Simple Optical Charecter Recognition for Handwritten Charecter

Farhad Hossain MD

Department of Computer Science and Engineering
American International University Bangladesh.

***Abstract*—Optical Character Recogniton is a part of Computer Vision which allows a device or computer to recognise about an given input character [1]. Throughout this project, the objective is to develop a Optical Character Recognition for handwriting input. From a Dataset [2] the training data for KNN Classifier [?] is loaded. KNN Classifier determines which group does a sample belongs to, so the unknown sample is recognized.**

I. Introduction

The objective is to take hand-writing as input and make the device or computer recognise all the character written in the input. To reduce the workload of typing, it may work as be a great initiative. The Charecter recognition process requires various steps. There are lots of established work on Handwritten charecter recognition. A simple processing using Morphological operations, image cropping and KNN classifier is used in this project. All the work is done using Matlab [3] software.

II. Methods

To recognize various steps is followed in this project. For KNN Classifier's Training data, at first from a image dateset where each image represents a single charecter, Traning data is loaded. Then the image is resized for memory optimization. This resized image is passed throgh a binary image converting function so that Morphological operation can be done. Morphological skeleton operation is done to obtain the skeleton of the charecter. Then the skeleton of the charecter is devided into four equal blocks. Total pixel value of each block is calculated and stored in a array. Finally, this total pixel value for each block added to training array as a training data. All the procedure is followed again for loading sample image and the processed unknown image is sent to KNN classifier as a sample data. KNN classifier determines the class of the unknown image. The output of the classifier is displeyed in matlab command window.

III. Resluts

From the dataset image 5 image is taken as sample where the charecter is '7','8','9','A' and 'B'.
Figure 1 shows the result obatained from the Matlab. The overall correction rate is 60%.

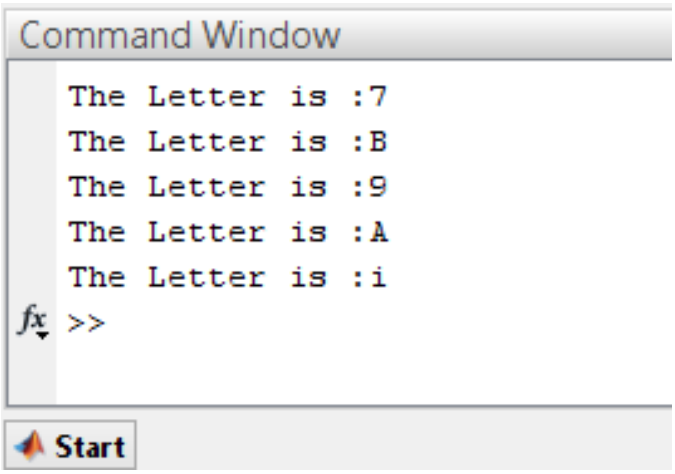

Fig. 1. Simple Optical Chareter recognition using KNN Classifier

IV. Conclusion

To optimize the time in this project a very samll amount of training data and only a simple feature extraction method is done, for which the corrrection rate is not optimum. With proper training data and feature extraction the correction rate will increase. Future work can be done with more feature extraction and training data.

References

[1] [Online]. Available: https://en.wikipedia.org/wiki/Handwriting_recognition
[2] [Online]. Available: http://www.ee.surrey.ac.uk/CVSSP/demos/chars74k/#download
[3] [Online]. Available: https://www.mathworks.com/products/matlab.html