/papers/310-316,Tesma412,IJEAST.pdf> 9/03/2023.

[5] B.Mehligh-Deep learning and neural network <http://physics.gu.se/frtbn/joomla/media/mydocs/NeuralNetworks.pdf>. 15/03/2023

[6]<https://towardsdatascience.com/common-loss-functions-in-machine-learning-46af0ffc4d23>. 01/04/2023

[7] J.Brownlee-supervised and unsupervised learning https://scholar.google.com/citations?view_op=view. 01/04/2023

[8] Karpathy-Convolutional Neural Network <https://cs.stanford.edu/people/karpathy/convnetjs/>.

2/04/2023

[9] J.Brownlee-Keras https://scholar.google.com/citations?view_op=view_citation. 04/02/2023