

Handwritten English Word Recognition

Vishakha Satpute
Computer Engineering
Shree L.R. Tiwari College of
Engineering
Thane, India
vishakhasatpute17@gmail.com

Nikita Dane
Computer Engineering
Shree L.R. Tiwari College of
Engineering
Thane, India
nikitadane.7.99@gmail.com

Sonali Mamidipaka
Computer Engineering
Shree L.R. Tiwari College of
Engineering
Thane, India
sonalimamidipaka@gmail.com

Abstract— Handwritten English Word Recognition is evolving fields within computer vision. Humans are able identify the handwritten text accurately due to strong visual cortex system but it is difficult for a computer to identify this text. In Handwritten English Word Recognition, we give an input image as input, preprocess it, apply classification schema, and train the system. In this approach, we are using Convolution Neural Network, Recurrent Neural Network and Connectionist Temporal Classification

Keywords— Neural Network, Convolution Neural Network, Recurrent Neural Network, Connectionist Temporal Classification, LSTM.

I. INTRODUCTION

Handwritten Text Recognition is a field which has lots of applications such as form-based applications, Bank cheques, etc. The handwritten English word recognition consist of following steps like image acquisition, preprocessing of image, segmentation, classification and post processing. Handwriting English Word recognition is a software which has the capability of receiving the handwritten input image and convert it into digital text.

We use Convolutional Neural Network and Recurrent Neural Network to train a model that can correctly classify words and for the latter, we use Long Short Term Memory networks (LSTM) with convolution to create bounding boxes for each character, the system is trained. We are also using Connectionist Temporal Classification. This software takes input as the image of a handwritten word and converts it into its corresponding digital text

II. MOTIVATION

The development in the field of handwritten text recognition has many and various application areas such as data acquisition in bank, security, text-voice conversion, automated postal address etc. As the expenses of scanners is decreasing we have seen increases demand for handwriting text recognition system. Some handwritten recognition system have been published in the market. The difficulty occurs in recognizing bad handwritings. Moreover, many application demand high accuracy in recognizing the text.

III. PROPOSED SYSTEM

The proposed system of handwritten English word recognition consist of image acquisition, pre-processing, segmentation, classification and recognition, post processing. Figure below is the proposed system,

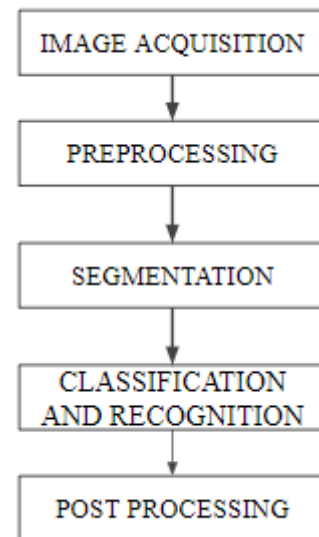


Fig 1. Proposed System

A. Image Acquisition

In Image acquisition, the system takes in a scanned image as an input image.

B. Preprocessing

The pre-processing enhances the image rendering and the output of preprocessing is given as the input to segmentation.

C. Segmentation

In the segmentation stage, an image consisting of sequence of characters is disintegrated into individual character. Here CNN algorithm is used.

D. Classification and Recognition

The classification stage is the supervisory stage of the Word Recognition system. It analysis the input and recognizes the word present in the input image.

E. Post Processing

In Post-processing stage, the system prints the recognized word. In this stage, CTC algorithm is used.

IV. METHODOLOGY

We are using Neural Network for developing this system. It consist of convolution neural network, recurrent neural network and connectionist temporal classification.

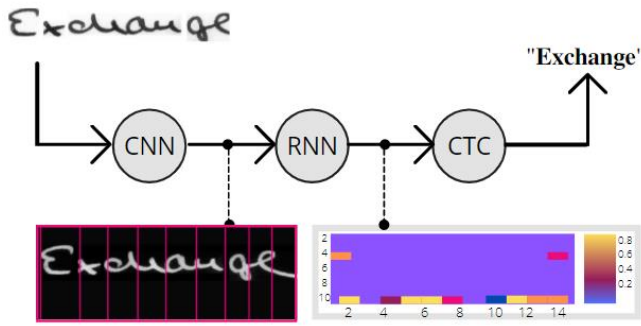


Figure 2. Overview of Handwritten English Word recognition

A. Convolution Neural Network

Whenever a human see an object, we immediately recognize the corresponding object. Visual Cortex System of humans are so well trained that they can recognize an image correctly even if it is slightly blurred but this is difficult for a machine to recognize a blurred image. After years of research, French scientist “Yann Lecum” discovered an algorithm so that systems can recognize an object, this algorithm is known as CNN.

In CNN, there are two main categories, Convolution and Pooling. The image is given as a input to CNN layer. There are five CNN layers used. These layers collect the information regarding the image by scanning the group of pixels with the help of a filter. These filter can be of various size. In this project we apply filter of size 5x5 in first two layers of CNN and in the last 3 layers of CNN we use filter of size 3x3. We apply rectified linear activation function. Also, in pooling layer image regions is analysed and it outputs a downsized version of the input.

B. Recurrent Neural Network

RNN works on the principle of saving the output of the layer and giving this back as the input in order to predict the output of layer. It is the only algorithm with internal memory. Because of interior memory RNN is able to retain important factors about input they acknowledged, which allows them to be specific in forecasting what comes succeeding. That is why they're the preferred algorithm for sequential data like time series, text the autocorrect we use during texting

So here we are using LSTM implementation of RNN as it is capable of learning long term dependencies by remembering information for long periods in the default behaviour.

So the feature sequence contain 256 features per time-step. The RNN output is converted into a matrix of size 32x80.

C. Connectionist Temporal Classification

CTC layer is used to calculate the loss for our stochastic gradient descent to train our LSTM. The idea is to try to align the prediction to the target. The output and the ground truth text of RNN layer is given as input to the CTC layer. CTC layer decode it into recognized text.

CTC works on three concepts:

- Encoding the text:

- Loss calculation:
- Decoding

We use best path algorithm. Best path algorithm consist of following two steps:

1. Calculate the best path by considering the character with maximum probability at every time-step. This step involves removing blanks and duplicate characters which results in the actual next
2. Best path decoding is used as only an approximation.

V. IMPLEMENTATION

The implementation consists of 4 modules:

1. SamplePreprocessor.py: It train the images from the IAM dataset for neural network.
2. DataLoader.py: DataLoader module reads the sample of image, place them into group and offers an iterator interface to travel through the image.
3. Model.py: Model module creates the implementation for CNN, RNN and CTC as described above, also provide GUI, manages the Tensorflow and provides an interface for training and inference.
4. Main.py: Main module implements the GUI which makes it easy for user to work with the software and also puts all the modules that are created above together for the recognition.

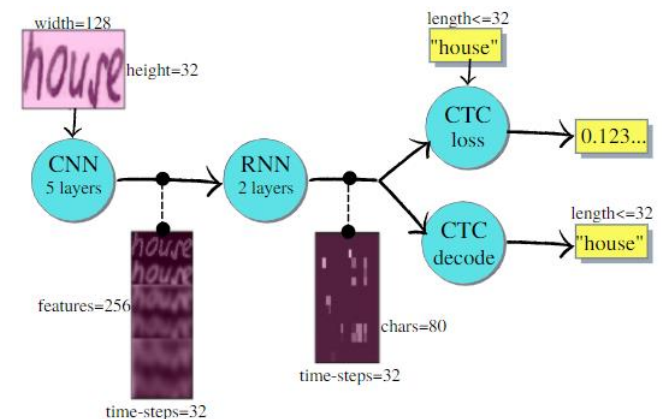


Figure 3. Overview of model.py module

VI. RESULT

The image below is the input given to the system which predicts the output by loading the modules explained above in implementation.



Figure 4



Figure 5

Figure 5 is the output given by the system when image in figure is given as input to the system.

VII. CONCLUSION AND FUTURE WORK

The project is developed by using neural network. This algorithm will provide effective result for the recognition. The project gives correct output for the words who have less noise. If we increase the training data or create our own dataset, we can get more accuracy.

Future Work:

In future we are planning to recognize a sentence or a complete document. For this we have to follow the below points:

- Increase input size of a Neural Network
- Add more CNN layers
- Replace LSTM by 2D-LSTM
- Increase dataset size by applying further transformation to the input image.

ACKNOWLEDGMENT

This work is partially supported by the grant from the Shree L.R. Tiwari College of Engineering. We thank Mrs. Dipali Bhole for her guidance in this project.

REFERENCES

- [1] Jayaram Kinthali, Kowshik Kotha, Kanithi Santosh Kumar and Jagilinki Jayalaxmi, "Handwritten Text Recognition using Deep Learning with TensorFlow", International Journal of Engineering Research & Technology (IJERT).
- [2] R. Plamondon and S.N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey. Pattern Analysis and Machine Intelligence", IEEE Transactions on 22(1):63-842000.
- [3] J.Pradeep, E.Srinivasan and S.Himavathi, "Neural Network based Handwritten Character Recognition system without feature extraction", International Conference on Computer, Communication and Electrical Technology – ICCET 2011.
- [4] Junqing Yang, Peng Ren and Xiaoxiao Kong, "Handwriting Text Recognition Based on Faster R-CNN"
- [5] Bala Mallikarjunarao Garlapati, Srinivasa Rao Chalamala, "A System for Handwritten and Printed Text Classification", 2017 UKSim-AMSS 19th International Conference on Modelling & Simulation.