Neuro Recognizer: Neural Network Based Hand-Written Character Recognition

K. Radha Revathi[1*], A.N.L Kumar[2], Andey Krishnaji[3]

[1*,2,3]*Department of MCA,*

Swarnandhra College of Engineering & Technology, JNTUK University, Narsapuram,

www.ijcseonline.org

Abstract: This paper presents an Artificial Neural Network (ANN) based approach for the recognition of handwritten characters in the presence of noise because now a days the handwritten recognition place a crucial role in various industrial applications. In the Handwritten recognition noise has been regarded as one of the major issue that degrades the performance of character recognition system. So to recognize the handwritten characters in different noise levels, In order to overcome this limitation, the back propagation (BP based ANN) is designed for the handwritten character recognition. The recognition system is designed and tested in JAVA under different noise levels. Experimental results indicate that the proposed approach can obtain very high recognition rate for all handwritten characters in the presence of noise.

Keywords: OCR, MCR, ANN, Character Recognition, Back Propagation Algorithm.

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 [1]. Several pattern recognition approaches have been applied to both online 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. In the problem of hand written character recognition neural networks are heavily used for efficiently predicting the hand written symbols [3].

Character recognition has long been a critical area of the Artificial Intelligence. Recognition is a trivial task for humans, but to make a computer program that does character recognition is extremely difficult [2]. Recognizing patterns is just one of those things humans do well and computers don't [2].

One of the primary means which the computers are endowed with human like abilities is through the use of neural network. Neural Networks are particularly useful for solving the problems that cannot be expressed as the series of steps such as recognizing patterns, classifying them into groups and series prediction and data mining. ANN approach for character recognition is now gaining importance because of ANN's highly parallel architecture and fault tolerance [14].

II. Literature Review

Literature survey is the most important step in software development process. It is necessary to determine the economy, time factor and company strength before developing the tool. After these things are satisfied, next step is to determine which language and operating system can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. Before building the system the above consideration are taken into account for developing the proposed system.

Character recognition is an art of detecting segmenting and identifying characters from image. More precisely Character recognition is process of detecting and recognizing characters from input image and converts it into ASCII or other equivalent machine editable form. It contributes immensely to the advancement of automation process and improving the interface between man and machine in many applications. Character recognition is one of the most interesting and fascinating areas of pattern recognition and artificial intelligence [1].Character recognition is getting more and more attention since last decade due to its wide range of application. Conversion of handwritten characters is important for making several important documents related to our history, such as manuscripts, into machine editable form so that it can be easily accessed and preserved. Lots of independent work is going on in Optical Character Recognition that is processing of printed/computer generated document and handwritten and manually created document processing i.e. handwritten character recognition. The character recognition system consists of feature extractor and classifier with the stored patterns in the database. The same characters differ in sizes, shapes and styles from person to person and even from time to time with the same person. Like any image, visual characters are

subject to spoilage due to noise near the edges. Also, there are no hard-and-fast rules that define the appearance of a visual character. Thus, classical methods in pattern recognition are not perfect for the recognition of visual characters. 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 online 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 [1]**.**

Character recognition is the process to classify the input character according to the predefined character class. With the increasing interest of computer applications, modern society needs that the computer should read the text. The text may be in the form of scanned handwritten document or typed text in various fonts or a combination of both. The character recognition system helps in making the communication between a human and a computer easy. Classical methods in recognition are not perfect for the recognition of visual characters due to the following reasons [2].
1. The „same" characters differ in sizes, shapes and styles from person to person and even from time to time with the same person. The source of confusion is the high level of abstraction: there are thousands styles of type in common use and a character recognition program must recognize most of these.
2. Like any image, visual characters are subject to spoilage due to noise. Noise consists of random changes to a pattern, particularly near the edges. A character with much noise may be interpreted as a completely different character by a computer program [2].
3. There are no hard-and-fast rules that define the appearance of a visual character. Hence rules need to be heuristically deduced from the samples. Character recognition system is useful in license plate recognition system, smart card processing system, automatic data entry, bank cheque /DD processing, and money counting machine, postal automation [1]. Character Recognition has gained impetus largely due to its application in the Fields of Computer Vision, Intelligent Text Recognition applications and Text based decision making systems. The approach taken to solve the OCR problem was based on psychology of the characters as perceived by the humans. Thus the geometrical features of a character and its variants were considered for recognition.

Another approach, namely Recognition using Correlation Coefficients was based on the Cross Correlation of input characters or their transforms, with the database templates; so as to accommodate minor differences was used. It introduced False or Erroneous Recognition among characters very similar in shape, such as 'I' & 'J',

'B' & '8', 'O', 'Q' & '0' etc. The solution to this problem lies in ANN, a system that can perceive and recognize a character based on its topological features such as shape, symmetry, closed or open areas, and number of pixels [4]. The advantage of such a system is that it can be trained on 'samples' and then can be used to uniquely identify each character is called its Vector. Thus, a Vector Database is utilized to train the network, so as to enable it to effectively recognize each character, based on its topological properties to recognize characters having a similar (not exact) feature set. The ANN used in this system gets its inputs in the form of Feature Vectors. This is to say that every feature or property is separated and assigned a numerical value. The set of these numerical values that can be used to uniquely identify each character is called its Vector. Thus, a Vector Database is utilized to train the network [4].

III.   Problem Statement

Given a handwritten character, the system needs to predict the type of the character. In other words if we can write the character "A" the system predict the character that it is truly "A" or the input character is nearer to "A" or something else. The purpose of this project is to take the hand written characters as an input process the character, train the neural network effectively by using the algorithm to recognize the pattern.

IV.   Objectives
The proposed system will have the following objectives:

- The system should offer effective training method.
- To provide an ease of use interface for the users to draw the characters with free hand.
- System should be able to effectively predict the hand-written characters in the presence of noise.
- System should draw the learning graph.

V.   Methodology

A. Neural Networks
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. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. There, the Network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are used, in this supervised learning, to train a network. An artificial neural Network as the backend is used for performing classification and recognition tasks. In the off-line recognition system, the neural networks have

emerged as the fast and reliable tools for classification towards achieving high recognition accuracy [3],[13],[8].
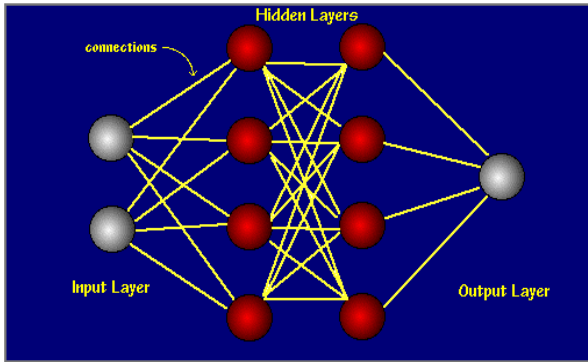

Figure-1: Neural Network

B. Neuron

A Neural Network consisting of number of neurons that is also called Artificial Neuron. An artificial neuron is a mathematical function conceived   as   a model of biological neurons. Artificial neurons are the constitutive units in an artificial neural network. The artificial neuron receives one or more inputs (representing dendrites) and sums them to produce an output (representing a neuron's axon)[13], [8]. Usually the sums of each node are weighted, and the sum is passed through a non-linear function   known   as   an activation function or transfer function. The basic structure of the neuron is shown in figure-2
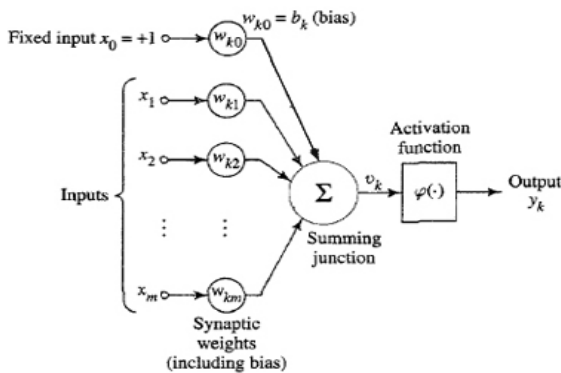
Figure-2: Basic structure of a Neuron

The output of the neuron can be computed using the formula, $y_k=f(x_1w_{k1}+x_2w_{k2}+x_3w_{k3}+BIAS)$, where x_1,x_2 and x_3 are the input values whereas w_{k1} ,w_{k2} and w_{k3} are the weights in that layer [14].

Character is identified by analyzing its shape and comparing its features that distinguish each character. Also an analysis has been carried out to determine the number of hidden layer nodes to achieve high performance of back propagation network.

C. Back Propagation Neural Network

A Back propagation (BP) network consists of at least three layers of units, an input layer, at least one intermediate hidden layer, and an output layer. When a Back propagation network is cycled, an input pattern is propagated forward to the output units through the intervening   input-to-hidden   and   hidden-to-output weights. Back   propagation   learns   by   iteratively processing a data set of training values, comparing the networks prediction for each set with the actual known target values. For each training dataset, the weights are modified so as to minimize the mean squared error between the networks prediction and the actual target value. These modifications are made in the "backwards" direction, i.e. from the output layer, through each hidden layer. Although it is not guaranteed, in general the weights will eventually converge, and the learning process stops [12], [5].

Neural network uses Back propagation which is a technique and a supervised algorithm that learns by first computing the output using a feed forward network, then calculating the error signal and propagating the error backwards through the network. The BP algorithms most specific feature is the error that the neural network gets on its output. The method is used in multi-layer forward network with gradient –descent based on delta –learning rule along with a supervised training learning method. The main objective of this network is to train the network to balance between the input patterns that are used in training and their responses to the input patterns [14].
In back propagation Method, the measure of error function is given by Mean square error 'E'

E=(Target-output)^2…………………equation-1
Partial derivatives of the error function w.r.t the weights
For Output Layer Neurons:

$\delta y =f'(net)(target\ y-output\ y)$……………equation-2

$\partial E/\partial w_{yx}=-\ output_x\delta_y$……………………....equation-3

Where y=output neuron
        x=neuron in last hidden layer
For hidden Layer Neurons:

$\delta y =f'(net).\Sigma(\delta_kw_{ky})$ ………………………equation-4

$\partial E/\partial w_{yx}=-\ output_x\delta_y$ ………………..…..equation-5

Where y=hidden neuron
        x=neuron in previous layer
        k=neuron in next layer [14]

D. Proposed Model
The Proposed Model design steps involved in designing ANN based noisy character recognition system are summarized in the form of a flow chart as shown in figure3
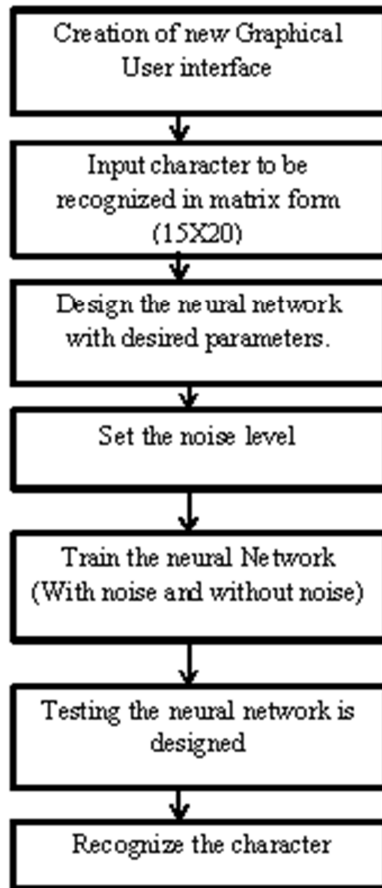
Figure-3: Flow chart for Proposed System

VI.   Results and Discussions

In this section, we describe the experimental results we obtained for the prediction of handwritten characters using proposed handwritten character recognition system. The Handwritten character recognition system has been implemented in Java programming language for the prediction of characters by using the character dataset.

A. Character Dataset

Handwritten Character data are very much essential for this work for training and testing the proposed Handwritten Character Recognition System. Handwritten Character data are to be developed only in the editors in 15X20 matrix form. In English alphabets 26 characters and as well as in numerical 10 characters are to be developed in different forms. For example take the character "A" that character can be written in ten different various forms. For Training the character data we can use Back Propagation Neural Network algorithm. For this reason, the character data is to be taken and recognize what character it is. The following figures describe the characters dataset.
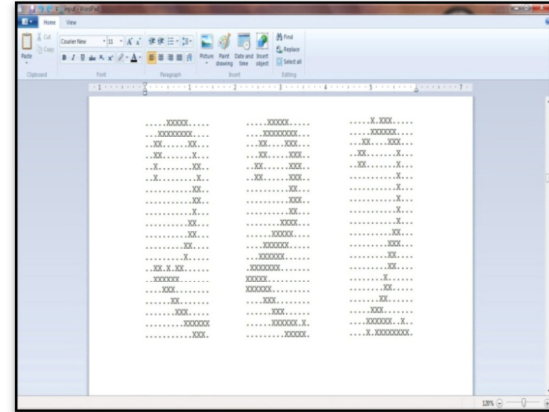

Figure-4: Character Dataset

The screen shots in figures show the experimental results of Handwritten character recognition system when the program ran on character data.

Figure-5 shows the Login window. Login window consists of User name and password. If login is successful we can enter into the system otherwise we can reenter the user name and password until login is to be successful.
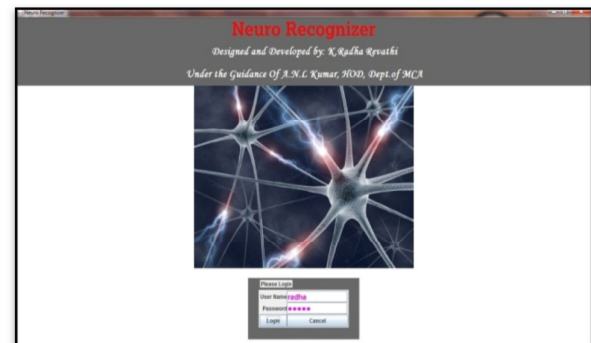

Figure-5: Login window

The below figures represent how the character recognition process is occurred. Figure-6 explains the opening screen that means it consists with all training samples and consists of Image canvas. We can draw the input pattern on the Image canvas.
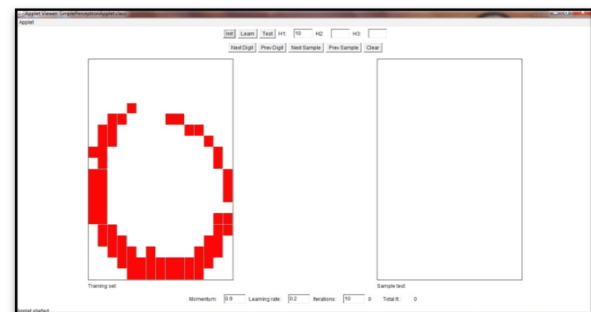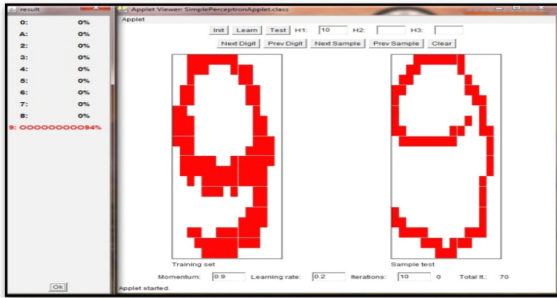

Figure-6: Opening Screen

Figure-7: Character Matching and percentages

Figure-7 explains how the character is to be drawn on the Image canvas and learn the character database and calculate percentage. Based on percentage we can recognize the character. Figure-8 explains how the character set is to be learned based on that we can display the learning graph.
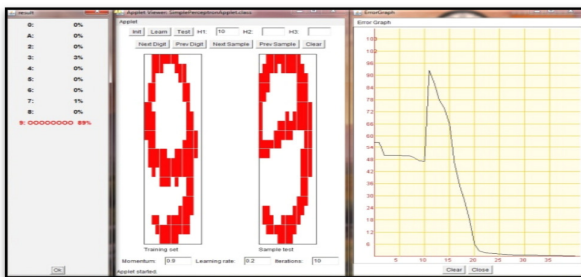

Figure-8: Character Matching and Learning graph

VII.   Conclusion

Neural Network provides an efficient way for character recognition in the presence of noise .In this work, Neural Network is designed & tested for the recognition of handwritten characters under noisy environmental conditions. The results of JAVA Simulation are presented for recognition of handwritten characters. The ANN based system has shown promising results due to the fact that despite being trained only on a single set of templates, it is not only gets trained in 21 seconds, but also can recognize the fonts with high efficiency. This ANN approach is useful method for the recognition of handwritten characters to a great extent. The proposed method has been applied on different unknown characters. Neural network based method gives the accuracy 85 %.

VIII.   Acknowledgement

IX.   References

[1] Mansi Shah and Gordhan b jethava "A literature review on hand written character recognition" indian streams research journal,vol -3 , issue –2, march.**2013**,ISSN:-2230-7850.

[2] Suruchi G. Dedgaonkar, Anjali A. Chandavale, Ashok M. Sapkal "Survey of Methods for Character Recognition" International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 5, May **2012** .

[3] Ankit Sharma,Dipti R Chaudhary "Character Recognition Using Neural Network" International Journal of Engineering Trends and Technology (IJETT) - Volume4Issue4- April **2013**.

[4] Vivek Shrivastava and Navdeep sharma "Artificial Neural Network based optical character recognition" an international journal (sipij) vol.3, no.5, october **2012**.

[5] Anjali Chandavale, Suruchi Dedgaonkar, Dr. Ashok sapkal "An approach for character recognition using pattern matching with ANN "International journal of scientific & engineering research, volume 3, issue 10, october-**2012** 1 ISSN 2229-5518.

[6] J.Pradeep, E.Srinivasan and S.Himavathi "Diagonal based feature extraction for handwritten alphabets recognition system using neural network" International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb **2011**.

[7] N. k. Bose, P. Liangneural "Network fundamentals with graphs, algorithms, and applications".

[8] David Kriesel "A Brief introduction to Neural Networks".

[9] Madhu Shahi, Dr. Anil K Ahlawat, Mr. B.N Pandey " Literature Survey on Offline Recognition of Handwritten Hindi Curve Script Using ANN Approach" International Journal of Scientific and Research Publications, Volume 2, Issue 5, May **2012** 1 ISSN 2250-3153 .

[10] Line Eikvil, "Optical Character Recognition".

[11] Majida Aliabed,"Simplifying Handwritten Characters Recognition Using a Particle Swarm Optimization Approach".

[12] Brandon Maharrey,"A Neural Network Implementation of Optical Character Recognition".

[13] Girish Kumar jha, "Artficial Neural Networks".

[14] prerna Kakkar, Umesh Dutta, "A Novel Approach to Recognition of English Characters Using Artificial Neural Network" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (An ISO 3297: 2007 Certified Organization)Vol. 3, Issue 6, June **2014**.

Author Profile

Ms. Radha Revathi Komanduri is a student of MCA Final year from Swarnandhra college of Engineering and Technology, Narasapur. Her areas of interest is Core Java, Neural Networks, Data Mining. She is very much interested in designing websites, application software.

Mr. A.N.L.Kumar, working as a Head, Department of MCA, Swarnandhra college of Engineering and Technology, Narasapur. His areas of interest is Software Engineering, Image processing, Computer graphics, Neural Networks, Design and Analysis of Algorithms.

Mr. Andey Krishnaji is a project coordinator at Department of Computer Applications in Swarnandhra College of Engineering and Technology. His areas of interest are Neural Networks, Data mining, Fuzzy logic Computational analytics. He is a popular trainer in the areas of Core Java, Advanced Java, c and data structures. He has developed a good number of software applications. He can be reached through his blog www.ijice.org which contains a good number of articles on Technology and Programming.