HANDWRITTEN TEXT RECOGNITION MODEL

BY

RADHA RANI KOMMINENI

PS ID: 2156574

July 27, 2022

INTRODUCTION

Background

Deep learning has been one of the booming tech stack one can use to build and deploy models using several tools like, tensorflow, keras, pytorch frameworks so on and so forth. In contrast against machine learning, Deep learning acts like a complex system where its internal workings or tasks are not so understood, that is, one does not have to do several acts on a dataset, for example classification and feature extraction which is something one has to do when working with Machine learning.

In machine learning, the phase of feature extraction is so tiresome as one has to do some notes on different features in a dataset. For example, one doing building a facial recognition model in machine learning, has to extract features from the face images or non-image dataset. In such scenario features may include but not limited to, age, color, face edges face shapes, et-cetera. After does the extraction he or she is required to pass the features in a machine learning classifier and perform some classification on its own. However, as opposed to deep learning, such scenario is simple and precise. Why? This is because deep learning doesn't consider some tasks being done by the modeler since deep learning models has the ability to perform the classification and all that other workings on its own. Therefore, in that scenario, one has to consider defining faces or non-faces from the images or dataset. So then, when passed to model it has the ability to compile and train without any glitch.

What is Deep Learning in simple terms? Well, deep learning is a subset of Machine learning. It is based on Neural network layers modelling, that is, it has to contain three or more neural network layers depending on your architecting. The neural networks are said to be a replica of human brain, well, they try to simulate behaviors of a human brain. This ability is said to be much big enough than human brain abilities hence enabling them to learn from a large couple of datasets.

Deep Learning applications in most of the industries, has been the norm for driving the industries in adapting the new tech. Deep learning has been widely used in most of the departments or industries like

automotive engineering, medic industries, in agriculture and as well as other industries like network providers. Deep learning is well seen to be in use in automated driving, object detection, human detection and pedestrian detections which in turn helps to solve some good problems and bad problems as well.

Talking of deep learning, we are going to create a handwriting text recognition model. This is all about making a model to help us in determining what and what word or text is in a picture.

LITERATURE SURVEY

Introduction

The world is widely changing at an amazing rate and internet technology ballooning is hence considered to be the main driver for these changes. This chapter theoretically reviews on a handwritten text recognition model, the existing look like model, and the major gaps that exists on other researcher's research work as well as the methodologies they have used in their research work.

SURVEY

Character Recognition is nothing but machine learning simulation of human reading, Gaurav Kumar.

The existing current system to research on is known as character recognition model. The model uses various preprocessing techniques for character detection. The model used a dataset containing images containing a handwritten texts and images of different kinds. The images were colored and had some complicated background and as well as several with different light intensity.

The offline character recognition is proposed by Sing, a diagonal extraction feature. It supported an ANN model. The model was simply using a multi-layer feed forward neural network. There are several approaches to use while assembling such model, we can for example a 54 or 69 feature-based model. During a paper proposed by A. Brakensiek, J. Rottland, A. Kosmala, J. Rigoll, reviews a system for offline handwriting recognition which was described. The system used a model that used a Hidden Markov Technique (HMM) model that used hybrid and discrete modelling technique.

The above described models have however had some cons in doing predictions. They tend to fail. Many models therefore, don't seem to be so accurate and successful in predictions of a long and huge couple of text data. However, with the deep learning capabilities of handling a large amount of data, It is therefore evident that Neural networks can be reliable and efficient in predictions.

Research gaps involved in the above type of model's research is that, a lot of preprocessing doesn't have proper and healthy handling, of data such as, feature extraction and classification of data when using machine learning models, and layers feeding is mishandled when dealing with a lot of deep learning modelling.

PROPOSED WORK

This is simply a show of how the model proposed and how the project is all about. Most of the OCR systems out there usually use image captures and images with texts in subject to detect a text via a camera. In this project, we are going to have a CNN neural network model to help us detect and recognize a character and form a word if the model is subjected into real life scenario.

The Model will use a dataset consisting A to Z alphabet written hand written images. The pixel size is said to be 28 by 28 pixels. The images is set to 20 by 20 box-pixeled inside. The dataset contains a total of 372451 images.

Dataframe Info is:

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 372450 entries, 0 to 372449

Columns: 785 entries, 0 to 0.648

dtypes: float32(785)

memory usage: 1.1 GB

Since input layers are mostly small, the model is often runnable using a CPU and better in GPU. The flowchart below helped in the technical part of modelling:

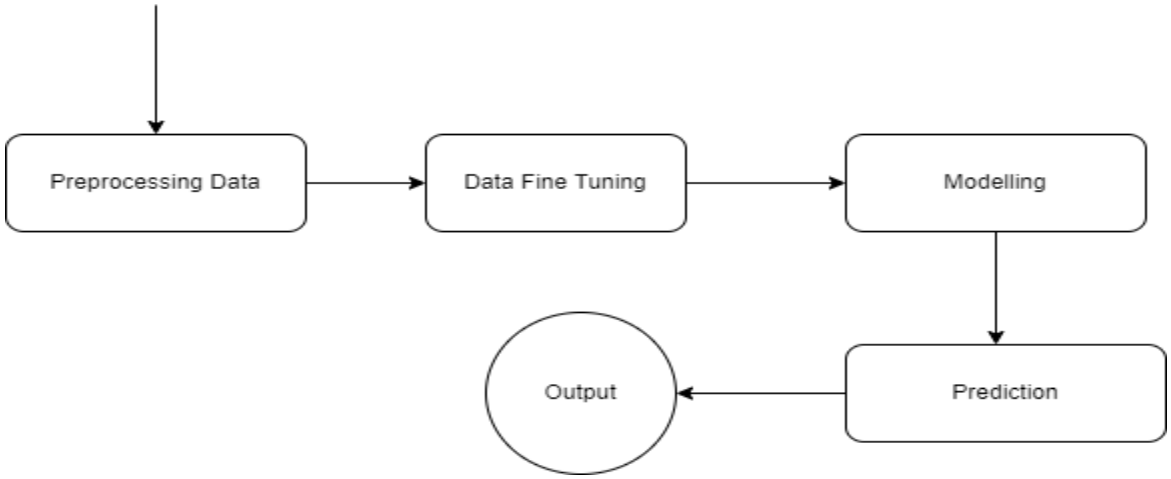


Fig 1.0. Flowchart to illustrate the modelling of the Handwritten text prediction detection model.

EMPIRICAL ANALYSIS

Introduction

This part simply helps researchers to understand and validate the idea proposed. The part explains more in detail of how the modelling was done to have the desired output proposed.

Get Started

Getting started is simply the way of knowing our model and what constitutes our model.

The modelling was done and tested in a google colab notebook. However, It could be runnable on a jupiter notebook as well.

To get started we shall have a look at the following list of basic tasks done in deep learning modelling.

1. Data Preprocessing
2. Data Fine Tuning
3. Model Crafting
4. Performing Predictions

Data Preprocessing

This happens after Modelling environment preparation.

This part involves data reading. Reads may include, Data frame, dataset information, description, shape as well as type of data it is.

The dataset used here was having a shape of set; **372450 x 785**

Data Fine Tuning and Manipulation

This stage is all about doing some manipulations on our data, as well as preparing the data to a data understandable by our model.

To start with fine tuning, we have to do a splitting of our dataset into two subsets, X and Y.

Several tasks that was done here includes but not limited to:

Reshaping Dataset – This helps to get a fine dataset our model could understand. Reshaping our dataset, helps in converting our data into imagery format. X and Y dataset had to be split further into two. This helps us to get training dataset and test dataset. That is, to have x train split and x test split, y train split and y test split.

Alphabet mapping - This involves mapping of the alphabets on characters in the data frame by replacing floating key values with the alphabets. These helps well by making the characters in the data frame get labeled.

Alphabet mapping was plotted as below:

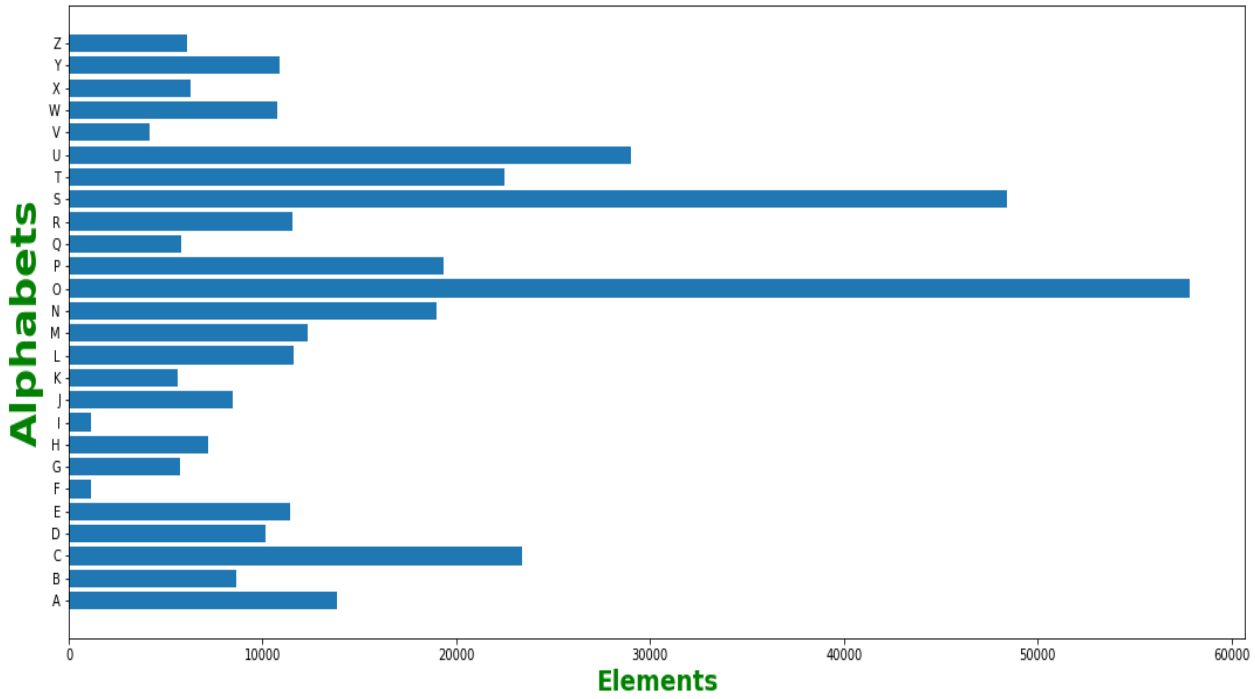
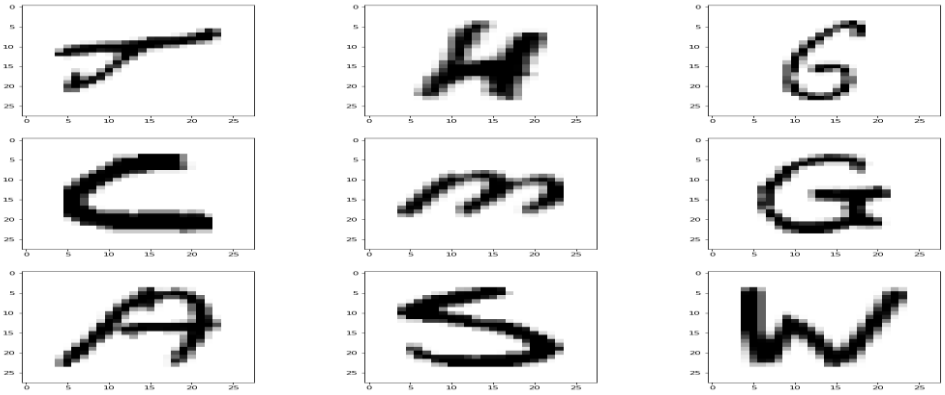


Fig 1.2. Alphabet Distribution in the dataset using a dynamic Alphabet dictionary

Shuffling of dataset values and Threshold application – This helps well in making optimizable model. This helps in retaining representative values in our dataset splits. Threshold application in turn helps in improving our model. By default, threshold is always set to be 5. To apply a threshold, depends on the modeler to decide what threshold best suits the model. It is advisable in most cases to apply the default value (5).

After shuffling, the alphabet representation was plotted as shown below.



Converting Float values in our dataset to Categorical – This helps in getting rid of float values, and as well as maintain uniform data.

Model Crafting

This refers to building our model. Model is based on neural network layering. It is a sequential model which gets us to know tensorflow and its capability of deep learning. The model was trained on an Adam optimizer and against accuracy metrics.

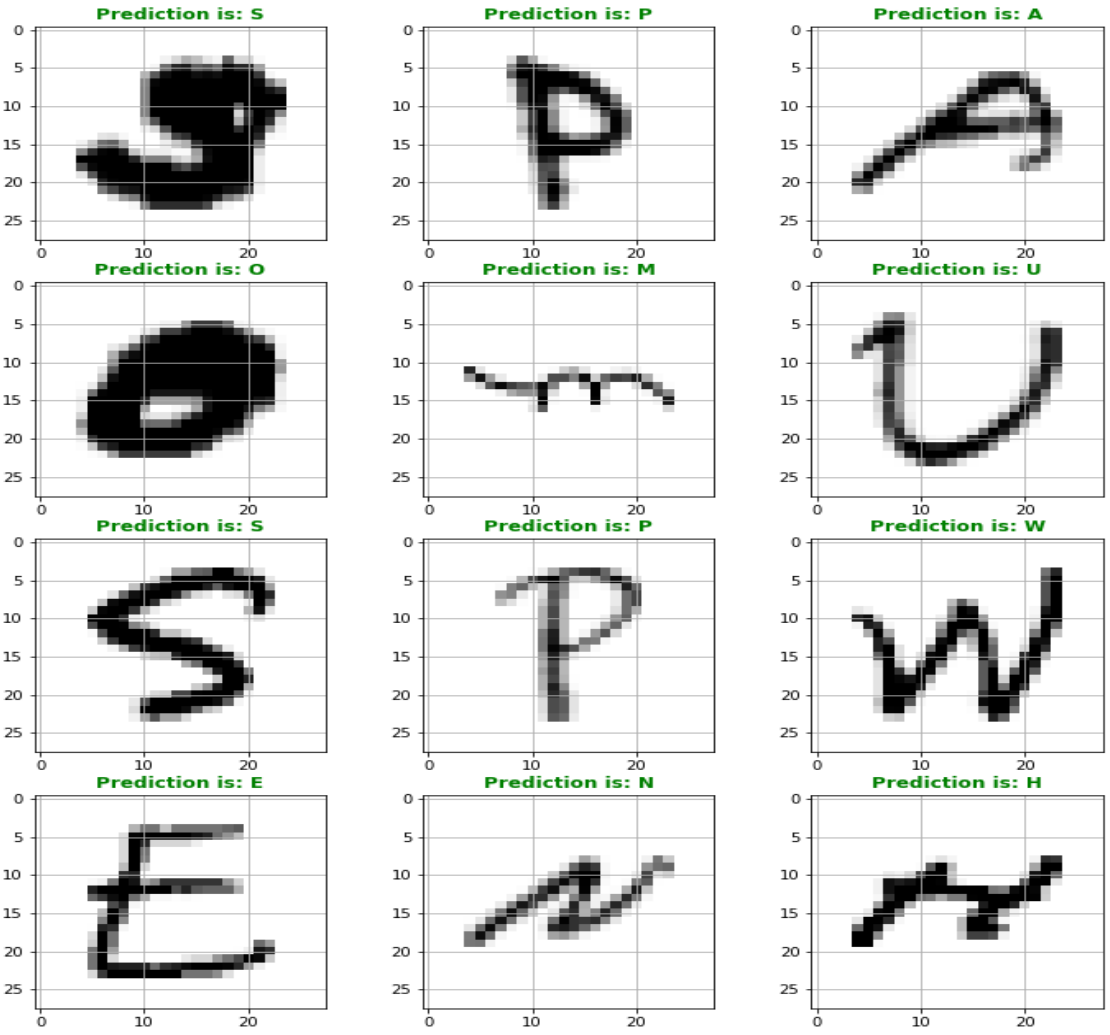
Our model had 10 layers which constituted of;

1. **Conv2D(filters = 32, kernel_size = (3, 3), activation = 'relu', input_shape = (28, 2 8, 1)),**
2. **MaxPool2D (pool_size = (2, 2), strides = 2),**
3. **Conv2D (filters = 64, kernel_size = (3, 3), activation = 'relu', padding = 'same'),**
4. **MaxPool2D (pool_size = (2, 2), strides = 2),**
5. **Conv2D (filters = 128, kernel_size = (3, 3), activation = 'relu', padding = 'valid'),**
6. **MaxPool2D (pool_size = (2, 2), strides = 2),**
7. **Flatten (),**
8. **Dense (64, activation = "relu"),**
9. **Dense (128, activation = "relu"),**
10. **Dense (26, activation = "softmax")**

Performing Predictions.

Performance of our model seems so perfect and accurate to even detect a long couple of text.

The following diagram below how the predictions were made by our model.



References

Anil. K. Jain and Torfinn Taxt, "Feature extraction methods for character recognition-A Survey," Pattern Recognition, vol. 29, no. 4, pp. 641-662, 1996

Pal, U. and B.B. Chaudhuri, "Indian script character recognition: A survey," Pattern Recognition, vol. 37, no.9, pp. 1887-1899, 2004.

[2] T. Bluche, J. Louradour, and R. Messina Scan Attend, and Read: End-to-End Handwritten Paragraph Recognition with MDLSTM Attention (2016)