Character Recognition Using Back-Propagation Method

Asst. Prof. Dipti R Chaudhary[1] Prof. Ankit Sharma[2]
Department of Electrical Engineering ITM Universe Vadodara
Department of Instrumentation and Control Institute of Technology Nirma University
diptichaudhary84@gmail.com
ankit.sharma@nirmauni.ac.in

Abstract:- This paper describe how handwritten English character Recognition (HCR) processed, trained and then recognized using Back propagation method. Size and fonts are different in training the data and testing the data. In the present paper, we have given a method to recognize a handwritten character using back propagation method. It is developed for isolated handwritten English Characters (A to Z). Preprocessing of Recognition is used binarization, thresolding and segmentation method. Image are first converted into gray scale then features are extracted which are in form of 0 and 1. 780 hand written characters are used in database for characters then test the data and find the Recognition accuracy.

Keywords:- Neural network, back propagation method, Binaryzation, Recognition, HCR.

I. INTRODUCTION

Handwriting recognition is undoubtedly one of the most challenging areas of pattern recognition. It is extremely useful in a wide range of real world practical problems, including documentation analysis, mailing address interpretation, bank check processing, signature verification, document verification and many others**.** Several pattern recognition approaches have been applied to both on-line and off-line handwriting recognition, including statistical methods, structural and syntactic methods, and neural networks. Some reading systems identify strokes, others **try** to identify Characters, groups of characters, or entire words **.**Computerized processing has been growing rapidly since the 1980's because of the exponentially increasing amount of daily received documents and the more powerful and affordable computer systems. Intuitively, the conversion of textual blocks into ASCII codes represents one of the most important tasks in document processing. Our strategy of reclassifying characters is to incorporate typographical structure analysis which categorizes characters in the first step, and therefore it reduces the scope of character Recognition

The recognition of handwritten characters has many applications such as automatic Postal sorting, automatic bank cheque processing etc. In the work on character recognition has been reviewed. The earlier systems known as Optical Character Recognition (OCR) systems that had been developed were confined to recognize only the printed or handwritten characters of fixed size and fonts. But, the present study aims at producing a system, which could recognize characters of any arbitrary size, shape and fonts. There are numerous approaches that address the problem and they vary in the features extracted from the graphical representation of the Characters.

II. IMAGE PREPROCESSING

The first phase in our character recognition process is converting the image to Binary image by thresholding the given character image. Binary images are images whose pixels have only two possible intensity values. They are normally displayed as Black and White. The converted Character image has pixel value zero for Black and one for white. Thus the color of the character is White and the background **i**s black.

Preprocessing techniques are needed on color, grey-level or binary document images containing text and/or graphics. In character recognition systems most of the applications use grey or binary images since processing color images is computationally high [5].Such images may also contain non-uniform background and/or water marks making it difficult to therefore; the desired result from preprocessing is a binary image containing text only. Thus, to achieve this, several steps are needed, first, some image enhancement techniques to remove noise or correct the contrast in the image, second, thresholding to remove the background containing any scenes, watermarks and/or noise, third, page segmentation to separate graphics from text, fourth, character segmentation to separate characters from each other and, finally, morphological processing to enhance the characters in cases where thresholding and/or other preprocessing techniques eroded parts of the characters or added pixels to them. The above techniques present few of those which may be used in character recognition systems and in some applications; few or some of these

techniques or others may be used at different stages of the HCR system.

Pre-processing

Pre-processing covers all those functions carried out prior to feature extraction to produce a cleaned up version of the original image so that it can be used directly and efficiently by the feature extraction components of the character recognition.

The steps in pre-processing involves

Binearization: it is process of converting a gray scale image into binary image by thresholding. **Smoothing**: the erosion and dilation smooth the Boundaries of objects without significantly changing their area.

Edge detection: morphological gradient operators are used in edge detection because they enhance intensity of edges of characters.

Segmentation: the characters are always written in "print fashion", not connected, horizontal histogram profile (for line segmentation), vertical histogram profile (for word segmentation) and connected component analysis are able to handle the character segmentation problem.

II. SEGMENTATION

After scanning the document, the document image is subjected to pre-processing for background noise elimination, and binarization to generate the bit map image of the text.

The pre-processed image is then segmented into lines, words and characters. The details of line, word and character segmentation are discussed in the following sub-sections.

Line Segmentation
Word Segmentation
Character Recognition

Line segmentation-To separate the text lines, from the document image, the horizontal projection profile of the document image is found.

Word segmentation-The spacing between the words is used for word segmentation.

Character Recognition-It is provide the Spacing between the characters. so it is called segmentation.

HCR system more robust mainly through accurate image enhancement, noise removal, image thresholding, skew detection/correction, page segmentation, character segmentation, character normalization and morphological techniques.

III. NEURAL NETWORK

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. We can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements.

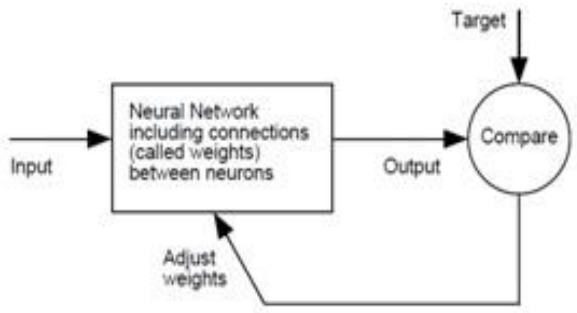

Fig. 1: Neural Network

Back propagation was created by generalizing the Windrow-half learning rule to multiple layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with output vectors. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a number of discontinuities. Standard back propagation is a gradient descent algorithm, as is the Windrow-Ho_ learning rule, in which the network weights are moved along the negative of the gradient of the performance function. The term back propagation refers to the manner which the gradient is computed

There are a number of variations on the basic algorithm that are based on other standard optimization techniques, such as conjugate gradient and Newton methods. The Neural Network Toolbox implements a number of these variations.

IV. ARCHITECTURE OF NEURAL NETWORK

This section presents the architecture of the network that is most commonly used with the back propagation algorithm with the multilayer feed forward network. Basically the neural network consists of: Neurons Interconnection Neurons: Similar to the human brain neurons. Neurons in the network transport the incoming information on their outgoing connections to the other neurons. Weights: outgoing connection from neuron to another neuron are called weights. The following information is simulated with special values stored in those weights.

FEED FORWARD NETWORK

Feed forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range -1 to +1. On the other hand, if you want to constrain the outputs of a network (such as between 0 and 1), then the output layer should use a Sigmoid transfer function (such as logsig).

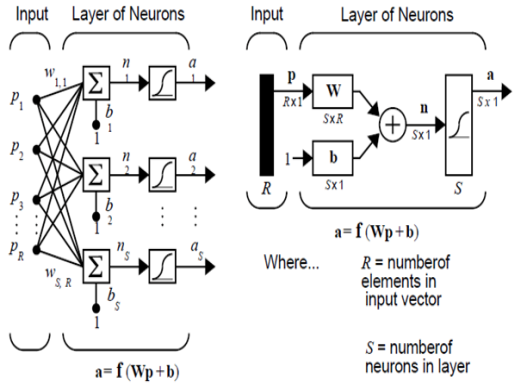

Fig. 2: Back Propagation Method

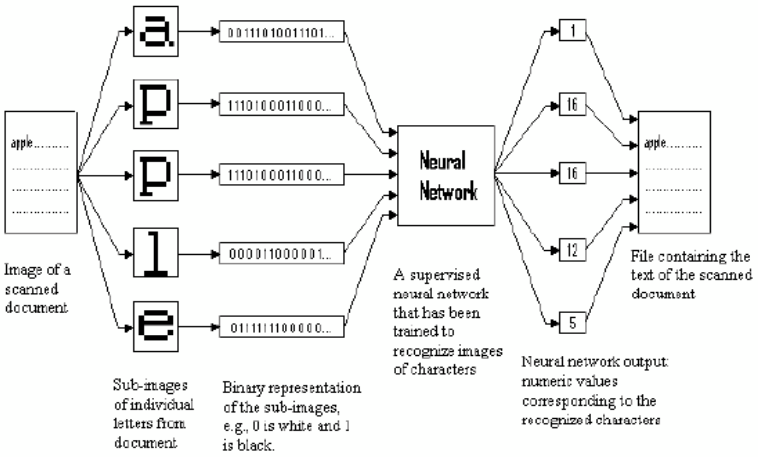

Fig. 3: Process of Neural Network

V. EXPERIMENTAL SET-UP

The databases consist of 780 samples of hand written character. The samples are separated into 2 set of training and testing data. The trained neural network is tested with testing data se

DATASET OF CHARACTERS

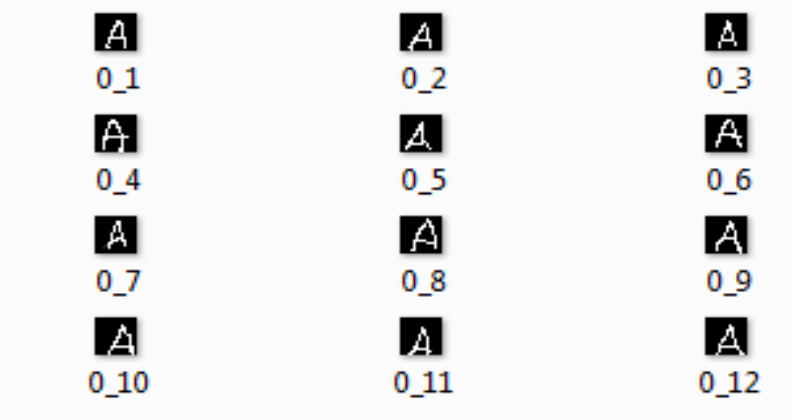

Fig.4: Samples of hand- written character used in this experiment.

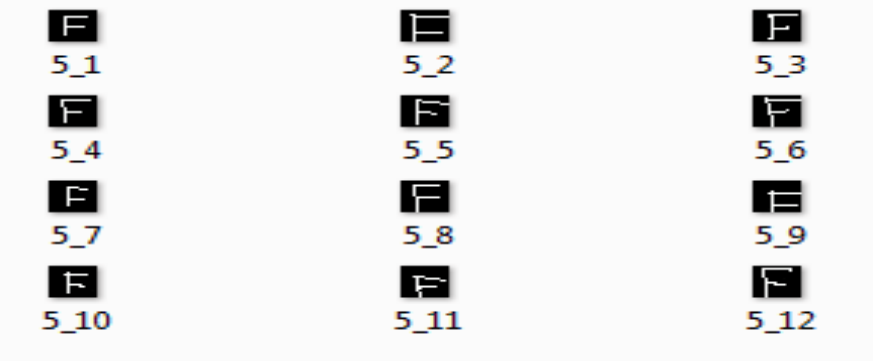

Fig.5: Samples of hand- written character used in this experiment.

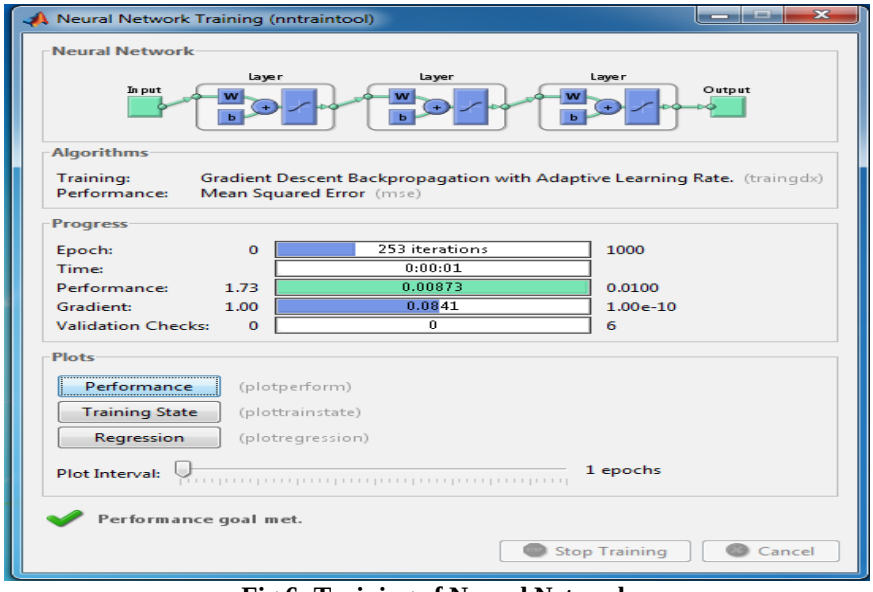

Fig.6: Training of Neural Network

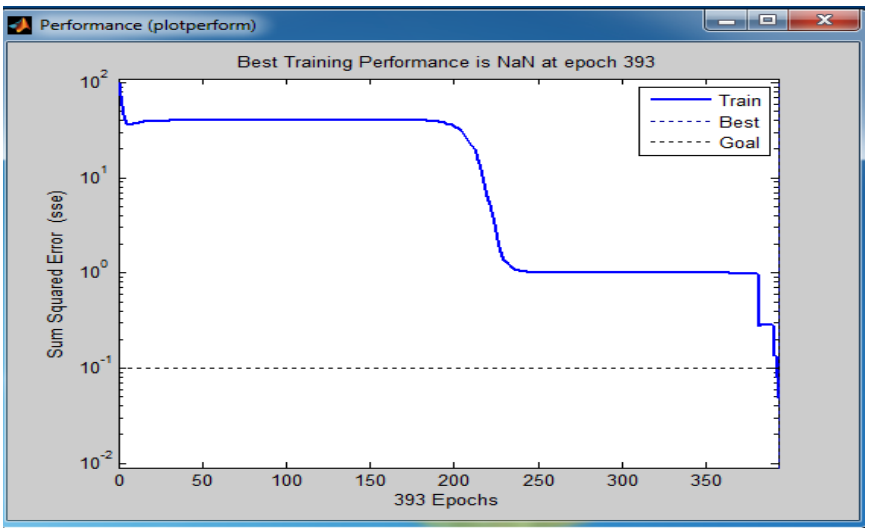

Fig.7: Training performance

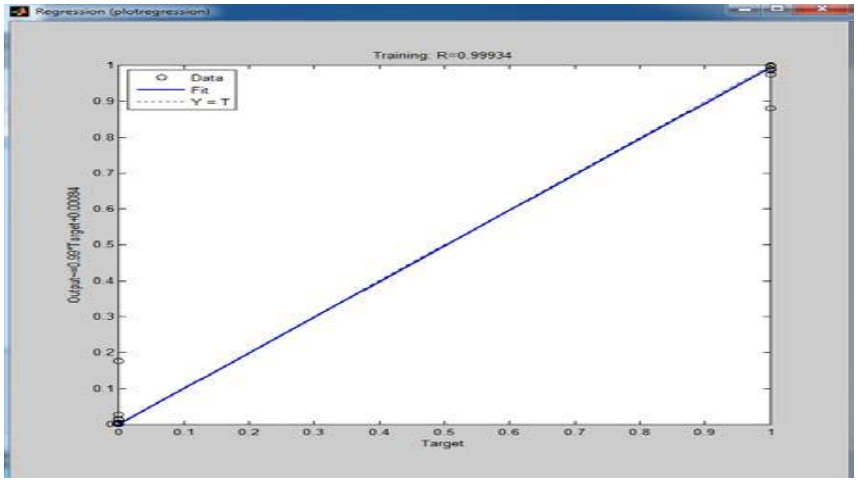

Fig.8: Regresor performance

VI: RESULTS

The experimental results are shown in Table 1. The best accuracy achieved by the Network is 85%.

Input Character	No of training sample	Correct Identification	Accuracy percentage
A	30	24	80
B	30	24	80
C	30	22	70
D	30	21	70
E	30	28	90
F	30	24	80
G	30	28	90
H	30	24	80
I	30	28	90
J	30	28	90

VII: COCLUSION

It is gives a useful method for the Recognition of handwritten characters to a great extent. The proposed method has been applied on different unknown characters of English language. Handwritten English Character sets are taken. These steps are followed to obtain best accuracy of input hand written English character image from the HCR system. Proposed method gives the accuracy approximately 85 percent for the Character. First of all, training of

system is done by using different data set or sample. And then system is tested for few of the given sample, and accuracy is measured.

REFERENCES

[1]. Verma B.K,\Handwritten Hindi Character Recognition Using Multilayer Perceptronand Radial Basis Function Neural Network",IEEE International Conference on Neural Network, 4, pp. 2111-2115, 1995.
[2]. Sutha .J, Ramraj.N,\Neural Network Based Online Tamil Handwritten Character Recognition Syste, IEEE International Conference on Computational Intelligence and Multimedia Application, 2007, 2, 13-15, Dec.2007, Page(s): 446-450, 2007.
[3]. Yuelong Li Jinping Li Li Meng,\Character Recognition Based on Hierarchical RBF Neural Networks" I ntelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference, 1, On Page(s): 127-132, 2006.
[4]. Y.Y. Chung, and M.T.Wong,\Handwritten Character Recognition by Fourier De-scriptors and Neural
[5]. Network", Proceeding of IEEE TENCON - Speech and Image Technologies for C.
[6]. Gonzalez, Rafael C., and Richard E.Woods. Digital Image Processing. Addison-Wesley, 1992. p. 518.