ECS 271: Project Proposal

Shivam Pandey | Michael Yang | Aniket Banginwar

February 9, 2022

1 Problem Statement and Motivation

The research and development on handwritten sentence recognition are to a large degree motivated by many application areas, such as data acquisition in banks, text-voice conversion, security, etc. Many applications demand very high recognition accuracy and reliability. *Google's Tesseract*, despite being developed by a tech beast, its output will have poor quality if the input images are not preprocessed to suit it. This motivates us to research and develop a handwritten sentence recognizer with a promising accuracy.

2 Design and Technical Approach

This project focuses on recognizing a single-line handwritten English sentence using Deep Learning methods such as *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN) using *Long Short Term Memory* (LSTM) units and *Language Model* (LM). The input given to the recognition system is an image of a single-line handwritten English sentence, which is processed through a series of aforementioned deep learning techniques to finally output the sentence in machine-printed format. The entire recognition algorithm can be broadly broken down into three major parts as follows -

2.1 Line Segmentation into Words

Firstly, our algorithm involves segmenting the input sentence image into a sequence of words' images. We plan to approach this problem using *Kadane's algorithm* and *k-means clustering*. We obtain the images of individual words which will be fed into our neural network model.

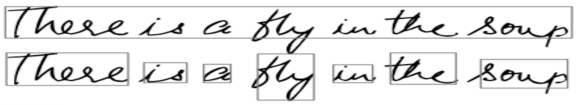

Figure 1: Line Segmentation into Words

2.2 Handwritten Word Recognition

Secondly, the task is to recognise each of the word from the sequence of images. We will build upon an existing Neural Network (NN) architecture [2] which will be trained on word-images from the *IAM* dataset [1]. This NN will consist of CNN layers, RNN layers with LSTM units and a final Connectionist Temporal Classification (CTC) layer. The NN will output a character-probability matrix. This matrix will either used for CTC loss calculation or for CTC decoding. Eventually, we will obtain the machine-printed words from the images provided as input to the NN.

2.3 Improving Accuracy of Recognized Sentence

Thirdly, we deal with improving the accuracy of the NN output using some techniques such as LM, deslanting, text-correction (i.e. search for the most similar word, if the recognized word is not in dictionary), etc. To our understanding, this is a subtle part of the algorithm.

3 Contributions of Team Members

The *three* parts of this project would be distributed in the following manner -

- Aniket Banginwar - Line Segmentation into Word
- Shivam Pandey - Handwritten Word Recognition
- Michael Yang - Improving Accuracy of Recognized Sentence

References

[1] IAM. URL: https://fki.tic.heia-fr.ch/databases/iam-handwriting-database.

[2] Harald Scheidl. URL: https://towardsdatascience.com/build-a-handwritten-text-recognition-system-using-tensorflow-2326a3487cd5.