**Handwritten Text Recognition using Machine Learning Techniques**

Submitted in partial fulfillment of the requirements of the degree

**BACHELOR OF ENGINEERING** IN **COMPUTER ENGINEERING**

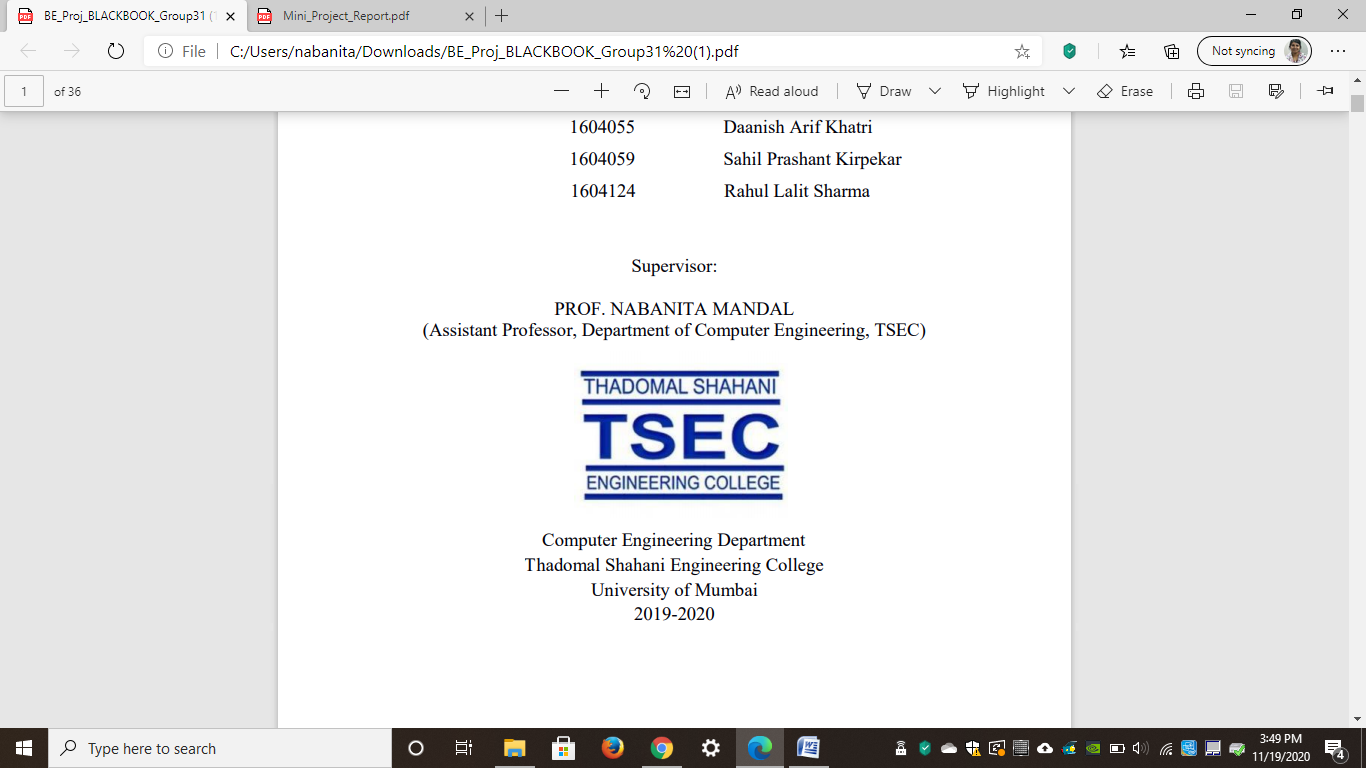
By

**Nishita Matlani 1902096 Shanvi Mehta 1902100**

**Isha Mehta 1902099**

Supervisor

Dr. Jayant Gadge



**Computer Engineering Department**

**Thadomal Shahani Engineering College**

**Bandra(w), Mumbai - 400 050**

**University of Mumbai**

**(AY 2021-22)**

# CERTIFICATE

This is to certify that the Mini Project entitled “**Handwritten Text Recognition using Machine Learning Techniques”** is a bonafide work of **Shanvi Mehta (1902100), Nishita Matlani (1902096), Isha Mehta (1902099)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **“Bachelor of Engineering”** in **“Computer Engineering”.**

### Dr. Jayant Gadge

Supervisor

### (Prof. ) (Prof. )

Head of Department Principal

# Mini Project Approval

## This Mini Project entitled **Handwritten Text Recognition using Machine Learning Techniques”** by **Shanvi Mehta (1902100), Nishita Matlani (1902096), Isha Mehta (1902099)** is approved for the degree of **Bachelor of Engineering** in **Computer Engineering.**

**Examiners**

**1………………………………………**

(Internal Examiner Name & Sign)

### 2…………………………………………

(External Examiner name & Sign)

Date: Place:

# Contents

### Abstract ii

### Acknowledgments iii

### List of images iv

### Introduction 1

* 1. Introduction
  2. Motivation
  3. Problem Statement & Objectives
  4. Organization of the Report

### Literature Survey 11

* 1. Survey of Existing System
  2. Limitation Existing system or research gap
  3. Mini Project Contribution

### Proposed System (e.g., New Approach of Data Summarization) 18

* 1. Introduction
  2. Architecture/ Framework
  3. Algorithm and Process Design
  4. Details of Hardware & Software
  5. Experiment and Results
  6. Conclusion and Future work.

### References 32

# Abstract

Handwritten text recognition (HTR) is the process of detecting and converting characters from photographs, papers, and other sources into a machine-readable format for further processing. It's still difficult to recognize intricately formed compound handwritten symbols accurately. By learning discriminating qualities from enormous volumes of raw data, recent breakthroughs in convolutional neural networks (CNN) have made significant progress in HTR. The characters from a test dataset are recognized using CNN in this article. The major goal of this project is to look at CNN's capacity to detect characters from an image collection, as well as the recognition accuracy during training and testing. CNN distinguishes the characters by examining their shapes and contrasting the qualities that distinguish them.

# Acknowledgement

We would like to express our gratitude and thanks to **Dr. Jayant Gadge** for her valuable guidance and help. We are indebted for her guidance and constant supervision as well as for providing necessary information regarding the project. We would like to express our greatest appreciation to our principal **Dr. G.T. Thampi** and head of the department **Dr. Tanuja Sarode** for their encouragement and tremendous support. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of the project.

Nishita Matlani

Shanvi Mehta

Isha Mehta

# List of diagrams

Name page no

Text recognition process block diagram 13

Text recognition architecture 14

Landing page 16

Output shown on the screen 16

# 

# Chapter 1

# Introduction

This chapter explains the aim, objectives and scope of the proposed system.

## **Introduction**

Handwriting recognition is a computer technology or feature that receives and interprets comprehensible handwriting data from sources such as paper documents, touch screens, and photographs. It is a capacity of a computer to receive and analyses understandable handwriting data from sources

It is one of area pattern recognition. The goal of pattern recognition is to categories or classify data or objects into one of several classes or categories.

The input is usually in the form of an image such as a picture of handwritten text that is fed to a pattern-recognition software. The purpose is to accurately recognize input letters or images, which will subsequently be examined by numerous automated process systems. We continue to investigate the challenge of identifying handwritten text and converting handwritten material to digital representation.

## **Motivation**

## Much work has been done in the area of ​​character recognition, but less work has been done to analyze a complete document. Recognizing the text of a document is useful in a variety of applications, such as reading doctor's prescriptions, banker's checks, and other official documents. It is also used in detective and police departments in applications such as handwriting-based personal identification and identification of genuine documents from forged documents. Nevertheless, it's a crucial problem to solve for multiple industries like healthcare, insurance and banking.

## The digitalization of patient prescriptions is a serious issue in the healthcare/pharmaceutical business. Every day, Roche processes millions of petabytes of medical PDFs. Patient enrollment and form digitalization are two further areas where handwritten text recognition has a significant influence. Hospitals/pharmaceuticals can dramatically improve user experience by integrating handwriting recognition to their toolbox of services.

## **Problem Statement & Objectives**

The purpose is to accurately recognize input letters or images, which will subsequently be examined by numerous automated process systems. We continue to investigate the challenge of identifying handwritten text and converting handwritten material to digital representation. We took on the challenge of identifying the visual of any handwritten word, whether it was in cursive or block script, for this research.

It is difficult to store, access, and analyze physical data in an effective manner. It is necessary to update manually, and work is required to maintain correct data structure. For a long time, we have experienced serious data loss as a result of the old way of data storage.

* 1. **Organization of the Report**

This report consists of three chapters. The first chapter deals with introduction of the topic, problem statement, motivation behind the topic and objectives. The second chapter is the Literature Survey. It includes all the research work done related to this topic. All information related to study of existing systems as well as learning of new tools is mentioned in this chapter. The third chapter is about the proposed system which is used in this project. The block diagram, techniques used, hardware and software used screenshots of the project are presented in this chapter. All the documents related to development of this project are mentioned in References

# Chapter 2

**Literature Survey**

This chapter explains the concepts used in this project, study of existing system and contribution of this project

**2.1 Survey of Existing System**

<https://www.concordia.ca/research/cenparmi/publications.html?utm_source=redirect&utm_campaign=publications.html>

https://data-flair.training/blogs/handwritten-character-recognition-neural-network/

To recognize numbers, it uses a feature-based method. They utilized a statistical approach to give weights to each attribute and then used those weights to calculate scores. It also recognizes handwritten words using a feature-based technique. Because words may be italicized but numbers cannot, our technique differs from theirs.

**2.2 Limitation of existing system**

## A Word can not always be broken down into characters. the collection of alphabetic traits is bigger than the set of numerical features. It is easier to identify the numeric features in the image compared to the alphabetic features because of the complexity in different handwriting with includes cursive handwriting.

## **2.3 Mini Project Contribution**

People write checks on a regular basis, and checks continue to play an important role in most cashless transactions. Current check processing technologies in many developing countries include bank tellers reading and manually entering information about checks and validating data such as signatures and dates. Banks process a large number of checks every day, so using a handwritten text recognition system can save you money and labor costs.

Large insurers receive about 20 million documents daily, and delays in claim processing can have a significant impact on the organization. Claim documents can contain a variety of handwriting styles, and fully manual automation of claim processing can significantly slow down the pipeline. People working in these sectors can use the system to ease their work and reduces the time for doing the same.

# 

# Chapter 3

**Proposed System**

This chapter consists of detailed description about the methodology used, the hardware and software components, the tools used and also the screenshots of the project

**3.1 Introduction**

The project is built using python and its different framework which are explained below:

Python: Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small- and large-scale projects.

TensorFlow: TensorFlow makes it easy for beginners and experts to create machine learning models for desktop, mobile, web, and cloud. TensorFlow is a Python library for fast numerical computing created and released by Google.

It is a foundation library that can be used to create Deep Learning models directly or by using wrapper libraries that simplify the process built on top of [TensorFlow](https://machinelearningmastery.com/tensorflow-tutorial-deep-learning-with-tf-keras/).

OpenCV: OpenCV-Python is a library of Python bindings designed to solve computer vision problems.

NumPy:  NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, Fourier transform, and matrices.

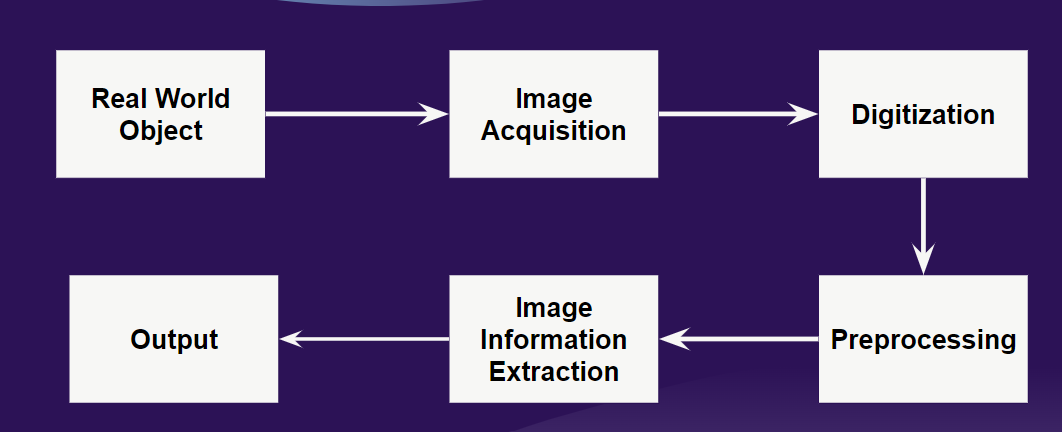
Neural networks used for the implementation are listed below:

* **Convolutional NN (CNN) layers:** In a deep CNN, convolutional layers are where filters are applied to the original picture or other feature maps. In the network, here is where the majority of the user-specified parameters are located.
* **Recurrent NN (RNN) layers:** The RNN (Recurrent Neural Network) is designed to recognise patterns and sequences in applications such as speech, handwriting, and text. The input sequence is taken into account by the RNN's logic. In hidden units, each prior input data is kept in a state vector, which is then used to compute the outputs.

**LSTM:** LSTM is an RNN technique that prioritizes feedback links for use as general purpose computers. The LSTM network is made up of cells, and the main function of the cell is to determine what to remember and what to exclude from memory. LSTM is state-of-the-art for many difficult problems as it is effective in capturing long-term time dependencies without suffering from the optimization hurdles that plague simple recurrent networks (SRNs). Has been used to advance. This includes, among other things, handwriting recognition and generation, speech acoustic modeling, speech synthesis, protein secondary structure prediction, and analysis of audio and video data.

* **Connectionist Temporal Classification (CTC) layer:** Connectionist speed classification loss or CTC loss is designed for tasks that require alignment between sequences but are difficult to align. Calculates the loss between a continuous (non-segmented) time series and a target sequence.

## **3.2 Architecture**



(3.2.1) Text recognition process block diagram

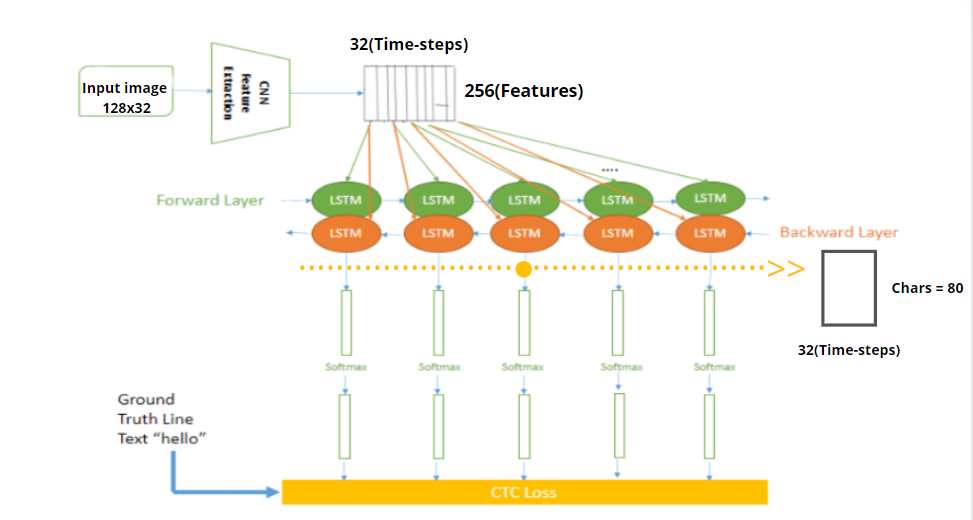
First, we will take image as input then we will retrieve the image from the hardware source like camera, etc. Then image will convert (analog to digital) into a format that is understood by the computer. Image acquisition is where the analogue image is converted into digital form and then the digitization is done on the image where the conversion takes place and the output formed is again in the digital form that can be processed by a computer

Then the digital image undergoes preprocessing where the steps are taken to format the images before they are used for model training

This is followed by preprocessing, where an image is subjected to various operations like, normalization, smoothing, and skeletonization. The result of this preprocessing can be given as an input to feature generation. Segmentation of an image is done to isolate the characters of an image into different sub-images.

## **3.3 Algorithm and Process Design**

Write some few lines for each of the steps given below. Mention what you have done in each and every step.



(3.3.1) Text recognition architecture

A word cannot always be broken down into letters. As a result, a continuous procedure should be used to match the set of features to the database, taking into consideration the order in which new features are presented and old features are removed. Furthermore, the collection of alphabetic traits is bigger than the set of numerical features.

NN i.e., Neural networks are used in this system. It consists of convolutional NN (CNN) layers, recurrent NN (RNN) layers and a final Connectionist Temporal Classification (CTC) layer.

First the input image i.e., the analog image is fed into the CNN layers. These layers are trained to extract relevant features from the image. Each layer consists of three operations.

The feature sequence contains 256 features per time-step, the RNN propagates relevant information through this sequence. The popular Long Short-Term Memory (LSTM) implementation of RNNs is used.

while training the NN, the CTC is given the RNN output matrix and the ground truth text and it computes the loss value. Model is trained using the IAM dataset.

**Altogether there are 3 steps:**

1. Multi-scale feature Extraction   
   → Convolutional Neural Network 5 Layers
2. Sequence Labeling (BLSTM-CTC)   
   → Recurrent Neural Network (2 layers of LSTM) with CTC
3. Transcription   
   → Decoding the output of the RNN (CTC decode)

The project is made using TensorFlow, OpenCV and NumPy

**3.4 Details of Hardware & Software**

Hardware: 8gb/512gb

i5 10th gen

Software: Python

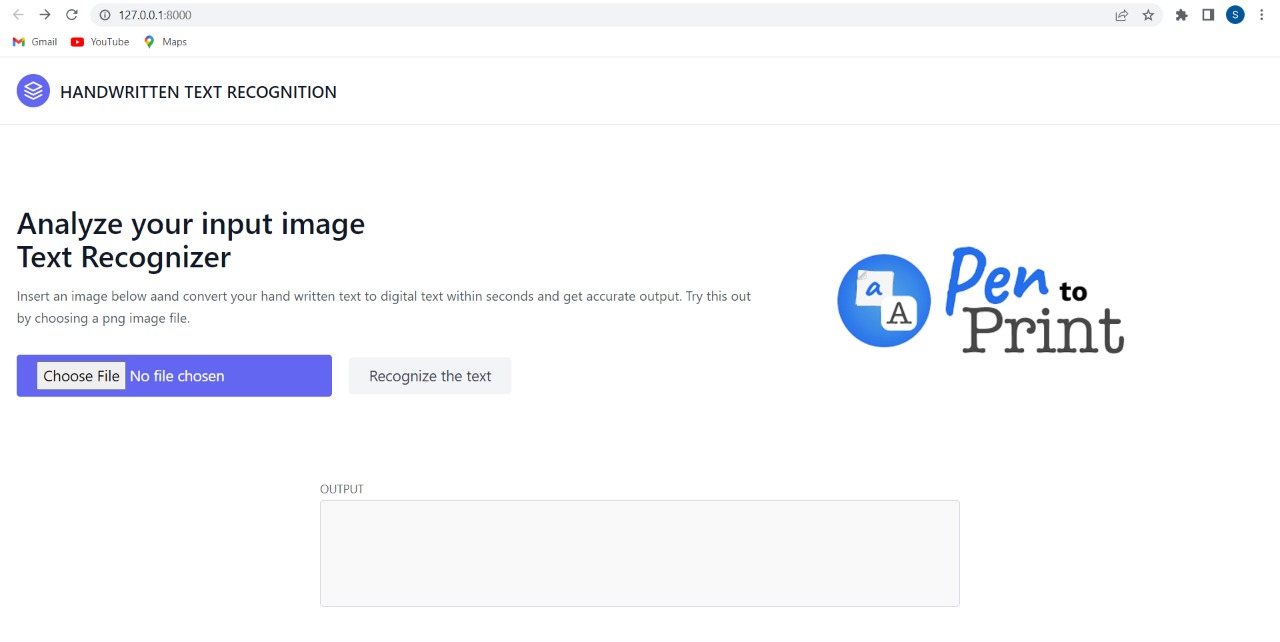
TensorFlow

OpenCV

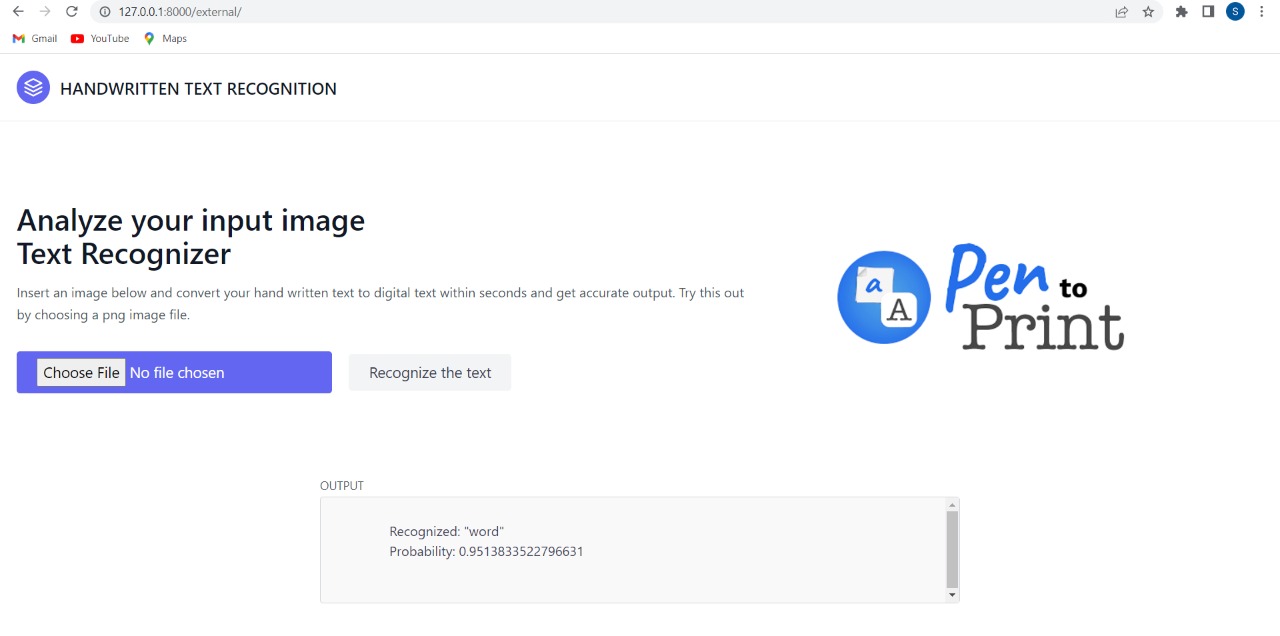
NumPy

Django

**3.5 Results**

Images below show the UI of the Text Recognizer

(3.5.1) Landing page

The above image is the landing page with an option to choose file of the input image and the output box will give the converted digital text from the image

(3.5.2) Output shown in the screen

The output for the image consisting of the word “word” is shown above.

**3.6 Conclusion and Future Work**

Conclusion: A neural network (NN) that can detect the text in a picture. An NN produces a character probability matrix by combining 5 CNN layers and 2 RNN layers. This matrix may be used to calculate CTC loss or to decode CTC. There is a TF implementation included, as well as some useful code.

Future work: Recognition accuracy can be improved. Addition of Voice recognition along with voice as an output for the text recognition.

**References**

<https://repositum.tuwien.at/retrieve/10807>

<https://repositum.tuwien.at/retrieve/1835>

<https://fki.tic.heia-fr.ch/databases/iam-handwriting-database>