Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 11, Number 7 (2018) pp. 547-554

http://www.ripublication.com

Optical and Handwritten Character Recognition System Using Segmentation and Neural Network

Suraj Saha[1]

*MCA (Student), VIT University, Vellore,
Tamil Nadu, India.*

[2]Dr. Mythili G.Y.

*Assistant Professor, School of Advanced Sciences,
VIT University, Vellore, Tamil Nadu, India.*

Abstract

Optical character recognition has been a hot topic for decades and has been continuously being developed and improved. The following paper describes how English alphabets can be distinguished from one another based on process of segmentation and how density of marked contours helps neural networks to analyze each data. Based on experiments conducted on four skeletalized English alphabets, the following article highlights a brief description of how errors and anomalies occur during the process of handwritten character recognition.

Keywords: Feed-Forward Neural Networks, Optical Character Recognition, Skeletalization, Binarization, Splitting of Characters

INTRODUCTION

The enlightenment of this paper is based on the survey research of [1]. Both Optical and Hand Character recognition has been one of the hot and challenging topics of this era and the research on this still continues. Optical characters are easier to recognize than handwritten ones as there is scope of variance of shapes and structures depending on the type of handwriting of a person. The aim is to develop a technique that recognizes both of them without compromising both space and time complexity of the system. In this paper we have used first four English alphabets for classification and used Neural Networks instead of classes of images in databases in order to lower time complexity of the algorithm as stated in [2].

Overview The algorithm starts with image preprocessing, normalization then Skeletalization followed by segmentation and at last classification and neural networks. A stepwise approach is maintained to present in this paper.

Image preprocessing: In these method basic preparatory operations has to be performed before recognition of the text starts that begins with obtaining the image. Then operation is performed on the input scanned image such as filtration and noise removal including smoothening of image borders. Removal of Salt and Pepper noises is done by Median Filtering technique which is a non linear filtering technique as proposed in [3]. This causes the text in the image to be distinct and free of errors which prevents the image to be hazy. Smoothening of the borders of each character is done due to distinction between boundaries and which in turn will help in classification of each character.

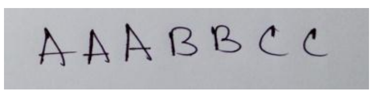
Normalization: These processes are divided into major three parts:

1) **Conversion to Grayscale.** Conversion of color images to gray scale images are done due to simplification of calculation during conversion into binary image.
2) **Binarization.** Binarization involves conversion of an image (here a grayscale image) with a binary valued image i.e., its pixel should have either 1 or 0 value. Binarization of image has always been great importance in field of scanning documents[7] and extracting features[6].

pi[x]=y; where if $x<110$ y=1; and if $x>=110$ y=0;

a> Original Image **b>Binary Image**

3) Splitting of Characters. This portion of the code deals with splitting of each character from a image, so that it can be processed individually in neural networks for identification.



Original Image

Splited image

Algorithm: Splitting is done by scanning the whole image vertically row by row.

01. Lets xmax=0, ymax=0, ymin=img.height, xmin=img.length

02. for(j=0 to img.length) {

03. for(i=0 to img.height){

04.if(p[i][j] equals to ‘1’){

05. xmin=MIN (j,xmin)

07..xmax=MAX (j,xmax)

07. ymin=MIN (i,ymin)

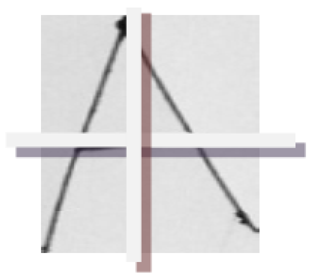
08. ymax=MAX (i,ymax) }}}

Skeletalization: Skeletalization is a morphological operation that is used to remove selected foreground pixels from binary images. Skeletalization is normally mainly applied to binary images and produces binary images as output. Skeletalization has been a part of morphological image processing and has played an important part in pre-processing phase of OCR systems. Many algorithms for vectorization by Skeletalization has been devised and applied to a great variety of pictures and drawings for data compression, pattern recognition and raster to vector conversion as given in [5].

Original Image

Skeletalized Image

Segmentation or Celled Projections: In this process each character sliced image is equally divided into four regions as shown below:



Segmented Image

As per [4], precession and accuracy in recognition of character increases when each character gets celled or segmented into certain regions. The main ides behind is that features extracted via zonings, crossings or by histogram projections would give unbiased values for each characters when it is in handwritten, but celled projections or segmentations takes a ratio of the spread of geometry over the image that tends to give constant value for each character in most cases.

Classification and Recognition: This process succeeds from the segmentation method, which calculation is based on the four regions and finding the ratios of filled is to unfilled regions of each region for a particular character do always tends to a constant. Let p[][] be the binary image of the original image of a particular character. Suppose "A". We previously encountered the rules in a binary image:

pi[i][j]=x;

if x<110 then y=1;

if x>=110 then y=0;

Now, we divided an image of character into half regions for both vertically and horizontally. Let no. of cols for a particular character after splitting be n and no. of rows be m. So we divide the image into n/2 and m/2 regions:

For region1, we calculate following variables:

$X=\sum_{0}^{n/2}\sum_{0}^{m/2} p[i][j]$, where p[i][[j]=1, for dark region

$Y=\sum_{0}^{n/2}\sum_{0}^{m/2} p[i][j]$, where p[i][[j]=0, for empty region

X/Y=K1 (a constant)

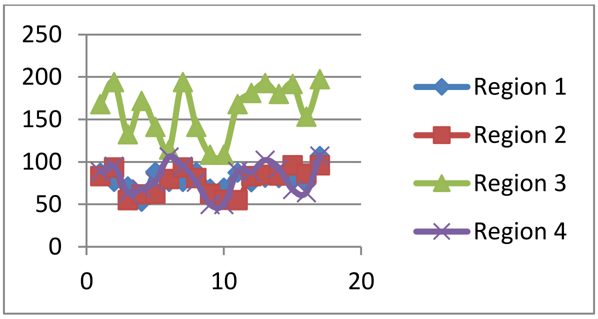
Like this we find k2, k3 and k4 for all three regions respectively. For alphabet "A", these four constants do remain almost same for any handwritten character of "A". Now we put k1, k2, k3 and k4 as input of our feed forward neural network in order train them and later after satisfactory training, we predict the output to satisfy the idea of this project.

Backpropagation: A Neural Network processes its categorization capabilities via run and error method known as training. There are many neural network models to run an operation. A feed forward neural network has been a one of the basic training models in order to perform operations for machine learning. It first set its weights from various training data provided to it by the supervisor and when the training seems to provide satisfactory results, testing process gets started. We have conducted our experiments on various models having different activation functions such as linear function, Hyperbolic Tangent function and Sigmoid function. Even the number of hidden layers in the models gets varied from 30 to 55. A total of 578 images are trained and then 56 images get tested via those models which let us to come a conclusion of which process gives a more accurate and time consuming results.

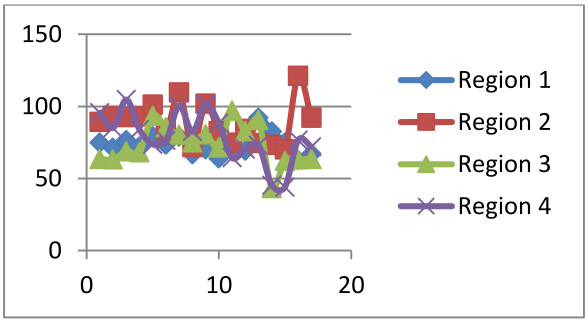
Table 1: Accuracy in percentage is shown for each activation function and number of hidden layers for both 4-segmented and 8-segmented images of both optical and handwritten characters.

	Linear activation function		**Hyperbolic Tangent Function**		**Sigmoid or Logistic Function**	
	30-45	46-55	30-45	46-55	30-45	46-55
4 Segmentation	68.11%	71.55%	72.69%	77.44%	82.86%	85.21%
8 Segmentation	77.32%	73.41%	78.32%	81.33%	84.19%	89.6%

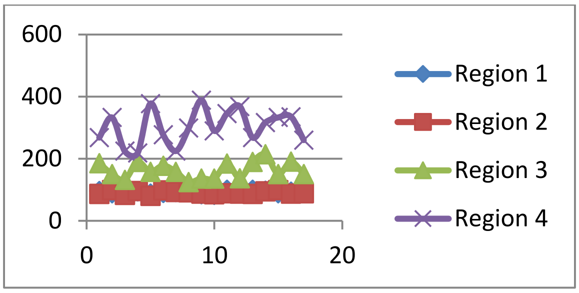
Analysis of Data: Each character is analyzed based on its density in each region and then tries to observe the causes of anomalies. In most of the text images, character A and C seems to get right prediction due to their unique geometry. We take the 4-segmented image data of handwritten characters for analyzing the geometry of each character.



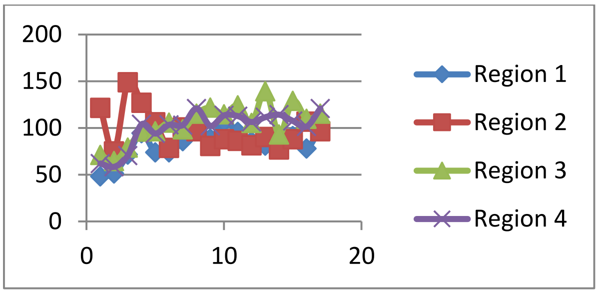
Data spread of "A"



Data Spread of "B"



Data spread of "C"



Data spread of "D"

From the graphs we can observe that characters "A" and "C" do have their unique geometry which easily make them recognizable from others. But characters "B" and "D" have an overlapping geometry that makes them recognizable to each other, which causes the error in the recognition procedures in some cases of handwritten characters, which we name as B-D anomaly.

EXPERIMENTAL RESULTS

We have taken total 578 images for training and out of which 204 are optical characters images and left are the handwritten ones. We have used 12 different writers for the variance of handwriting. 56 after images which constitute both handwritten and optical characters images get trained. While training both space and time complexities are to be kept in mind. Comparing images as in case of [8] and [9] do increases the speed up to an extent where data compression is involved, but it even reduces the accuracy of character recognition. That is why neural network seems to be an optimal solution. Increasing the number of hidden layers do increases accuracy but it increases the time complexity to a minimal level. An accuracy of 89.6% seems to be a fair result comparing with Tesseract, an OCR software which provides an accuracy of 93.1%.

CONCLUSION

The main aim of this project is to compare different situation in which the proposed algorithm seems to work accurately and in a time consuming manner. Each step in the paper is described in a detailed and simplistic manner as much as possible and experimented images along with diagram have been provided for better understanding. The future work would include even more complicated models of Neural Networks such as PNN and CNN.

ACKNOWLEDGEMENT

The author would like to thank Miss Maithili. This work is partially supported by Vellore Institute of Technology University, Vellore, India.

REFERENCES

[1] Noman Islam, Zeeshan Islam, Nazia Noor, A Survey on Optical Character Recognition System, *Journal of Information & Communication Technology-JICT Vol. 10 Issue. 2, December 2016*

[2] Sukhpreet Singh, Optical Character Recognition Techniques: A survey, ISSN: 2278 – 1323 *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 6, June 2013*

[3] Sarbjit Kaur, Er. Ram Singh, Image De-Noising Techniques: A Review Paper, ISSN (Online): 2347 – 4718, International *Journal For Technological Research In Engineering Volume 2, Issue 8, April-2015*

[4] Recognition M. Zahid Hossain, M. Ashraful Amin, Hong Yan, Rapid Feature Extraction for Optical Character, *Manuscript Draft, June 4, 2012, arXiv:1206.0238v1 [cs.CV] 1 Jun 2012*

[5] Gupta Rakesh and Kaur Rajpreet, Skeletonization Algorithm for Numeral Patterns, *International Journal of Signal Processing, Image Processing and Pattern Recognition* 63

[6] Badawy, W. "Automatic License Plate Recognition (ALPR): A State of the Art Review." (2012): 1-1.

[7] Ntirogiannis, Konstantinos, Basilis Gatos, and IoannisPratikakis. "A Performance Evaluation Methodology for Historical Document Image Binarization." (2013): 1-1

[8] Ayatullah Faruk Mollah, Nabamita Majumder , Subhadip Basu and Mita Nasipuri, Design of an Optical Character Recognition System for Camerabased Handheld Devices, *IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011 ISSN (Online): 1694-0814*

[9] Fatima Ezzahra Zdadou Supervisor: Dr. Falah, The Optical Capture Recognition Mobile Application, EGR 4402 Capstone Design, Fall 2015,

file:///C:/Users/Xander/Desktop/Masters%20College%20Details/The%20Optical%20Capture%20Recognition.pdf