OFFLINE HANDWRITTEN TEXT RECOGNITION
PROJECT ON MACHINE LEARNING

Prachi Goyal ,Anna Simon ,Krittika Chaturvedi

August 31, 2020

INTRODUCTION

Despite the abundance of technological writing tools, many people still choose to take their notes traditionally:with pen and paper. However, there are drawbacks to handwriting text such as it's difficult to store and access, search through them and to share them with others in an efficient manner.

Handwritten text recognition is the ability of a machine to receive and interpret the handwritten input from an external source like image.

This project aims to style a system that can efficiently recognize the actual handwritten text to convert handwritten text into the digital format.

OBJECTIVE

- To provide an easy user interface to input images.
- System should be able to pre-process the given input to suppress the background.
- System should detect text regions present in the image.
- System should retrieve text present in the image and display them to user.

CHALLENGES

- Huge variability and ambiguity of strokes from person to person .
- Handwriting style of an individual person also varies time to time and is inconsistent.
- Cursive handwriting makes separation and recognition of characters challenging.
- Collecting a good labelled dataset to learn is not cheap compared to synthetic data.

TECHNOLOGY STACK

- Python 3
- Tensorflow
- Keras
- Window 10 and Ubuntu

# PROJECT FLOW
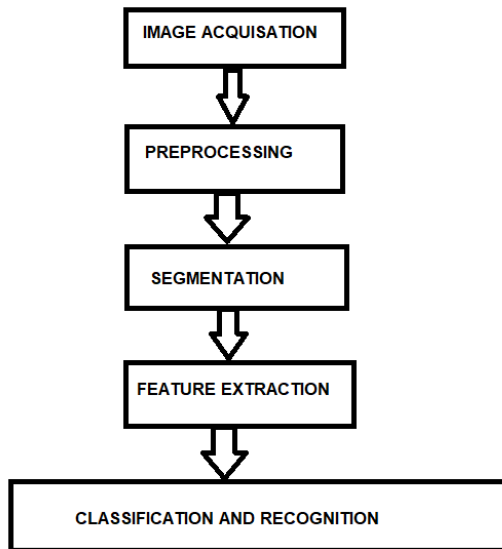
TRAINING MODEL : IAM DATASET

The IAM Handwriting Database is structured as follows:

- 657 writers contributed samples of their handwriting .
- 1,539 pages of scanned text.
- 5,685 isolated and labeled sentences .
- 13,353 isolated and labeled text lines.
- 115,320 isolated and labeled words .

MODEL OVERVIEW

- The input image is a gray-value image and has a size of 128x32 .
- 5 CNN layers map the input image to a feature sequence of size 32x256 .
- 2 LSTM layers with 256 units propagate information through the sequence and map the sequence to a matrix of size 32x80.
- While training the NN, the CTC is given the RNN output matrix and also the ground truth text and it computes the loss value.
- While inferring, the CTC is just given the matrix and it decodes it into the ultimate text.
- Both the bottom truth text and also the recognized text are often at the most 32 characters long.
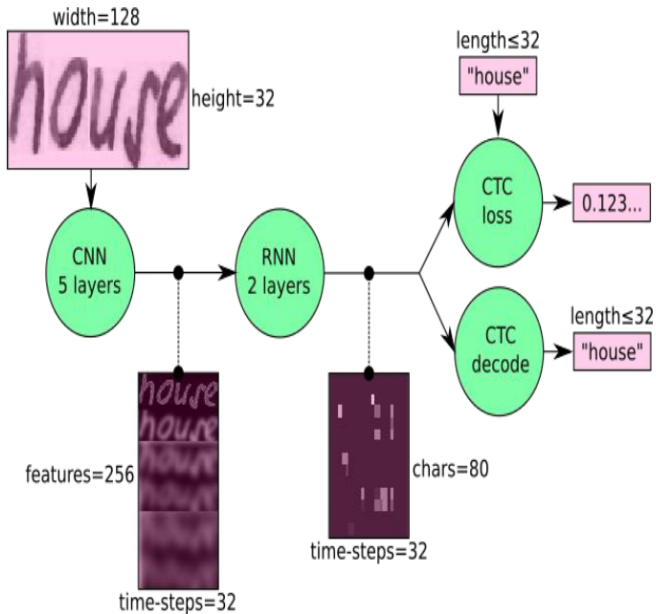
Result

Our model is trained and tested on 50,000 and 5000 data entries respectively producing accuracy of 66.3% with learning rate of 0.01 .

LEARNINGS

- Got introduced to the field of Machine Learning .
- Got conceptual understanding of various neural networks such as CNN,RNN .
- Learned implementing various ML models using Tensorflow and Keras .
- Learned to use and collaborate using Gitlab .
- Learned to use Latex .

REFERENCE PAPER

Reference Paper Link :
https://drive.google.com/file/d/
1lIg77GrQXWQ2YIyxkfAMjMGgdE0xXz-v/view?usp=sharing

Thank You