An Analysis of Effect of Proportionality among Hidden Layers in Standard Back Propagation Neural Network.

Rakesh Kumar Bhujade, Stuti Asthana, Amit Kant Pandit

PhD Research Scholar, CSE Department, Suresh Gyan Vihar University, Jagatpura, Jaipur, Rajasthan, India.
SECE, Shri Mata Vaishno Devi University, Katra, Jammu, India.

Abstract-- **In this paper, an analysis is performed on the importance of proportionality among multiple hidden layers of BPNN. In case of any discrepancies in the network, maximum two layers are enough to train the whole neural network to get the desired result. But in some situations where accuracy is the chief criteria and training data is similar in data sets, (Like in Multiscript Numeral Recognition where shapes of different numerals resemble different values), in such conditions accuracy matters the most. This paper describes a five layer hidden approach to get the most possible accurate results in Multiscript Pin Code Recognition System using MATLAB. In order to get the desired output, the proportionality of hidden layers are semi-optimized to precise level but it may vary depending upon the training data set & the type of problem.**

Index Terms-- **Multiple Hidden layers, Back propagation Neural Network, Multiscript Pin Code Recognition.**

I. INTRODUCTION

This paper analyses the proportionality of neurons in multiple layers of Back Proportion Neural Network. Recognition is a machine learning technique i.e. data mining technique used to find out group membership for data instances. The purpose of neural networks is to make simpler the problems of prediction & recognition. Neural networks are easy models of biological neuron system. However, there are different learning mechanisms exist to enable the neural network acquire knowledge. This biological neuron system is an equally distributed system made up of highly interconnected neurons thereby making it available for use & acquires knowledge. Based on the learning mechanisms & other features neural networks architectures have been classified into various categories. Training has been defined as the method of learning & the ability to solve a problem with the help of knowledge acquired is known as presumption.

The design & configuration of neural network is inspired by Human Brain, it is imitation of central nervous system, hence, performs functions similar to human brain. The functioning & structural constituent of human brain known as neurons performs computations such as cognition, pattern recognition, and logical inference.

Therefore, relating to this the technology which has been built on the structure like human brain has been given the name Artificial Neural Network (ANN) or Artificial Neural System (ANS) or simply neural network. Additional names for this technology are termed as Connectionist networks, Neuro Computers, parallel distributed processors etc. Neural networks are known by other names also neurodes, Processing Elements & nodes.

II. PROPERTIES OF NEURAL NETWORKS

1. The NNs have the capability to map the input & output patterns i.e. they have mapping capabilities, they map the input patterns associated to the desired output patterns.
2. Examples are very important in training process of NNs. Thus, NN architectures could be trained with recognized examples of a problem before they are tested for their inference capability on unknown instances of the problem.
3. Neural networks posses the ability to generalize i.e. they may predict new outcomes from past trends.
4. NNs models real world problems i.e. they are robust & fault tolerant. They may recall full patterns from noisy patterns also.
5. NNs are multitasking also i.e. they may process information in parallel at a high speed & in undistributed manner.

Neural Network Characteristics:

Neural network has the definition in the name itself; the term network describes the interconnection between neurons in the various layers of the system. Every NN system is basically a 3 layered structure Input layer, Hidden layer & Output layer. The input layer has got the input neurons which transfers the data to the hidden layer via synapses. In the same way the hidden layer transfers this data to the output layer via more synapses. The functioning of synapses is to store the value known as weights; it helps in manipulating the input & output to various layers. ANN can be defined on the basis of three characteristics:

- The Architecture: The number of layers & the number of nodes in each of the layers.
- The learning mechanism which is applied for updating the weights of the connections.
- The activation functions have been used in various layers.

Standard BPNN Algorithm

BPNN has the capability of self-learning, which has two parts. The first one is forward transfer of information & the other one is the reverse transfer of error between expected output & the actual output. The structure of BPNN has three layers: Input layer, hidden layer & output layer.
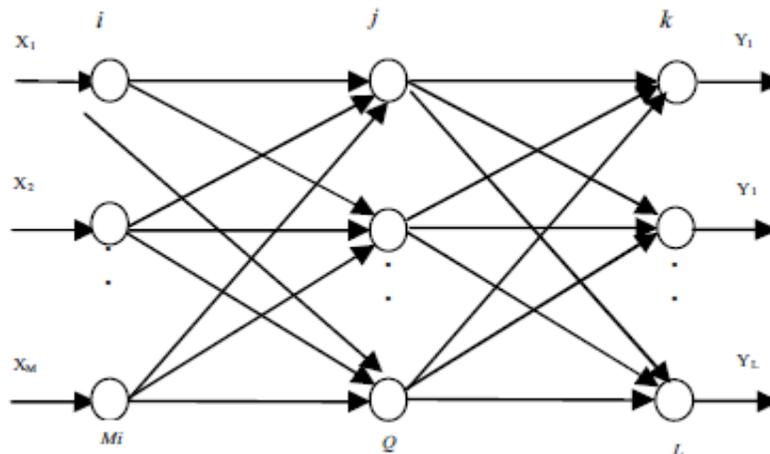
The model of BPNN is shown in figure 1.

Figure 1. BP Neural Network Model

The input layer of network has M neurons, the hidden layer has Q neurons & output layer has L neurons. The input vector of neural network is X = [x1, x2……..xM]. The output vector of neural network is Y = [y1, y2……yL].

The weighted value between input layer & hidden layer is w ij & the weighted output between hidden layer & output layer is w jk. The transfer function of neural network is unipolar sigmoid function which is

$$f(x) = \frac{1}{1 + e^{-x}}$$

The function has the characteristic which is:

$$f'(x) = f(x)\,[1 - f(x)]$$

According to the gradient descent method, the data transmitted from input layer to hidden layer is *l = xl*. After the hidden layer receives data from input layer the first thing to be calculated is the weighted sum which is

$$net\ j = \sum_{i=1}^{M} w_{ij}x_i$$

And then transfer the data to the output layer through the transfer function. The output of hidden layer is

$$0\ j = f(net\ j) = f\left(\sum_{i=1}^{M} w_{ij}x_i\right)$$

III. LITERATURE REVIEW

Many scientists have researched about numeral character recognition till now & gave their reviews about the same. We have prepared the reviews about some of them.

Stephen Trenn considered the sufficiently smooth multivariable functions with a multilayer perceptron(MLP). For a given approximation order, explicit formulas for the necessary number of hidden units & its distributions to the hidden layers of the MLP are derived. The concept of approximation order encompasses Kolmogorov-Gabor polynomials or discrete Volterra series, which are widely used in static & dynamic models of nonlinear systems. His main contribution is that he has given the explicit formula for the necessary number of hidden units in an MLP to achieve a given approximation order. For linear and quadratic approximation, only one layer is needed. The correlation between approximation order & approximation accuracy was studied. A sufficient condition was given for the activation function for which a high approximation order implies high approximation accuracy.

Jinchuan Ke[1] and Xinjhe Liu[2] presents the neural network model, a new procedure and structure in nonlinear interpolation and demonstrates a potential of stock market prediction with incomplete, imprecise & noisy data. It presents a sensitivity analysis of optimal hidden layers and hidden neurons in neural network modeling for stock price prediction. To perform the sensitivity analysis forty cases with various hidden neurons are examined to estimate the training & the generalization errors. During the network training, the weights have to be updated in the forward & backward processes. In one forward & backward epoch, each weight must be perturbed twice & the error must be reevaluated for each perturbation, so the error measurement in each hidden layer takes O(n) complexity for compute time, the total complexity required scales like $O(n^m)$ for m hidden layers, where n is the number of weights in the network. It is very robust & simple to implement.

Diamantino Caseiro and Adrej Ljolje presented techniques for building faster & more accurate recurrent neural network language modeling (RNNLMs). RNNLM have been shown to outperform most other advanced language modeling techniques. An error rate reduction of 5.9% was observed on a state of the art multi-pass voice mail to test ASR system using RNNLMs trained with the proposed algorithm. He also showed that Brown clustering of the vocabulary is much more effective that other techniques. They also presented an algorithm for converting an ensemble of RNNLMs into a single model that can be further tuned or adapted. They had shown that the word clustering algorithm used to decompose the output layer of a RNNLM is critical, and that using algorithms such as Brown clustering that take into account the context of words leads to substantial reductions in perplexity relative to simpler word frequency bining approaches.

When evaluated on a state of the art ASR system, this technique proved to be superior to linear interpolation & achieved a 5.9% error rate reduction relative to the n-gram base line.

Prasanna kumar Muthukumar, Alan W Black introduced Speech Synthesis Systems. Synthesis systems are typically built with speech data & transcriptions. It is tried to build synthesis systems when no transcription or knowledge about the language are available. An automated way of obtaining phones and phonetic knowledge about the corpus is proposed by making use of Articulatory Features (AFs). An Articulator feature predictor is trained on a bootstrap corpus in an arbitrary other language using a three hidden layer neural network. This neural network runs on the speech corpus to extract AFs. Hierarchical clustering is used to cluster the AFs into categories i.e. phones. Phonetic information about each of these inferred phones is obtained by computing the mean of the AFs in each cluster. A Text-to-Speech synthesis system or synthesis in a speech to Speech translation system must work from text rather than raw phones. At this point, the ease with which the inferred phones can be predicted from text is unknown. This is an issue that is planned to investigate in the future work.

Gnana Sheela K and Dr. S.N Deepa analyzed back propagation algorithm for neural networks. The objective is to analyze the parameters such as momentum & learning rate using back propagation algorithm in artificial neural networks. The results are obtained using two variation of back propagation algorithm: simple back propagation & back propagation with momentum. Simulation result shows affect the performance of ANN. Also, Gnana Sheela K and Dr. S.N Deepa proposed a new algorithm to find hidden neurons in Radial basis function networks for wind speed prediction in renewable energy systems. To find the number of hidden neurons, 101 various criteria examined based on the error values of mean squared error, mean absolute percentage error and mean absolute error. The minimal error values are considered as the best solution to find hidden neurons in RBFN. The proposed algorithm is tested on real time wind data.

Yinyin Liu, Janusz A. Starzyk , Zhen Zhu proposed a novel and effective criterion based on the estimation of the signal to noise ratio figure(SNRF) to optimize the number of hidden neurons in neural networks to avoid over fitting in the function approximation. SNRF can quantitively measure the useful information left unlearned so that overfitting can be automatically detected from the training error only without use of a separate validation set. Without a priori knowledge of the noise characteristics of the training data may come with White Gaussian Noise(WGN) at an unknown level.

Shin'ichi Tamura and Masahiko Tateishi introduces Neural network theorems which state that only when there are infinitely many hidden units is a four-layered feed forward neural network equivalent to a three layered neural network. A proof is given showing that a three layered feed forward network with N – 1 hidden units can give any N input target relations exactly. Based on the results of the proof, a four-layered network is constructed & is found to give any N input –target relations with a negligibly small error using only (N/2)+3 hidden units. This clearly shows the difference in capabilities between a four layered feedforward neural network and a three layered feedforward neural network and indicates the superiority of the four layered network for the training of input target pairs.

David hunter and his fellows find out one of the major difficulties researchers facing using neural networks is the selection of the proper size and topology of the networks. A partial solution proposed to this problem is to use the least possible number of neurons along with a large number of training patterns. Neural network architectures vary in complexity & efficiency. The power of neural networks increases significantly with increased number of connections across layers. The most powerful are FCC architectures with the largest number of connections across layers per number of neurons. However, with increased depth more nonlinear transformations are made & this may require higher accuracy of computing and larger sensitivity to noise.

Ms Neha Sahu and Mr. Nitin Kali Raman proposed an efficient handwritten devnagri character recognition system using neural network. Various techniques have been implemented for this problem with many improvements so far but mainly Artificial Neural Network technique is used to preprocess, segment & recognize devanagri characters. The system was designed, implemented, trained & found to exhibit an accuracy of 75.6% on noisy characters. A small set of Devnagri characters using back propagation neural network is trained, then testing was performed on other word set. The accuracy of network was average. These characters are further analyzed & then tested with different set. Again the accuracy is analyzed. This process is repeated then again testing was performed on some new image sets written by different people. It was found that accuracy of the network increases in many cases. As the network is trained with more number of sets, the accuracy of recognition of characters increases.

Jayashree Prasad and Uday.V. Kulkarni implemented Adaptive Neuro Fuzzy Classifier(ANFC) for recognition of isolated handwritten characters of Guajarati. They aimed to compare the performance of ANFC with weighted k – NN classifier proposed by them.

Fuzzy classification is the task of partitioning a feature space into fuzzy classes. Authors used a supervised learning procedure based on Scaled Conjugate Gradient(SCG) algorithm to update parameters in an adaptive network. The recognition rates were recorded for 34 consonants of Guajarati. Maximum recognition efficiency of 88.89% was achieved for few characters using GPXNP feature.

Md. Musfiqur Rahman Sazal, Sujan kumar Biswas, Md. Faijul Amin, and Kazuyuki Murase recognized Bangla handwritten character using Deep Belief Network with a focus on automatic learning of good representations. Among different learning structures, the deep learning structures, we employ the deep belief network (DBN) that takes the raw character images as input and learning proceeds in two steps : an unsupervised feature learning followed by a supervised fine tuning of the network parameters. In this approach, there is no need to create handcrafted features like loops, strokes & curves.

Zhu Dan and Chen Xu introduced Back Propagation Neural Network based on offline handwritten digits recognition system The MNIST database of handwritten digits is applied to train and test the neural network. They introduced the principle component Analysis(PCA) for feature extraction which can improve the performance of the neural network and immensely shorten the training time. Comparisons being made depending on the recognition rate among three methods : neural network method , thirteen features method , Fisher discriminant analysis method. After comparisons , Fisher discriminant analysis method has the lowest recognition rate and proved to be time consuming which is not suitable for smart phone application. BP neural network has itself too many nodes , which costs a lot of time in training . The thirteen features method need to calculate the minimum of the weighted distance from unclassified sample to all the training samples , so it occupies much more RAM. In order to increase the recognition rate we acquire the real-time learning method to adapt to the written style of the specific software user.

Daniel Neil and Shih Chii Liu introduced the Minitaur spiking network accelerator , an event driven neural network accelerator , which is designed for low power and high performance. As an field programmable gate array-based system, it can be integrated into existing robotics or it can offload computationally expensive neural network tasks from the CPU. In addition to the system's performance of 18.73 million PSCs /second, it consumes just 1.5 W of power, enabling it to be used in embedded robotics applications. With proper weights ,the system is remarkably robust to noise.

Jarernsri Mitrpanont and Yingyot Imprasert introduced Thai Handwritten Character Recognition using Heuristic Rules Hybrid Neural Network. In this the neural network technique is applied to to identify and recognize the zigzag pattern. The proposed functions are mainly in Feature extraction enhancement to improve the feature conflict resolution rule and the specialized neural network based zigzag feature extraction and neural network based recognition . The result showed that the additional feature conflict resolution rule could achieve the feature extraction rate of 87.85% , specialized neural network based zigzag extraction could achieve 90.48% , neural network based recognition which combine both of the two proposed feature extraction functions could achieve 92.78%.

Authors Salim Ouchtati, Mohammed Redjimi and Mouldi Bedda proposed an offline system for the recognition of the handwritten numeric chains, their main objective is to determine the parameter vector who gives the best performances. They did their work and divided it into two parts :

The realization of system of isolated digits recognition , based mainly on the evolution of neural network performances , trained with the gradient back propagation algorithm and fed by several parameter vectors.

The extension of our system for the reading of the handwritten numeric chains by using vertical projection to segment the numeric chain at the isolated digit will be presented to the entry of the system achieved in the first part.

Below is the table showing the accuracy in '%' with the respective authors and the published year.

AUTHOR	YEAR	ACCURACY
Ms Neha Sahu & Mr. Nitin Kali Raman	2013	75.6%
Jayashree Rajesh Prasad & Uday V. Kulkarni	2014	86.3%
Md. Musfiqur Rahman Sazal, Sujan kumar Biswas & Md. Faijul Amin	2013	90.2%
Zhu Dan & Chen Xu	2013	91.2%
Daniel Neil & Shih Chii Liu	2013	92%
Jarernsri L. Mitrpanont & Yingyot Imprasert	2011	92.7%
Salim Ouchtati , Mohammed Redjimi & Mouldi Bedda	2014	95.2%
Husam Ahmed Al Hamad	2014	95.3%
Ravindra S Hegadi & Parshuram M. Kamble	2014	97%
M.M Farhad & Ripon Shah Mishu	2014	91.4%

IV. PROPOSED FIVE HIDDEN LAYERS ARCHITECTURE

The five hidden layer neural network architecture is shown in the Fig.2. The architecture shown is five hidden layers fully interconnected with different weights.

The first layer is hidden layer and the input to this layer is present in binary form viz 0, 1.

It is assumed that every numeral has the size of 15*10 so the input layer consists of 150 neurons. The bit patterns are chosen so that they correspond to certain characters e.g. all the boxes in the matrix are numbered from 1 to 35.
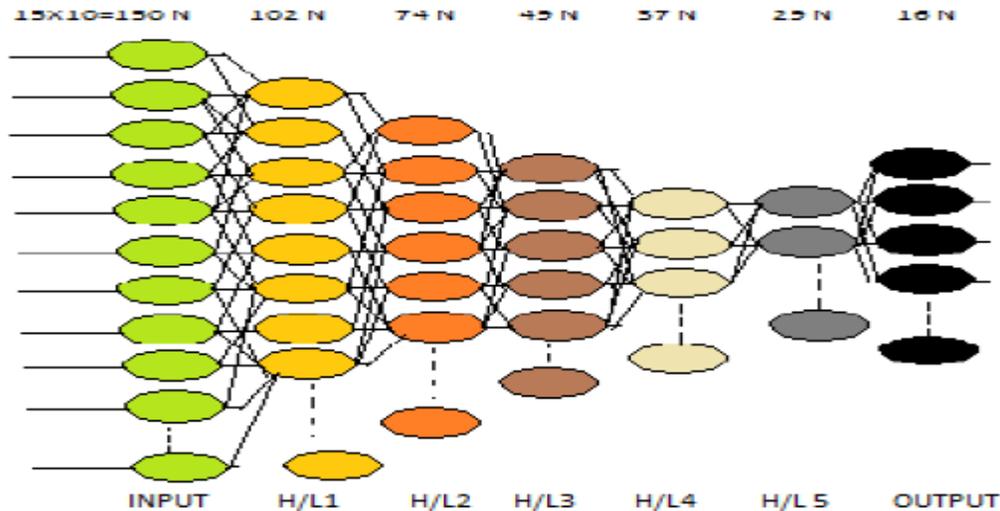

Fig. 2 Shows the five hidden layer architecture.

Now seen the architecture, the input pattern will have the form as shown. Say for character 'C' the pattern will be [1,1], [1,2], [1,3], [1,4], [1,5], [2,1], [3,1], [4,1], [5,1], [6,1], [7,1], [7,2], [7,3], [7,4], [7,5] will contain 1 bit and the rest will contain 0. In this way the input values are framed & fed to the network as input. The output layer has 16 neurons in order to support Unicode. As specified in the thumb rule, 3/4th number of neurons must be in hidden layer and the first hidden layer has 70-75 % neurons. Similarly in second hidden layer, the first hidden layer will act as input layer for it. Hence 75% of 102 is nearly equals 74 and so on for third, fourth & fifth layer. A very minute difference of 2 or 3 neurons can be adjusted while training, so this helps in better architecture of the network.

70-75% neurons of input layer = first hidden layer

70-75% neurons of first hidden layer = second hidden layer

70-75% neurons of second hidden layer = third hidden layer

70-75% neurons of third hidden layer = fourth hidden layer

70-75% neurons of fourth hidden layer = fifth hidden layer

Total neurons in hidden layers = 291 (less than twice of input)

This is proved theoretically if we consider rule of thumb that the network is pruned. But many researchers suggest that least number of hidden layers must be used to save training time & complexity but this will not result correctly when accuracy is the major criteria.

V. PROBLEM FORMULATION

The aim of this study is to figure out the merits of learning as well as recognition rate of handwritten numerals. The main reasons of less recognition rate of numerals is the extraction features from handwritten characters to train neural network so that its recognition accuracy is better and we have studied in the previous works that hidden layer plays a very important role in enhancing the learning rate of neural network & recognition accuracy .

The whole neural network is divided into two phases:

- Training or learning phase
- Recognition phase

VI. PROPOSED METHODOLOGY

Proposed five hidden layers BPNN architecture.

The proposed learning algorithm of five hidden layers BP-NN is explained below:

1. Declare all the variable to create simulation environment.
2. Preprocess the template image of handwritten numerals to train the neural network.
3. Create the neural network with proposed weights and neurons and initialize with proposed modes.
4. Start training of neural network with 5 hidden layers.
5. The training will repeat for given number of epochs.
6. Compare the network output with the target output and keep repeating the training process until the targeted output.
7. After achievement of targeted output save the neural network for further recognition process.

The Recognition algorithm for five hidden layer BP-NN approach:

1. Load the trained neural network for 5 hidden layers.
2. Browse handwritten numerals image for recognition.
3. Select area of image to recognize.
4. Pre-process image after selection.
5. Start comparison process using neural network.
6. After recognition display the recognized numerals.

There are two phases of the work:

1. Training phase
2. Testing phase

Training Phase

In the training phase, the neural network system has been provided with information patterns. The information patterns are in the form of 10*10 matrix . Each element can have only two values either a 0 or a 1.

Testing Phase

In the testing phase, the patterns for the alphabets, numerals or symbols are entered into the system in either correct form with some errors in the element values. These patterns are compared with the ones in the knowledge base that was prepared by the training phase of the network. It recognizes the patterns by comparing the weights of every element in the pattern matrix with the elements in the knowledge base of the corresponding matrix. The number of iterations depends upon the choice of number of hidden layers.

VII. CONCLUSION

In this paper we have reviewed that with the help of multiple hidden layer phenomena how a character recognition program is designed.

After studying & reviewing many approaches to recognize the handwritten numerals it is concluded that BPNN gives best results in comparison to other methodologies. BP-NN performs better results and we have seen keeping the proportionality the performance increases with the increase in hidden layers. It can be declared from the study that if the number of hidden layers increases than three, the recognition rate for handwritten numeral network is higher for various languages e.g Hindi, English, Tamil, Bengali, Arabic etc. In this section more than three hidden layers have been given, which will significantly give high recognition rate and lesser learning time.

REFERENCES

[1] Stephan Trenn, *Multilayer Perceptrons: Approximation order and necessary number of hidden units,* IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 19, NO.5, MAY 2008.

[2] Jinchuan Ke[1,] , Xinjhe Liu[2] , Empirical Analysis of Optimal Hidden Neurons in Neural Network Modeling for Stock Prediction , IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008.

[3] Diamantino Caseiro, Andrej Ljolje 'Multiple parallel hidden layers and other improvements to Recurrent Neural Network Language Modeling' , AT & T Labs Research , 180 Park Avenue , Florham Park , NJ 07982, USA.

[4] Prasanna Kumar Muthukumar, Alan W Black, 'Automatic Discovery of a Phonetic Inventory for Unwritten Languages for Statistical Speech Synthesis.' , IEEE International conference on Acoustic, Speech & Signal Processing(ICASSP).

[5] Gnana Sheela, Dr SN Deepa, 'Analysis of Computing Algorithm using momentum in neural networks'. Journal of computing, Volume 3, Issue 6 , June 2011, ISSN 2151-9167. K. Gnana Sheela, S.N Deepa, 'A new algorithm to find number of hidden neurons in Radial Basis Function Network for wind speed prediction in Renewable Energy Systems', CEAI, Vol.15, NO. 3, pp. 30-37, 2013.

[6] Yinyin Liu, Janusz A. Starzyk , Zhen Zhu , 'Optimizing number of hidden neurons in neural networks.' , School of Electrical Engineering and Computer Science, Ohio University.

[7] Shin'ichi Tamura, Masahiko Tateishi , 'Capabilities of a four layered feed forward Neural Network : Four layers versus Three' , IEEE transactions on neural networks, Vol. 8 , No.2 , March 1997.

[8] David Hunter , Hao Yu , Michael S. Pukish Janusz Kolbusz, Bogdan M. Wilamowski, "Selection of proper neural network sizes and Architectures" , IEEE transactions on Industrial Informatics, Vol. 8 , No.2 , May 2012.

[9] Ms. Neha Sahu & Mr. Nitin Kali Raman, 'An Efficient Handwritten Devnagari Character Recognition System Using Neural Network'.

[10] Jayashree Rajesh Prasad, Uday V. Kulkarni, 'Guajarati Character Recognition using Adaptive Neuro fuzzy Classifier' , 2014 International Conference on Electronic Systems, Signal processing and Computing Technologies.

[11] Md. Musfiqur Rahman Sazal, Sujan kumar Biswas, Md. Faijul Amin and Kuzuyuki Murase, 'Bangla Handwritten Character Recognition using Deep Belief Network' , International Conference on Electrical Information and Communication Technology(EICT) 2013.

[12] hu Dan and Chen Xu , 'The Recognition of Handwritten Digits based on BP Neural network and the implementation on Android' , 2013 Third International conference on Intelligent System Design and Engineering Application.

[13] Jarernsri L. Mitrpanont , Yingyot Imprasert , ' Thai Handwritten Character Recognition using Heuristic Rules Hybrid with Neural network' , 2011 Eighth International Joint Conference on computer science and Software engineering (JCCSSE).

[14] Salim Ouchtati , Mohammed Redjimi , Mouldi Bedda , 'Realization of an offline system for the Recognition of the Handwritten Numeric Chains'.