International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

Handwritten Text Recognition using Neural Network

Chintan A. Modi
Electronics &Communication Engineering Department, SCET, Surat
Maulin M. Joshi
Electronics &Communication Engineering Department, SCET, Surat

ABSTRACT

Handwritten Text Recognition (HTR) is very challenging and subject of much attention in the field of recognition. This is due to the fact that writing styles of people vary to a great extent and it becomes difficult for the computer to recognize the handwritten characters. Various techniques are proposed in the literature including restrictions like specific writing styles-uppercase, lowercase or numeral characters. A more difficult problem is therecognition of characters when the writing style is not known a priori. The paper discusses handwritten English character recognition system using neural network. Simulation results include using neural based recognition for combined text comprising of capitalalphabets, small alphabets and numerals. The results show that neural network based method improves the text recognition in terms of accuracy.

Keywords

Handwritten Character Recognition, Binarization, Neural Network, Zoning, SVM, BAM.

INTRODUCTION

In the recent years, Handwritten Text Recognition (HTR) has been a challenging and interesting research area in the field of pattern recognition and image processing. It is a challenging issue to develop a practical cursive, handwrittentext recognition system which can maintain high recognition accuracy and is independent of the quality of the input documents.

Character Recognition is the mechanical or electronic translation of images of handwritten, typewritten or printed text into machine-editable text. Handwritten Character Recognition (HCR) is the process to classify characters from the input handwritten texts, as per the predefined character classes. Whereas, HTR is the process to segment line from text and finallyclassify characters from line.

HTR assumes considerable importance because of its applicability to pen-based interfaces and recognition of handwriting on scanned documents. These domains of the applicability classify theminto on-line HTR and off-line HTR. Both mode of handwriting recognition entail different modes of input, representation, processing and recognition strategies. Online HTR involves the use of pen based input devices to capture the sequence of co-ordinate points as the character is written. This gives information of the number, order and direction and writing speed of strokes. Here a stroke is defined as the trace of the pen-tip, captured from a pen-down event till the following pen-lift event. Offline recognition refers to a process of recognition performed later than handwriting capture. In the offline recognition, the writing is usually captured optically by a scanner and the completed writing is available as an image.

Previous work [1] suggests that the online methods perform better than offline counterparts in recognizing handwritten characters. This is due to the temporal information available with the online methods. Moreover, human writing varies from person to person and even for the same person depending on mood, speed, environment etc. However, in the offline systems, the soft computing techniques have been successfully used to yield comparably high recognition accuracy levels. Soft computing techniques like artificial neural networks (ANN) and statistical classifiers to extract rules based on numerical data canbe employed for offline HTR. Offline HTR covers several applications including mail sorting, bank processing, document reading that

inherently require more accuracy indicating obvious need of offline handwriting recognition systems. Due to above facts, the offline handwriting recognition field has attracted new researches to improve recognition accuracy of present work [2].

Basic block diagram containing training and testing phases are shown in the fig. 1. During image acquisition stage, input samples are passed through a scanner to the system and then given to pre-processing, where it converts the image into a form suitable for subsequent processing. Next stage is segmentation, where the input image is segmented into individual glyphs. This step separates out lines from text i.e. line segmentation and subsequently words and letters from lines. In Feature Extraction, important features are extracted that forms a vital part of the recognition process. During training, such data is stored in the database where, during classification, a character is placed in the appropriate class to which it belongs.

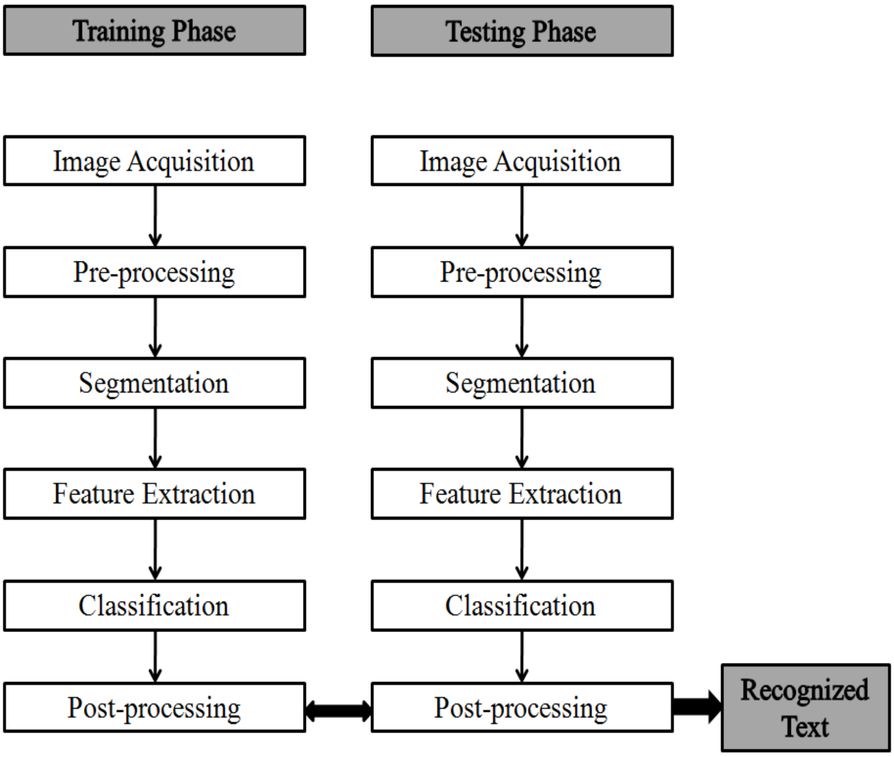

Fig 1: Basic Block Diagram of HTR System

Handwriting, as has always been, is assumed to be continued as preferred means of communication. Effective HTR systems need to be designed to convert these handwritten documents into an editable format. HTR systems aim at higher accuracy, with considerably reduced computational and storage space requirements.In [3], authors have described problems related to automatic text detection and recognition in imagery as well as how to solve major problems like text localization, verification, segmentation, and recognition by using different methods and how can we overcome challenges by using above methods.

Research aimed to investigate the identification and verification of freelance writers with more text scripts[4]. They have employed a set of run-length features which have shown promising results on a database of handwritten documents in two different languages, Greek and English. The performance measures used are the identification rate by using SVM as well as K-NN and the Equal-Error-Rate (EER) for verification task.

Our current work focuses on various approaches for offline handwritten Englishtext that comprises of Capital alphabets, small alphabets, numerals and rest of the paper is organized as follows. In section 2, Preprocessing is discussed. Section 3 discussed about the Segmentation. Section 4 discussed the Bi-directional Associative Memory approach. In Section 5, SVM based approach is presented. InSection 6, neural network approach is discussed. Section 7 shows simulation setup along with its results and finally, Section 8 concludes the paper.

PREPROCESSING

Preprocessing is the image Enhancement Technique. Preprocessing is important because it converts the image into a form suitable for subsequent processing and feature extraction. Major preprocessing steps are:

Binarization

Image under consideration is converted into a gray scale image. This grayscale image is subjected to binarization. Binarization is performed using Otsu's method.

Fig2: (a) Original Character(b) Binary Character [12]

This converts the gray-scale image into black and white image where in the pixel values of the image are either 0 or 1. Such an image is referred to as a binary image as seen in the fig. 2.

Cropping and Resizing

Once the image is free from noise, the extra portion present in the image other than the portion occupied by the character needs to be eliminated so that only the character can be processed. This process is known as cropping. In case of cropping an image, initially the top-leftmost black pixel of the character is first identified, and stored in a in a temporary variable. Similarly, the top-rightmost black pixel, bottom-leftmost black pixel and bottom-rightmost black pixel of the character are identified and stored. These values are fed to the cropping function in order to extract only the character from the image. After cropping the character image, the image is to be resized so that similarity is maintained in further processing.

SEGMENTATION

The most basic step in HTR is to segment the input image into individual glyphs. In present approach, this is needed in two different phases, with slightly different requirements. The first is during the training stage, where segmented glyphs are presented to the human supervisor for manual classification. The other is after the network is trained and we want to recognize a new image. In this case, we need to identify each glyph in the correct sequence before extracting features from it and classifying. To make things easier, especially for the second step, we first try to split the image into individual lines and then characters and for thatour input images are thresholded binary images and the method we came up is projection profile analysis.

BAM BASED APPROACH

A neural network classifier technique Bi-directional Associative memory (BAM) can be used for recognition. It is an associative neural networkthat retrieves an object or memory, based on part of the object itself. The following section describes the architecture and the working of the BAM.

Architecture

Fig.3 is BAM model that is a two layer associative recurrent neural network [5], in which neurons are connected by means of directional weight connection paths. In contrast to linear association, the connections are bidirectional, i.e., $W_{ij} = W_{ji}$, for i = 1, 2, ..., m and j = 1, 2, ..., n and units in both layers serve as both input and output units depending on the direction of propagation. The net iterates by sending a signal back and forth between the two layers until each neuron's activation remains constant for several steps.

BAM operation can be carried out by using:

$$W_k = X_k^T Y_k \qquad (1)$$

to store a single associated pattern pair and

$$W = \alpha \sum_{k=1}^{\Psi} W_k \tag{2}$$

To simultaneously store several associated pattern pairs.

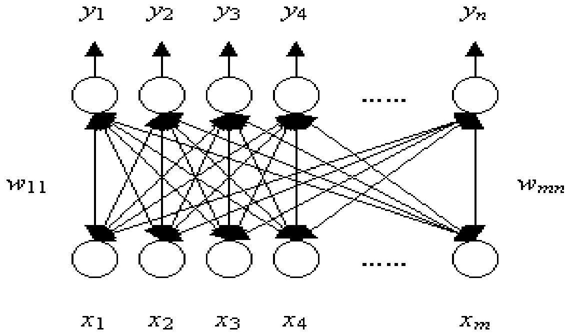

Fig3: BAM Model [5]

Algorithm

From the given image of handwritten text, after performing segmentation and feature extraction; critical features are available. Next is to define/decide input pattern for all characters. M noisy inputs (error) for each of alphabetic characters are created/ presented. The noisy inputs may be created by flipping p random locations. A weight matrix is created by auto associative property, as defined in equation (1). For testing purpose, actual/test patterns are compared to the stored patterns using Hamming distance and best match is found.

Current work is the extension of our earlierwork [5]. Our results proved that BAM approach fails to achieve high accuracies and henceforth discussion on theBAM is omitted.

SVM BASED APPROACH

Support Vector Machine (SVM) is primarily a binary classifier which classifies dataset by finding optimal hyper plane. It is a supervised learning method that can be used for classification, regression or other tasks [6].

Architecture

This approach is called linear classification because there are many hyper-planes that might classify the same set of data as can be seen in the fig. 4. The power of SVM lies in its ability to transform data to a high dimensional space where the data can be separated using a hyper plane. SVM is a well-developed technique to create optimal hyper plane which distinct two classes by maximizing the distance or margin between two classes.

The optimization process for SVM learning with different parameters of hyper plane needs to be understood. The SVM cannot be used directly to solve multi-class problem. The formulation to solve multi-class SVM problems in one step has variables proportional to the number of classes. Therefore, for multi-class SVM methods, either several binary classifiers have to be constructed or a larger optimization problem is needed.

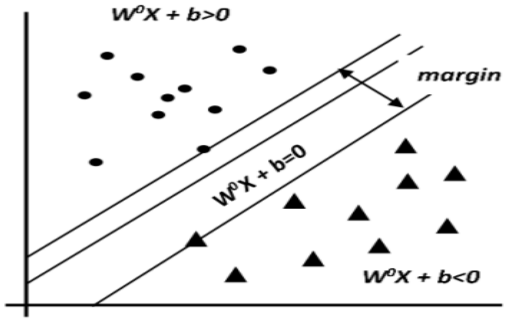

Fig 4: Two Class Support Vectors [7]

Hence, in general, basic SVM model is computationally more expensive to solve a multi-class problem than a binary problem with the same number of data [8]. In the context of multi-class problem, SVM can be used with variations. They are one-against-one and one-against-all.

Current work deals with combination of capital, small alphabets as well as numerals which is associated with multi-class classification problem. Due to same, discussion on SVM for multi-class problem is left for future studies and we focus on feed forward Back propagation Neural Network in next section.

NEURAL NETWORK APPROACH

Artificial Neural Network (ANN) [9-11], consisting of number of artificial neurons, is a machine learning approach that models human brain and it has found applications on various classification tasks due to its ability of generalization. Patterns are presented to the network via the 'input layer', which communicates to one or more 'hidden layers' where the actual processing is done via a system of weighted 'connections'. The hidden layers then links to an 'output layer' to obtain the desired output as shown in fig. 5.

This is the most common structure for neural networks: three layers with full interconnection. The input layer nodes are passive, doing nothing but relaying the values from their single input to their multiple outputs. In comparison, the nodes of the hidden and output layers are active, modifying the signals in accordance with fig. 5. The action of this neural network is determined by the weights applied in the hidden and output nodes.

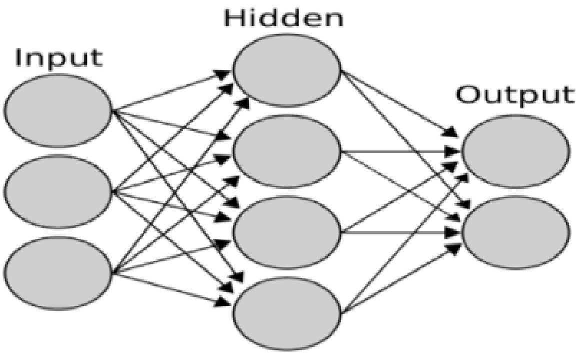

Fig.5: Basic Neural Network Model[5]

Each layer consists of one or more nodes, represented in diagram by the small circles. The lines between the nodes indicate the flow of information from one node to the next. In this particular type of neural network, the information flows only from the input to the output (i.e. from left-to-right). Other types of neural networks have more intricate connections, such as feedback paths. The nodes of the input layer are passive, meaning they do not modify the data. They receive a single value on their input, and duplicate the value to their multiple outputs. In comparison, the nodes of the hidden and output layer are active.

Feature Extraction

The main objective of feature extraction is to remove redundancy from data and hence it is very important to select features that allow effective and efficient recognition of pattern. Feature extraction technique is applied, when the pre-processing and the desired level of segmentation (line and character) have been achieved, some feature extraction technique is applied to the segments to obtain features.

It is essential to focus on the feature extraction phase as it has an observable impact on the efficiency of the recognition system. Feature extraction has been given as, “Extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the between class pattern variability”.

Thus, selection of a suitable feature extraction technique according to the input to be applied needs to be done with utmost careand one of the techniques, Zoning, is described next. Feature extraction is followed by application of classification and post processing techniques.

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

Zoning

Zoning is one of the feature extraction techniques used in the literature [12] to obtain information about local characteristics of the image under consideration. Steps involved in zoning are described as under.

In order to obtain the local characteristics instead of global characteristics, first the character image is divided into N*M zones and from each zone; features are extracted to form the feature vector.

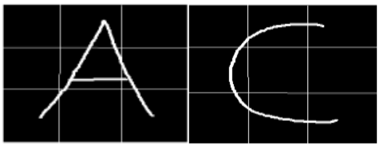

Fig6: Character A & C divided into zones [12]

- ***Euler Number:***Concept of Euler number is used to classify the characters. Euler number is defined as the number obtained by subtracting the number of the number of holes in the image from the number of objects in the image. E.g. as shown in the fig.6, the image is of uppercase character “C”. Here number of objects is ONE and number of holes is ZERO. Therefore, Euler Number is ONE.

<u>Euler Number</u> = Number of characters in the image - Number of holes in the character

- ***End Points:***Concept of finding the end points of characters is added to eliminate the problem arising from using only the Euler number. End points of each character can be noted down to which zones they lie into as shown in the fig.7.

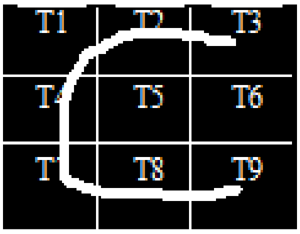

Fig 7: End-points of character “C” [12]

But this attempt too failed, as there were few characters which had the same combination for euler number and end point such as H and K as shown in fig. 8.To overcome this problem, aspect ratio in each zone was calculated and based on that; zone map of the character was created.

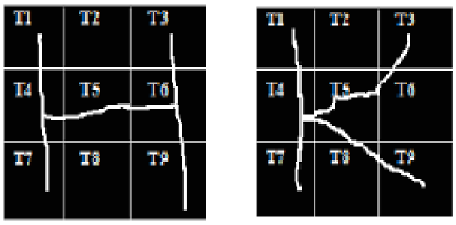

Fig 8:Character “H” and “K” with its EPs in T1, T3, T7 and T9 [12]

- ***Aspect Ratio:*** Aspect ratio in each zone can be calculated based on the which, the character can be found out uniquely. Total 5 types were created by [12] depending upon aspect ratio as discussed below:

Type 0: If aspect ratio = 0, it indicates no end point exists in the zone. It does not contain any pixels of the character. So, this zone will be assigned value 0.

Fig 9:Type 0: one of the zones out of 9 zones [12]

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 7
July 2017

Type 1:

Fig 10:Type 1: Case-1 and Case-2 [12]

<u>*Case 1:*</u> Pixels of character don't touch the bottom, top and right side of the zone and hence values of them are changed. It indicates the presence of an end point and the respective zone is assigned value 1.

<u>*Case 2:*</u> Pixels of character don't touch bottom, top and the left side of the zone and hence values of them are changed. It indicates the presence of an end point and the respective zone is assigned value 1.

Type 2:If aspect ratio >= 4, it prints value 2 for that zone in the zone map.

Fig 11:Type 2: Aspect Ratio >=4 [12]

Type 3:If 0 < aspect ratio <= 0.25, it assigns value 3 for that zone in the zone map.

Fig 12:Type 3: Aspect Ratio > 0 but <= 0.25 [12]

Type 4:If 0.5 <= aspect ratio < 4, it assigns value 4 for that zone in the zone map.

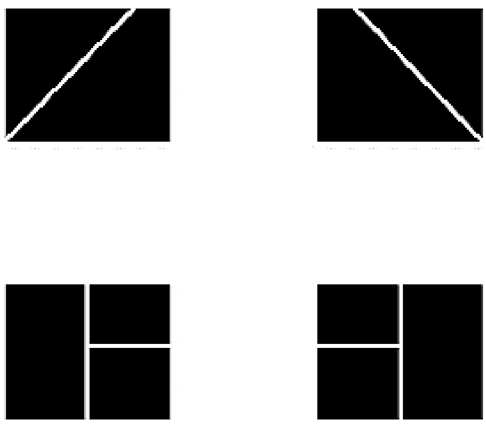

Fig 13:Type 4: Aspect Ratio >= 0.5 but < 4 [12]

Thus, the combination of euler number and zone-map of characters are used to recognize the characters.

Algorithm

- Collect samples (images) of handwritten English characters.
- Perform zoning on the character image i.e. Divide the character image into M x N and determine various feature inputs to neural network like Euler number, end points, and aspect ratio of the character etc.
- Train feed-forward back propagation neural network (FFBPN) using proper hidden layer, number of hidden neurons, activation functions, training pairs etc.
- Test the inputs using neural network to obtain the best match.

SIMULATION RESULTS

This section discusses the results obtained for handwritten text identification using neural network. Initial experimentations were conductedwith capital alphabets (A-Z) and then small alphabets (a-z) and numerals (0-9) were tested and in last combined them all in one.

Basic simulation setup for HTR using neural network considered feed forward back propagation neural network is for implementation with the feature i.e. character size 7x5 and 'Tansig' activation function is used for hidden as well as output layers. Measurement parameter considered is accuracy with the help of MSE.

Mean Square Error (MSE)

Mean Square Error (**MSE**) can be defined as average error squared between output and target,

$$\sum_{1}^{N} \frac{(e_i)^2}{N} = \sum_{1}^{N} (t_i - a_i)^2 \tag{3}$$

Where, t_i is the target and

a_i is the output of the output neuron i.

Accuracy

When different writers are considered, the accuracy is the most important task to maintain in different environment. Hence, Accuracy is the parameter that we have to improve among different writers and environment. Thus, the ratio of total number of correctly classified characters to the total number of characters gives us the recognition rate.Mathematically, Recognition Rate (%) can be calculated as,

$$= \frac{Total\ Number\ of\ Correctly\ Classified\ Characters}{Total\ Number\ of\ Characters} * 100 \tag{4}$$

Average Accuracy

As we know that the NN is one of the soft computing techniques that deals with the uncertainty, imprecision and tolerate to noise and the theory of NN tells us that the weights associated with different layers are changes and updated every time. Thus, accuracy is also changed every time and that’s why we looking towards the average accuracy that performs on N experimentation. Hence, Average accuracy is the ratio of total summation of recognition rate (%) to the total number of experimentation.Mathematically, Average Accuracy (%)can be calculated as,

$$= \frac{Total\ Summation\ of\ Recognition\ rate\ (\%)}{Total\ Number\ of\ Experimentation} \tag{5}$$

Simulation Results using NN

HTR using NN with character size 7x5

The architecture of the neural network includes an input layer with 62 inputs and 30 samples (24 samples for training and 6 for testing) of each input that helps to make total dataset of 1860, one hidden layer (decide from literature study) with 35 hidden nodes (found with experimentation) and an output layer with 6 neurons.

The gradient descent back-propagation method with momentum constant i.e. 0.5, adaptive learning rate i.e. 0.00100, number of epochs are 1000 with 6 validation checksand Tansig activation function is used for

training of the neural network. Neural network has been trained using 1488 known dataset with 0, 1, 2 and 3 errors located randomly. After training the network, the system was tested using 372 unknown dataset.

Table 1. HTR using NN with character size 7x5

Character Size	No. of Error(s)	Accuracy Level	Training Accuracy (%)	Testing Accuracy (%)
7x5	0	Min	85.4838	83.8709
		Avg	89.3548	88.7096
		Max	**91.9354**	**91.9354**
	1	Min	75.2688	66.1290
		Avg	83.4610	72.6612
		Max	**85.8198**	**84.4086**
	2	Min	54.0322	54.5698
		Avg	65.6115	56.5591
		Max	**80.2419**	**72.5806**
	3	Min	44.7044	40.9274
		Avg	51.4516	51.0215
		Max	**67.7096**	**67.0002**

Note: Results are calculated on training and testing dataset and average accuracy are computed from 10 experiments.

Results seem to be effective while there is no error in the input dataset (combined). Hence, NN is able to recognize the characters of the text with 91.9354% in both training and testing phase. Table1 depicts the accurate recognition rates where numbers of locating errors 1, 2 and 3, respectively.

CONCLUSION

This paper discusses free style handwritten recognition problem using various methods and detailed implementation using back propagation neural network. Wider perspectives of the problem were explored with different classification methods. Simulation results of NN highlight the fact that the efficiency for neural network approach is better compared to associative memory approach. This advantage is evident due the fact that in neural network it is possible to update the hidden layer with number of neurons, activation functions and bias values to hidden neurons according to requirements of inputs to generic training, an adaptive nature that is not possible in associative memory. Inclusion of hidden layer improves the accuracy of character recognition. In future, we plan to investigatetext recognition problem using SVM for multi-class classifications.

REFERENCES

[1] R Plamondon and S N Srihari, “On-line and off- line handwritten character recognition: A comprehensive survey,”IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63-84, 2000.

[2] N Arica and F.T. Yarman Vural, “An overview of character recognition focused on off-line handwriting,”IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and , vol. 31, no. 2, pp. 216-223, 2001.

[3] Q. Ye and D. Doermann, “Text Detection and Recognition in Imagery: A Survey,” in IEEE Transactions on

Pattern Analysis and Machine Intelligence, vol. 37, no. 7, July 1 2015, pp. 1480-1500.

[4] Chawki Djeddi, Imran Siddiqi, Labiba Souici-Meslati, Abdellatif Ennaji, “Text-independent writer recognition using multi-script handwritten texts,”Pattern Recognition Letters, Volume 34, Issue 10, 15 July 2013, Pages 1196-1202.

[5] Niki G. Mehta and Maulin M. Joshi, “Soft Computing Based Classifier Techniques for Offline Handwritten Character Recognition,” National Conference on Recent Trends in Computer and Communication Technology (RTCCT 2016), 10-11 May, 2016.

[6] Antonio Carlos Gay Thome. (2012) intechopen. [Online]. http://cdn.intechopen.com/pdfs/40722/InTechSvm_classifiers_concepts_and_applications_to_character_recognition.pdf

[7] Thome and Carlos Gay Antonio,“Svm classifiers-concepts and applications to character recognition,” in INTECH Open Access Publisher, 2012.

[8] Chih-Wei Hsu and Chih-Jen Lin, “A Comparison of Methods for Multi-Class Support Vector Machines,” in National Taiwan University by the National Science Council of Taiwan, NSC 89-2213-E-002-106, 2012.

[9] N. Sivanandam S. and N Deepa S., Introduction to Neural Networks Using Matlab 6.0, Tata ed., 2006.

[10] Neural Networks – A Comprehensive Foundation – Simon Haykin. [Online]. https://cdn.preterhuman.net/texts/science_and_technology/artificial_intelligence/NeuralNetworks-AComprehensiveFoundation-Simon Haykin.pdf

[11] Neural Networks (and more!) [Online]. http://www.dspguide.com/CH26.PDF

[12] R. R. Herekar and S. R. Dhotre, “Handwritten Character Recognition Based on ZoningUsing Euler Number for English Alphabets and Numerals,”IOSR Journal of Computer Engineering, vol. 16, no. 4, pp. 75-88, 2014.