Ecole Nationale Supérieure d'Informatique et d'Analyse des Systèmes - RABAT

End Of Year Project Report : Handwritten Digit Recognition using CNN Model

Option : ISEM

Realised by :

Dkheila Anass
Benayyad Abdessalam

Supervised by :

Pr. Faissal El Bouanani

Jury :

Pr. Faissal El Bouanani
Pr. Alaoui Ismaili

Academic year 2021/2022

Acknowledgment:

We would like to thank all the people who contributed to the realization and success of our project closing the 2nd year of engineering training in Embedded and Mobile Systems at ENSIAS

First of all, we would like to thank our project supervisor, **Pr. Faissal El Bouanani**, for his transmission of knowledge, his availability and his advice, which contributed to the smooth running of the project.

I would also like to thank the members of the jury for their careful reading of this report, as well as for the remarks they will address to me during the defence in order to improve my work.

Finally, I would like to thank all the teachers of our university "École Nationale Supérieure d'Informatique et d'Analyse des Systèmes" for sharing their knowledge and experience with us throughout this year.

Abstract

The field of machine learning and deep learning research is expanding rapidly these days. Handwriting recognition is an example of this field of work in everyday life. Data storage on digital media is a method that almost everyone now employs. Simultaneously, it has become a requirement for people to store their notes in digital media and even take notes directly in the digital environment. As a response to this need, applications that use machine learning and deep learning algorithms to recognize numbers, characters, and even text from handwriting have been developed. Furthermore, these applications can recognize and convert numbers, characters, and text from handwriting.

In this project, the performance of the deep learning algorithm **CNN**, which is commonly used in handwriting recognition applications is investigated in the case of handwritten digits. As a result of this study, the accuracy of the model developed reached **99.15%**. We also developed a handwriting recognition system for numbers similar to these mentioned applications. A Web interface was developed for end users where they can directly draw the digit and recognize it straight away in order to show the instant performance of some of this algorithm and allow them to experience the handwriting recognition system.

Key Words : Handwriting Recognition, Deep Learning, MNIST, CNN.

Table des matières

Table des figures

List of abbreviations

- **HDR :** Handwritting Digit Recognition.
- **DL :** Deep learning.
- **CNN :** Convolutional Neural Network.
- **MNIST :** Mixed National Institute of Standards and Technology.
- **ANN :** Artificial Neural Network.
- **OCR :** Optical Character Recognition.
- **SGD :** Stochastic gradient descent.

Introduction

Algorithms are becoming increasingly important in our lives as technology advances. These algorithms accelerate and reduce the workload. Deep learning learning algorithms, in particular, are constantly improving by mimicking human behavior.

Handwriting recognition as an evolving field, it is are becoming increasingly important for the following reasons : For data to be stored in digital media with recognition after being handwritten, and for previously written data to be transferred to digital media as optical characters. For example Handwriting recognition systems can be used for reading and archiving old documents, bank checks, and letters. Online handwriting recognition applications are also widely used in educational applications that support handwriting recognition in mobile devices, particularly applications that help ease the life of individuals with disabilities.

The computer recognition of handwritten letters, numbers, and characters is a process that might seem simple but it is difficult for computers. In other words, computers struggle to understand lines, symbols, and their combined shapes at the word level. Handwriting characteristics such as the presence of different characters, the different handwriting make it difficult for computer systems to recognize handwriting.

In this project, detailed information about the deep learning algorithm CNN ; Convolutional neural networks, which are a subset of deep neural networks, are frequently used for image analysis, but they may also be utilized for tasks like video recognition, recommendation systems, natural language processing, etc. Compared to previous image classification algorithms, convolutional neural networks require very less pre-processing ; used in the field of handwriting recognition was given and experiments were conducted to measure it's performance in this field and discover the most efficient parameters and compare the results on the MNIST dataset.

Chapitre 1

General Context of the project :

1.1 Introduction :

Handwriting recognition is a computer processing that aims to convert an image of a printed or handwritten text into text in a file in digital format.

In order to situate our-self in this section, we present the different aspects of handwriting recognition, then we define a system of recognition of handwritten numbers as well as its fields of application, focusing on the different problems of the recognition of handwritten numbers.

1.2 Aspects of handwriting recognition :

Handwriting recognition technologies can be divided into two categories. The first is an interactive (online) system, and the second is a non-interactive (offline) system.

Online and offline recognition are two types of handwriting recognition, each with its own set of acquisition tools and recognition algorithms. The first step in developing a recognition system that can recognize any character in any format is to convert the writing into numerical quantities that are suitable for the processing system, with as little degradation as possible. Depending on the mode of acquisition, there are several types of handwriting recognition systems :

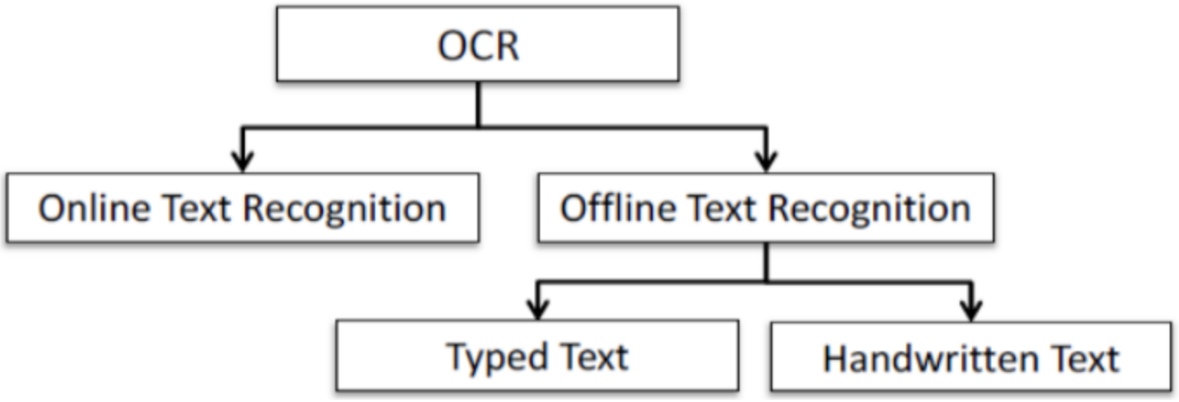

FIGURE 1.1 – Aspects of Handwriting Recognition

— **Online handwriting recognition :** is performed in real time, it is performed during character tracing, which makes it possible to obtain a good correction and modification, depending on the response given to the overlapping recognition phase to the acquisition phase. Commonly used means of online input are the graphic tablet with an electronic pen and the touch screen.
— **Offline handwriting recognition :** where information is in the form of a set of pixels that represents the image of an already existing text, obtained by a scanner or camera.

Offline writing recognition is more complex than online writing, due to the presence of noise in the image acquisition process and the loss of temporal information such as writing order and speed.

1.3 Handwritten digit Recognition System :

1.3.1 Definition :

A digit recognition system is a machine that recognizes numbers from various sources such as emails, bank checks, papers, images, vehicle plates, bank checks, and digital entries in hand-filled forms...

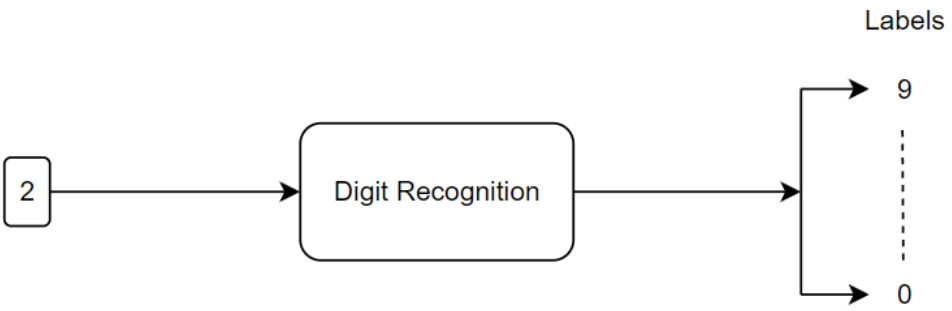

FIGURE 1.2 – Digit recognition flux.

1.3.2 Application Domains :

Because of a variety of potential applications such as :
— Manuscript address interpretation
— Credit authentication.
— Medical prescription reading.
— Social protection.
— Postal code reading.
— The processing of bank checks.

1.3.3 Project Problematic :

The main problems with automatic number recognition relies on handwritten numbers that are not always the same size and orientation because writing differs from person to person, so the general problem would be classifying the numbers based on similarities between numbers such as 1 and 7, 5 and 6, 3 and 8...

The problem is exacerbated when multiple people write the same number in different handwriting. Finally, the uniqueness and variety of handwriting of various individuals influence the formation and appearance of numbers as well.

1.3.4 Objective :

The goal of this project is to develop a model capable of recognizing and determining handwritten numbers from the MNIST data base using concepts from deep learning namely Convolutional Neural Networks.

1.3.5 Conclusion :

In this chapter, we defined the overall context of our end-of-year project by presenting firstly, the aspects of handwriting recognition then a definition of the digit recognizing systems. Finally, we discussed the research problem and goals.

Chapitre 2

Analysis and Conception

2.1 Introduction

In this chapter we are going to introduce some prerequisites and basic knowledge of deep learning to make it easier to realize this project.

2.2 Deep Learning

One of the machine learning techniques is deep learning. The main distinction between deep learning and the majority of machine learning methods is feature extraction. In fact, the algorithm must be aware of the characteristics that make up the data for machine learning to operate. The algorithm must receive these features as input in the form of feature maps. The output of running data through a filter is a feature map. In order to guarantee that the data is accurate and contains all necessary components, it is thus a processing of the data that is frequently done at the start of the process. Though it has been included into deep learning process,this aspect of feature extraction is not carried out automatically by machine learning algorithms.

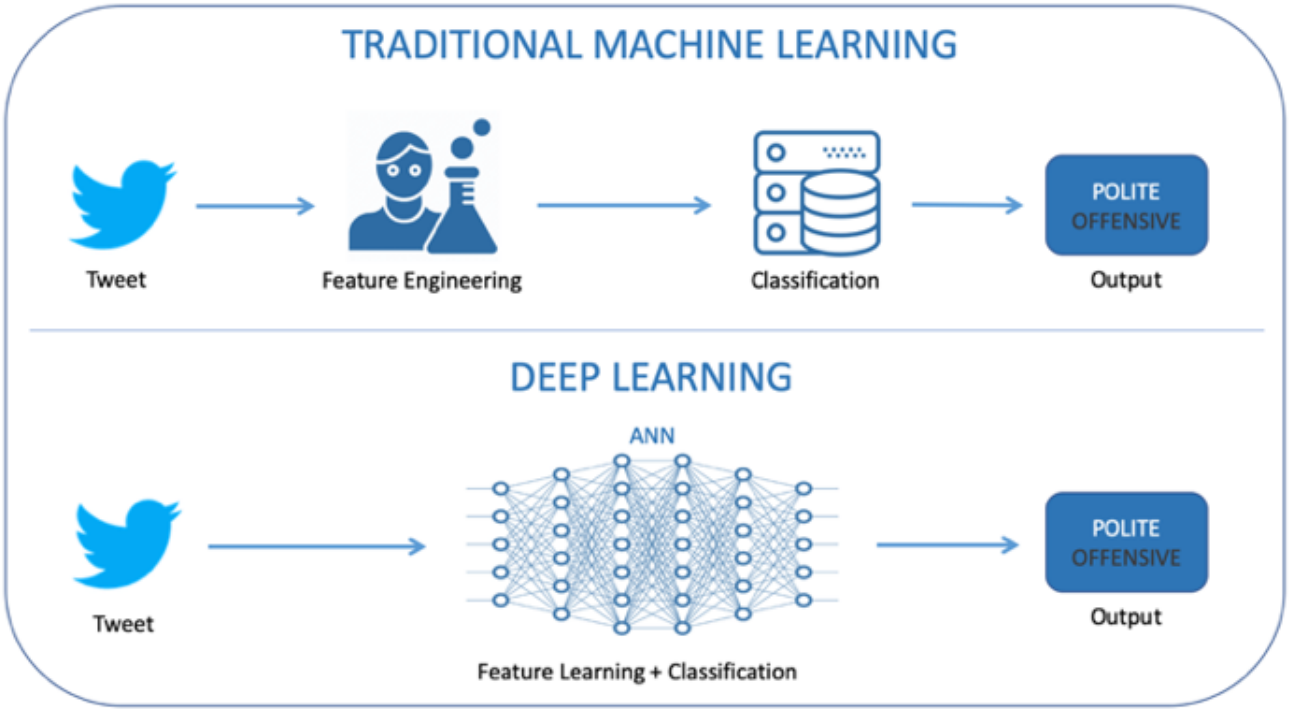

FIGURE 2.1 – Difference Between Machine Learning & Deep Learning

2.3 Neural networks

2.3.1 Definition :

Neural networks are a collection of algorithms that were influenced by how humans think and function. When you open your eyes, the majority of what you see is known as data, which is processed by the neuron (data processing cells) in your brain, which then helps you recognize your surroundings. That is how comparable the operation of neural networks is. They start with a big amount of data, process it (i.e., identify patterns), and then output the results.

2.3.2 Objectives :

Because they are artificial, unlike the neurons in your brain, neural networks are frequently referred to as Artificial Neural Networks (ANN's). They replicate the structure and operation of neural networks in an artificial way. In order to tackle certain issues, ANNs are made up of a significant number of closely linked processing components (neurones). Like people and children, ANNs also learn by doing. Through a learning process, an ANN is tailored for a particular application, such as pattern recognition or data categorization, image identification, or speech recognition.

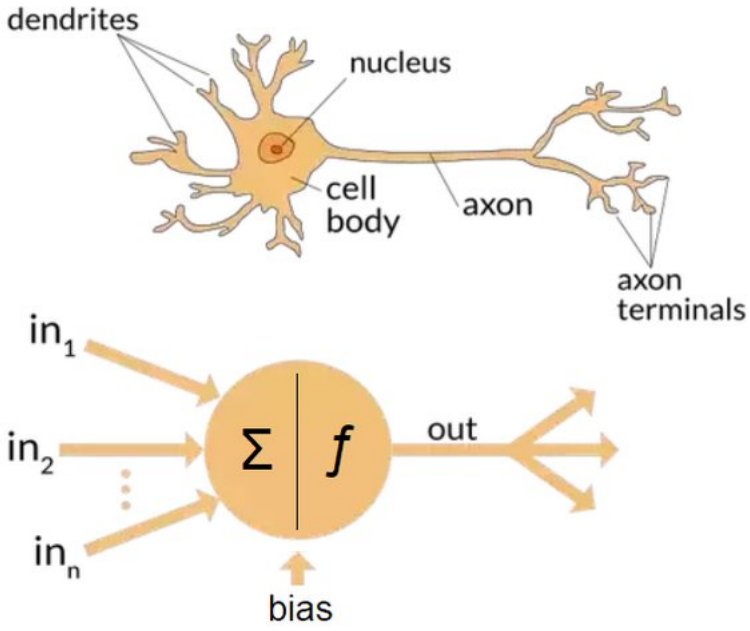

FIGURE 2.2 – Neural Networks

2.3.3 CNN : Convolutional Neural Network

Convolutional neural networks (ConvNets or CNNs) are one of the major types of neural networks used to recognize and classify pictures. A few applications for CNNs include object identification, facial recognition, and Handwriting Recognition...

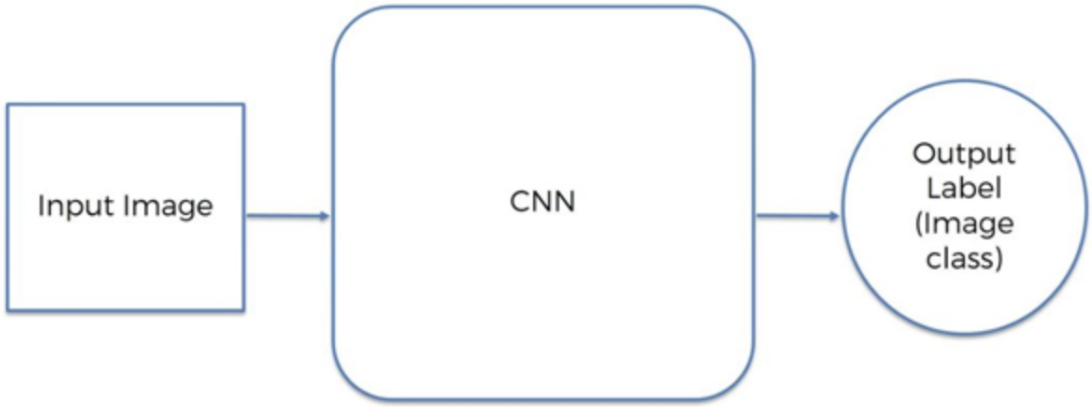

FIGURE 2.3 – CNN : Convolutional Neural Network

CNN image classifications takes an input image, process it and classify it under certain categories (Eg., Dog, Cat, Tiger, Lion). Computers sees an input image as array of pixels and it depends on the image resolution.

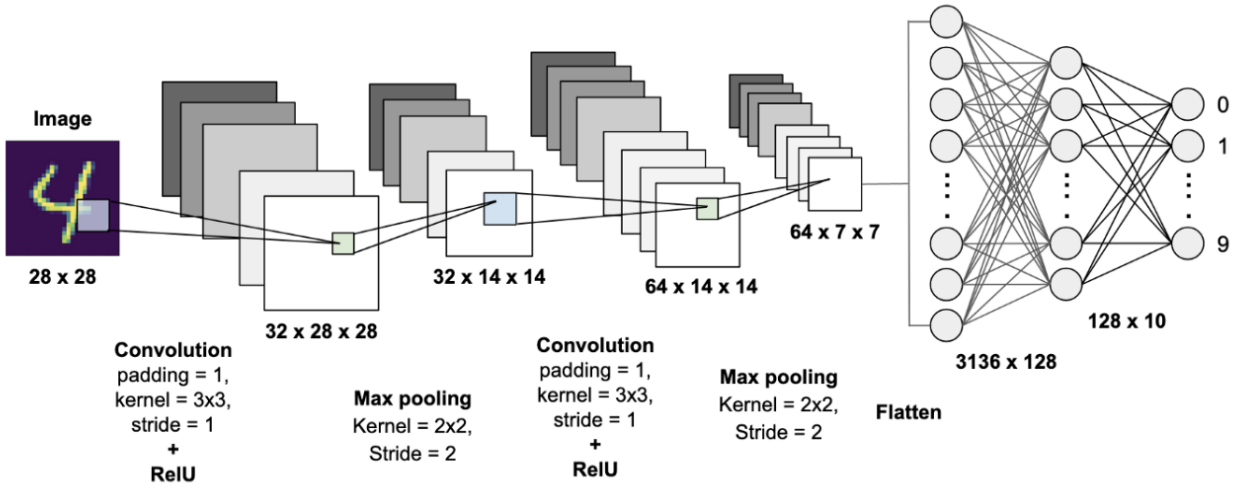

FIGURE 2.4 – MNIST Handwritten Digits Classification using a Convolutional Neural Network

2.3.4 Optimizer

In order to reduce losses, optimizer methods or algorithms are used to change the weight [of your network of neurons] among other attributes. The decision among a large number is also available here and will be based on the goal to be accomplished as well as the capabilities of the equipment being used. Among these, [Patterson et al., 2018] specifically describes the operation of the stochastic gradient descent (SGD) method, the RMSProp method, and the method Adam. The choice of optimizer depends on the rate of learning, which for some optimizers must be manually determined and is automatically determined for others.

2.3.5 Pipeline

To realise our project we follow the pipeline represented in the figure bellow.

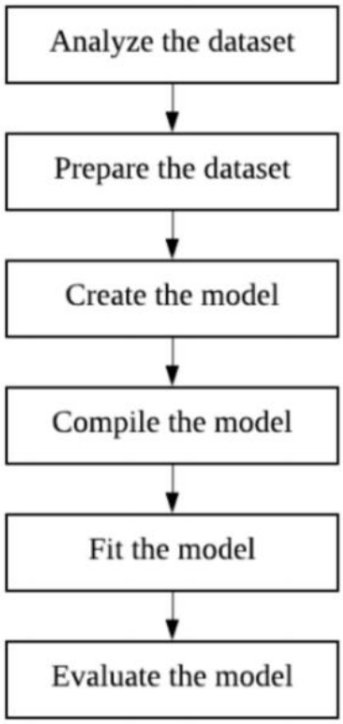

FIGURE 2.5 – Handwriting recognition processing steps

2.4 The MNIST dataset

MNIST is the most well-liked dataset among the thousands of datasets that are now on the market among machine learning and deep learning aficionados. The MNIST dataset includes more than **60,000** training pictures of handwritten digits from **0** to **9** and more than **10,000** images for testing. The MNIST dataset has 10 distinct classifications. The handwritten digit pictures are shown as a **28** by **28** matrix.

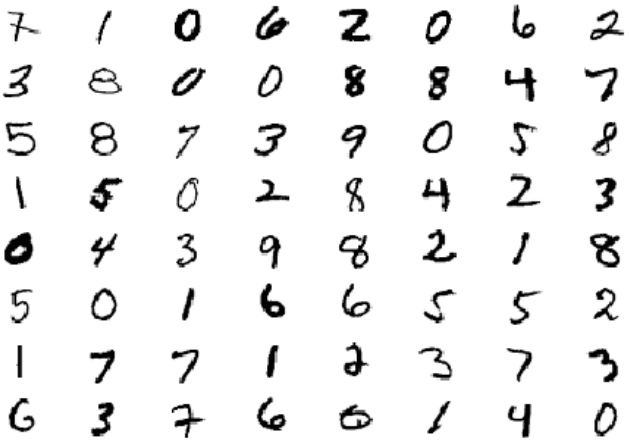

FIGURE 2.6 – Mnist dataset

This is the link to download data : **Mnist**.

2.5 Phases of Handwriting digit recognition

2.5.1 Import libraries and dataset

We import all the necessary modules at the start of the project so that our model may be trained. The Keras package already has many datasets, and MNIST is one among them, so we can quickly import the dataset and get to work. To obtain the training data with labels and the testing data with labels, we use the mnist.load data() method.

```
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense,Dropout,Conv2D,MaxPooling2D,Flatten,BatchNormalization
from tensorflow.keras.datasets import mnist

import tensorflow.keras.backend as K
```

```
#first we will load dataset
```

```
(x_train,y_train),(x_test,y_test) = mnist.load_data()
```

2.5.2 The Data Preprocessing

Model cannot take the image data directly so we need to perform some basic operations and process the data to make it ready for our neural network. The dimension of the training data is (60000*28*28). One more dimension is needed for the CNN model so we reshape the matrix to shape (60000*28*28*1).

2.5.3 Create the model

It's time to start building the CNN model for this data science project using Python. The two wheels of a CNN model are pooling layers and convolutional layers. The success of CNN in classifying images is due to the practicality of its use with grid-structured data. For the model compilation, we'll employ the Adam optimizer.

2.5.4 Train the model

To start the training of the model we can simply call the model.fit() function of Keras. It takes the training data, validation data, epochs, and batch size as the parameter.

The training of model takes some time. After succesful model training, we can save the weights and model definition in the 'model.h5' file.

2.5.5 Evaluate the model

```
score ,acc = model.evaluate(x_test,y_test)
print("Score is :",score)
print("Accuracy :",acc)
```

```
10000/10000 [==============================] - 2s 224us/sample - loss: 0.0360 - accuracy: 0.9915
```

FIGURE 2.7 – Accuracy Score

2.6 Conclusion

In the section,we have presented some general prerequisites and technologies that will help in the realization of this model then we showcased the pipeline taken to create and deploy a deep learning HDR model based on CNN.

Chapitre 3

Realization & Validation :

3.1 Tools used for the realization :

3.1.1 Visual Studio Code :

FIGURE 3.1 – VS Code Logo :

Visual Studio Code also known as VS Code is a free open source text editor by Microsoft. VS Code is available for Windows, Linux, and macOS. Although the editor is relatively lightweight, it includes some powerful features that have made VS Code one of the most popular development environment tools in recent times.

3.1.2 Python :

FIGURE 3.2 – Python Logo :

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect

existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. Python is commonly used for Artificial Intelligence applications namely Data science which makes a perfect tool to use for our project.

3.1.3 Tenserflow :

FIGURE 3.3 – Tenserflow Logo :

TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML-powered applications.
TensorFlow was originally developed by researchers and engineers working on the Google Brain team within Google's Machine Intelligence Research organization to conduct machine learning and deep neural networks research. The system is general enough to be applicable in a wide variety of other domains, as well.
TensorFlow provides stable Python and C++ APIs, as well as non-guaranteed backward compatible API for other languages.

3.1.4 Keras :

FIGURE 3.4 – Keras Logo :

Keras is a deep learning API written in Python, running on top of the machine learning platform TensorFlow. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result as fast as possible is key to doing good research.
Keras is :

- **Simple :** but not simplistic. Keras reduces developer cognitive load to free you to focus on the parts of the problem that really matter.
- **Flexible :** Keras adopts the principle of progressive disclosure of complexity : simple workflows should be quick and easy, while arbitrarily advanced workflows should be possible via a clear path that builds upon what you've already learned.
- **Powerful :** Keras provides industry-strength performance and scalability : it is used by organizations and companies including NASA, YouTube, and Waymo.

3.1.5 JavaScript & CSS & HTML :

(a) JS Logo (b) CSS & HTML

HTML is the markup language that we use to structure and give meaning to our web content, for example defining paragraphs, headings, and data tables, or embedding images and videos in the page.

CSS is a language of style rules that we use to apply styling to our HTML

content, for example setting background colors and fonts, and laying out our content in multiple columns.

JavaScript is a dynamic computer programming language. It is lightweight and most commonly used as a part of web pages, whose implementations allow client-side script to interact with the user and make dynamic pages. It is an interpreted programming language with object-oriented capabilities.

3.2 Validation & Results

3.2.1 GUI :

The figure below represents the interface :

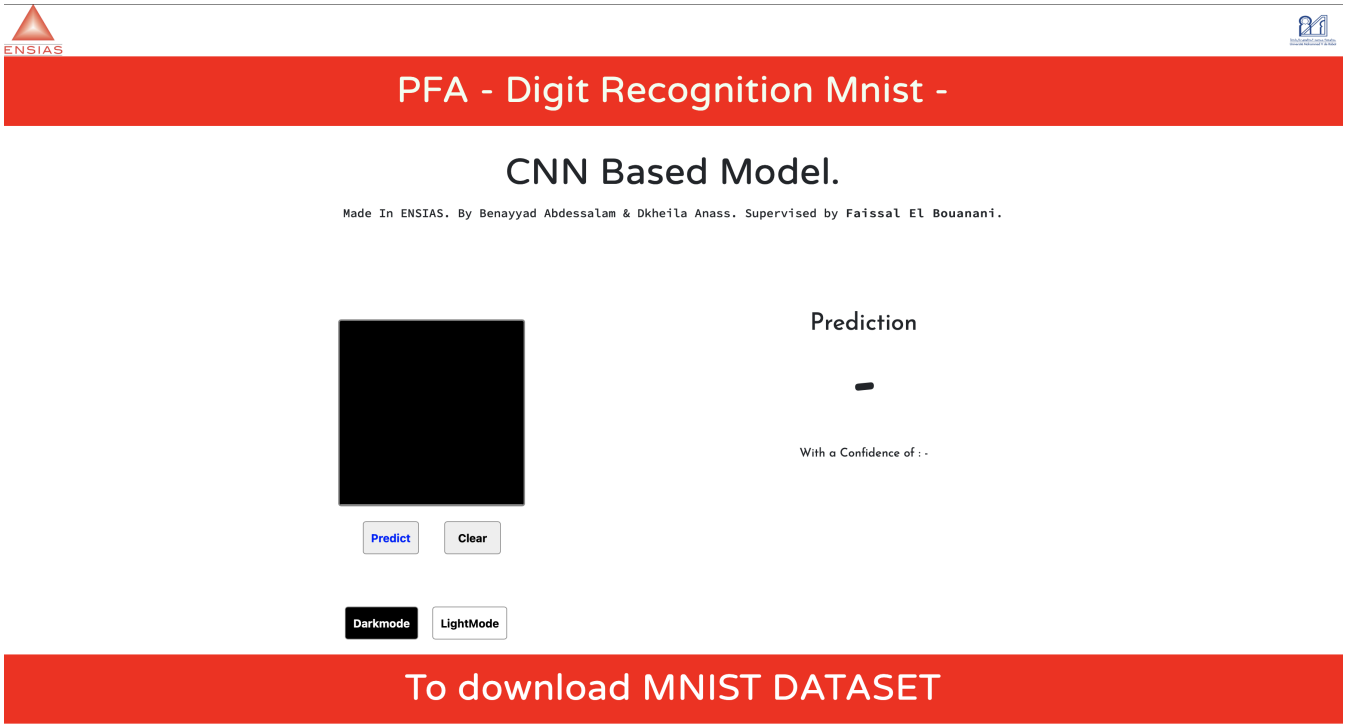

FIGURE 3.5 – Interface of the web app

3.2.2 Validation :

in these figures we validate the model by having it predict the various digits scaling from 0-9 drawn on the user interface by the users as shown in the pictures below :

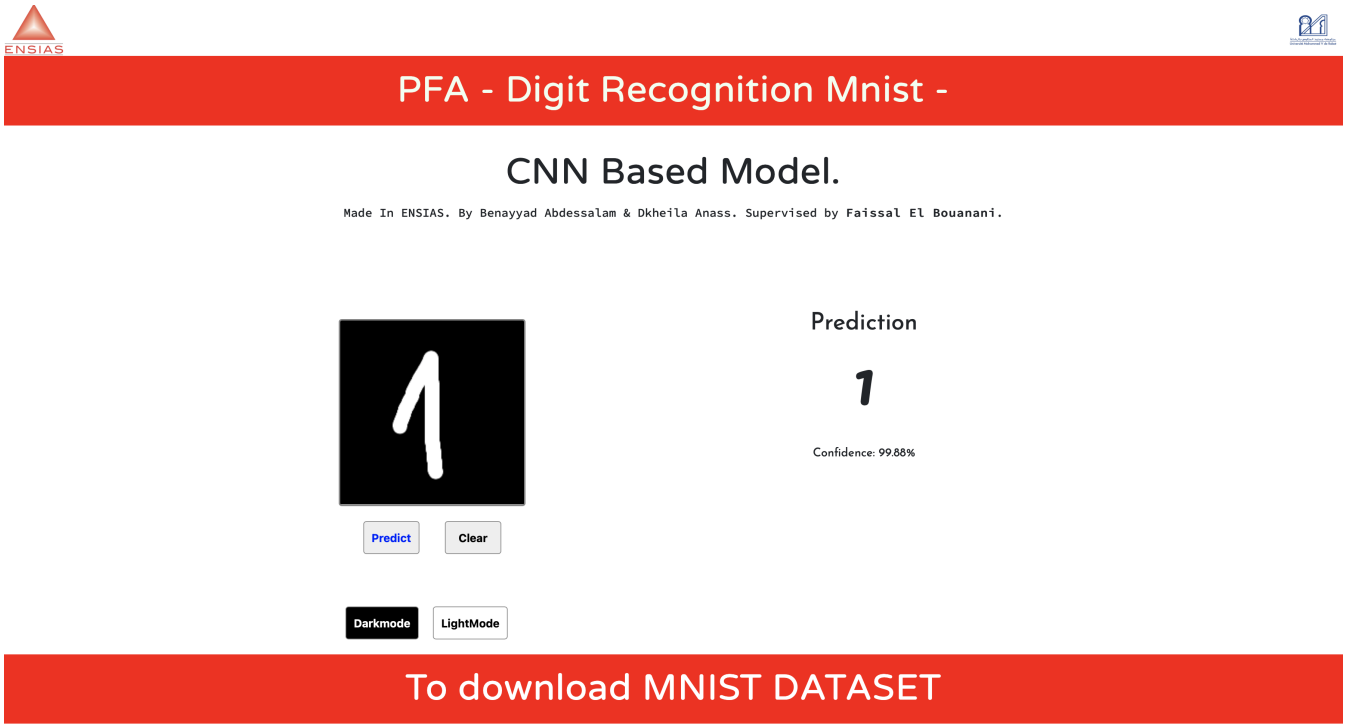

FIGURE 3.6 – Prediction with a confidence of 99.88%

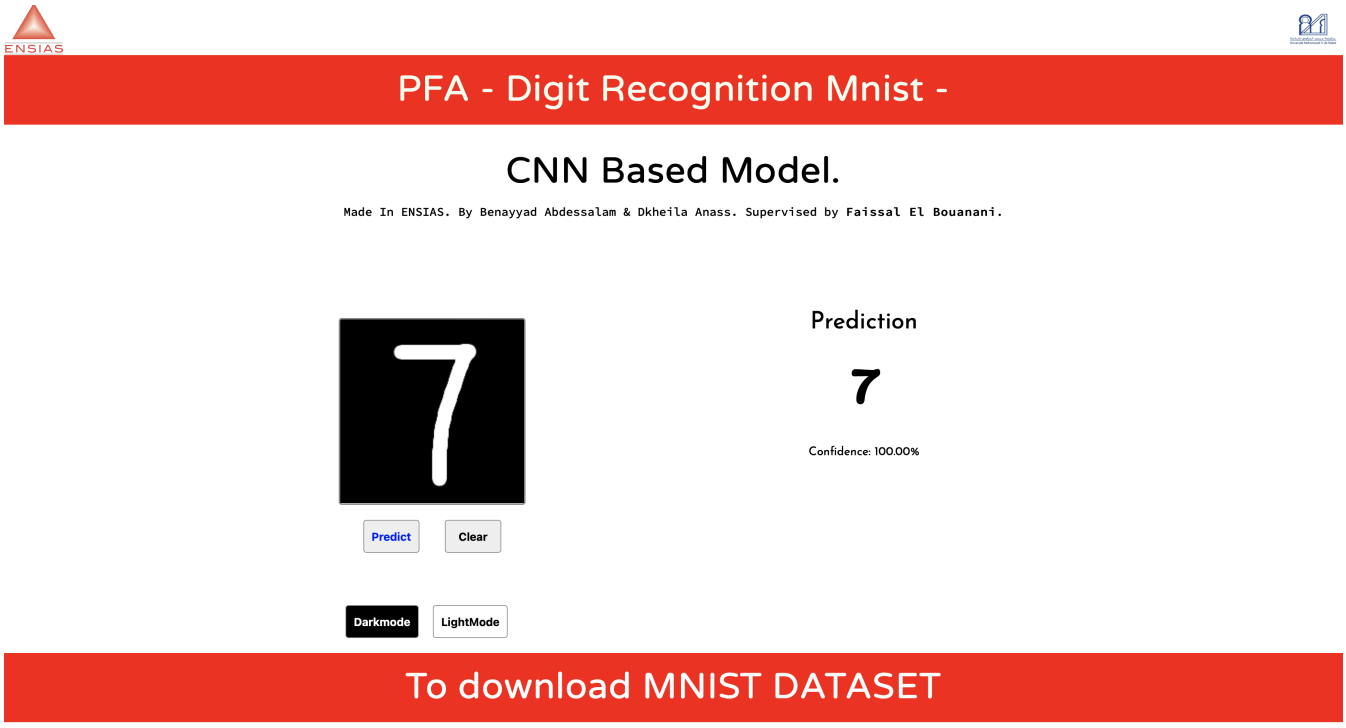

FIGURE 3.7 – Prediction with a confidence of 100%

3.2.3 Experimental Results :

In this section we're going to showcase the Accuracy results of our model and Compare the success rates of optimization parameters :

Optimization Parameters	Accuracy Rate
Adadelta	84.53%
Adam	99.15%

FIGURE 3.8 – Comparison of success rates of optimization parameters

Chapitre 4

Conclusion & Perspectives

In this study, we firstly introduced the general context of the project then we represented the various prerequisites and knowledge/techniques that can be helpful in realizing this project, finally we discussed the tools used for the realization of our HDR model and showcased some experimental results.

When it comes to future work, we already showcased a model that can identify handwritten digits, Character recognition and live handwriting analysis can be the next steps to take. The initial stage in the large field of artificial intelligence and computer vision is handwritten digit recognition. The experiment's findings show that CNN is one of the best performing algorithms in the digit recognition domain. More convolutional layers and hidden neurons can improve the accuracy of the output. It may entirely eliminate the necessity for typing. A fantastic approach to learn about neural networks and to construct more sophisticated deep learning methods is through the challenge of digit recognition. We intend to create a system for real-time handwritten text recognition in the future.