International Engineering Journal For Research & Development
Vol.1
Issue 2

AN ANALYSIS OF HANDWRITTEN CHARACTER RECOGNITION USING NEURAL NETWORK AND STATISTICAL METHOD

Sangita Bannore
Department of Electronics Engineering. Government Polytechnic, Thane Maharashtra, India
sangita_bannore@yahoo.com

ABSTRACT

Handwritten Character recognition is a critical area of research. In this paper pattern classification method for recognizing handwritten numerals in the database using Neural Network and Statistical method is presented. MATLAB has been used as a programming Tool. The data file consists of 15,000 handwritten numerals labelled into ten classes from 0 to 9 each of size 30 X 30 pixels. 600 dimensional depictions of each numeral image is cumbersome, hence few features are extracted from the numerals that best represent numerals and give desirable accuracy. The present study is focused on two methods for pattern classification. One is neural networks and the other is statistical approach. But more concentration is on neural networks, the statistical method used here is just for comparison.

The most widely used Multilayer Perceptron network with one hidden layer is selected for classification in neural networks, whereas Bayesian theory is selected for Statistical methods. Six features are extracted from the primary data set by calculating the skewness, kurtosis, mean, standard deviation, variance, and entropy for the X and Y coordinates of every image. The vector of obtained patterns is divided into training dataset and testing dataset. First dataset is used to coach the classifier using a back propagation training algorithm, whereas the second one to test the performance of the classifier. The results obtained from the neural network are then compared with the Bayesian classifier. It has been found that accuracy to classify the numerals is achieved more using neural networks as compared to statistical methods.

Keywords—Neural network, Multilayer perceptron network, Bayesian theory, Back propagation Training algorithm.

I. INTRODUCTION

Writing, which has been the most natural mode of collecting, storing, and transmitting information through the centuries, now serves not only for the communication among humans, but also for the communication of humans and machines. Since the advent of digital computers, machine simulation of human reading has become the subject of intensive research. A large number of references can be found at survey studies published on this topic. The reason for such was not only because of its challenge on simulation of human reading, but also, because it provides efficient applications e.g. for the automatic processing of bulk amount of papers i.e. transferring data into machines, such as bank cheques, commercial forms, government records, credit card imprints, and postal code reading. The ultimate success of neural networks depends on effectiveness in solving a variety of real life classification or recognition problems. Recent technical reports have demonstrated that neural networks can perform this task with the state of accuracy. Various people have written digits using a variety of different sizes, writing styles, instruments and with a widely varying accuracy. the majority of people write digits in such a way that it's very difficult to classify.

II. RELATED WORK

The field of Neural Networks has a history of some five decades but has found solid applications only in the past fifteen years and the field is still developing rapidly. Thus it is distinctly different from the fields of control systems or optimization where the terminology, basic mathematics, and design procedures have been firmly established and applied for many years. It is not to review the Neural Network as simply a summary of established procedures that are known to work well. Rather, it is a useful tool for Industry, Education, and Research, a tool that helps users find what works and what doesn't and will help develop and extend the field of Neural Networks.

Various methods for feature extraction and handwritten digit recognition were mentioned in the literature.

In [1], a pattern categorizing method for recognizing handwritten character digits in MATLAB is proposed. Designing a pattern identification system means establishing a mapping from measurement space into the space of potential labels. The basic operations in pattern identifications are feature unsheathing and categorizing. As proposed by the litterateur, the database consists of 15,000 handwritten digits labelled into 10 classes from 0 to 9, each of size 20 x 30 pixels. Each digit I considered as one sample in 600 dimensional features space and the data set is converted into a vector of patterns: array of features and array of labels (targets) for each feature sample.

Each digit is mapped to a more manageable low-dimensional feature space i.e. extract some numerical features from the images. Two numerical features are said to be extracted from the primary data set by calculating the skewness for x and y coordinates of each image. The vector of obtained patterns is split into training data set and the test data set. First data set is used to train the classifier whereas the second one to test the performance of the classifier.

In [2], the litterateur has extracted features from the image by adopting Fourier Transform as an appearance presentation scheme and derived 8 characteristic maps for describing different characteristics of co-occurrence relevance of image pixels in each channel of the colour space. Then they calculate the first and second moments of these maps as a depiction of the natural colour image pixel distribution, resulting in a 48-dimensional feature vector. The novel low-level feature is named colour appearance moments (CTM), which can also be considered as a certain augmentation to colour moments in 8 aspects through eight orthogonal templates. Their experiments show that this new feature can achieve good retrieval performance for Content Based Image Retrieval (CBIR). Primary issue in designing a CBIR system is to choose the most powerful image features to depict image contents.

Currently, the largely used features involve colour features, such as colour instants, colour histogram, and appearance features, such as Gabor wavelet feature MR-SAR. As the colour and appearance features capture different aspects of images, their combination may be useful. Therefore, some frontier works attempt to distinguish the colour and appearance information of an image in one feature depiction. In this paper, they propose a novel low-level feature, named Colour appearance moments, for representing image contents. It is able to integrate the colour and appearance characteristics of an image in one Compact form. The primary image is operated with eight templates derived from LFT and eight characteristic maps are obtained each of which characterizes some information on certain aspects of the primary image. Let {I(x,y) | x=0………....L-1, y=0……..…M-1)} denote the primary image. The eight-neighbourhood of pixels (x,y) are N_0, N_1, N_2, N_3, N_4, N_5, N_6, N_7 with anti-clockwise order. Assuming it is a periodic sequence with period of eight, we denote it as $I(x,y,n) = N_n$, n = 0 to 7. Intuitively, alike local parts of the appearance have alike series of I(x,y,n)and their Fourier transform coefficient in frequency field are alike correspondingly. So, it is possible to utilize the local Fourier Transform to extract features for representing the local grey-tone spatial dependency. On the other hand, the local Fourier Transform is equivalent to eight unique templates operating on the image respectively.

In [3], the litterateurs have presented an approach for handwritten digit recognition using perceptron neural networks. A single layer perceptron has been selected and trained to recognize handwritten digits. Database is obtained from UCT machine learning repository which consists of 1,934 cases of numeric digits. Each training case is in digitized format of 0's and 1's at a resolution of 32 x 32. The perceptron is designed to accept 1024 input units (32 x 32) = 1024 and 10 output units. Binary step function has been used as a perceptron activation function. For simplicity, default weights and bias (b) are set to 0 as well as learning rates (α) is set to 1. The network is then trained for the given dataset and a separate testing dataset. To get an unbiased estimate of the prediction error rate of the algorithm, ten-fold cross validation is done to evaluate the prediction error rate. Results are reported in terms of cross validated testing error rate. For the 32 x 32 map format the error rate is 9.98%, for the 8 X 8 block format, error rate is 11.07%. They have concluded that the algorithm runs best (in terms of lowest error rate). When setting threshold to zero and training directly on the primary bitmap format. Allowing non-zero threshold does not improve perceptron learning, nor does dimension reduction, although the later speeds up computation.

In [4], the litterateurs have proposed a new genetic learning algorithm to produce a nearby optimal feed forward neural network innovatively for the task of handwritten digit identification. Back propagation training algorithm is used. The major coverage of their paper is two-field. On one hand a systematic feature extraction method using orthogonal moments is presented, and the lofty robustness in appearance of noise and their use of universal information instead of neighbourhood information of the image. The suggested approach explores the application of moment method to assess a set of candidate features and to pick an information subset to be supplied as input data for Neural Network classifier. On the other hand, they have identified that network technology has a significant impact on the speed and execution of back propagation trained networks.

In [5], the litterateur has explored the application of feed forward neural networks to the goal of written number recognition. Differently occurring forms of numerals are explained to classify the digits. Good survey of recent adaptive system research results can be found according to the litterateur, where the performance of neural network techniques is contrasted to those of procedures such as Principal Component Analysis (PCA). Database is procured from the National Institute of Science and Technology. One concealed layer is used having function tan-sigmoid and output layer also having function tan-sigmoid. He comments that the numeral neurons allow the system to indefinite functions of larger complexity. The embryonic condition of the network also has a powerful influence on the training. The steepest fall algorithm is relatively sluggish and has a tendency to get in the domestic minima of the error surface (where the gradient is zero). Litterateur has implemented both the steepest fall back propagation and a most urbane method called conjugate-gradient back propagation.

In [6], a feed forward neural network with one concealed layer is used to understand handwritten digits. This problem is predominant in applications such as automatic zip code scanning on letters which is in present use by the United States postal service. A 5-fold cross validation testing is done to estimate the performance of a network. The input file is divided into 5 parts, each containing 100/5 = 20 instances. Out of 5 runs, 4 of the parts are used as training sets and the fifth part as test sets. Training is done for at least 200 epochs. After every 10 epochs, the sum of MSE on the training is computed. The training and evaluating steps are repeated after varying the number of concealed units in the network. The training plot and outcomes of 5-fold cross validation for each of these networks using the same training and evaluating sets are examined. A distinct graph has to be used for each network. The litterateur has identified what he considers to be the finest set of parameters (alpha, number of hidden units, number of iterations.)

III. METHODOLOGY

The working of the proposed system useful for predicting handwritten digit recognition is as shown in fig 1. Following are the pre-processing steps applied.

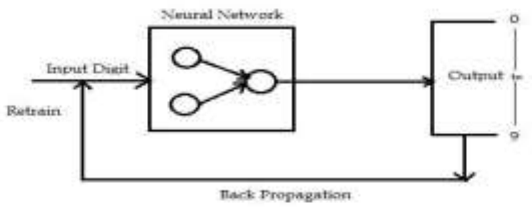

Fig.1. Working of the proposed Model

A. Feature Extraction:

It is the process of generating features to be used in the classification task. To accomplish the highest accuracy Feature extraction process should choose optimal subset of features from the dataset. Hence feature extraction procedure consists of mapping of M – dimensional vectors of consideration Y = (y1, y2, ….. ym) into a new K – dimensional feature interpretation X = (x1, x2,……… xk), K < M, which is more suitable for a given task. There are many algorithms for feature extraction. A few examples are:

i) Variance of digit along the x and y coordinates. Choose the mass '1' for black pixels and mass '0' pixels.

ii) High order moments for the x and y coordinates (Skewness and Kurtosis) – subdivide the digit image into several equal quadrants (2,4,8, etc.) and count black pixels in each quadrant.

iii) Visualize any pair of extracted features (if they are more than 2) as a 2-D scatter plot using different classes. Use the scatter command. Comment about separability of different classes in the dataset. Think, if the extracted features are invarient to any kind of transformations (e.g. translation, rotation, scale).

iv) Extract two numerical features from the original data set by calculating the skewness for x and y coordinates of each image. These will be the features used for further classification.

v) Divide the vector of obtained patterns into a training set and the test data set containing the rest of the data. First data set will be used to train the classifier, whereas the second one – to test the performance of the classifier.

B. Image Processing:

Four steps involved in extracting the features are

1) Draw the digit with the help of a mouse or read it from the database.
2) Convert it into binary image only when the digit is drawn with the help of a mouse.
3) Crop image.
4) Resize image.

IMAGE CROPPING: Creating a new image from a part of the primary image is known as image cropping. To extract a rectangular portion of image, 'imcrop' function accepts two primary arguments.

1) Image to be cropped.
2) Coordinates of a rectangle that define the crop area.

The 'imcrop' function displays the image in a figure window and waits to draw the cropping rectangle on the image. Program then calls 'inshow' to view the crop image.

RESIZING IMAGE: The image in the database of size 20 x 30 matrix. To extract the features more accurately, the image is resized by four times i.e. 80 into 120 matrix. Features are extracted from the digit image in the following manner.

If {I(x,y)/x=0……………L-1, y=0……………..M-1)} denote the primary image, the 8 neighbourhoods of pixel (x,y) are N_0, N_1, N_2, N_3, N_4, N_5, N_6, N_7.

f(x-1, y-1)	f(x-1, y)	f(x-1, y+1)
f(x, y-1)	f(x, y)	f(x, y+1)
f(x+1, y-1)	f(x+1, y)	f(x+1, y+1)

Table 1: Positive Review Table

Supposing it a periodic sequence with a period of eight, we denote it as follows.

$$I(x,y,n) = N_n \quad 0 \le n \le 7$$

If we consider any pixel, it has 8 neighbours. e.g. pixel f(x,y) has its 8 neighbours as shown in the Fig.2.

IV. RESULT AND ANALYSIS

Reliability of the neural network pattern recognition system is calculated by testing 930/2 i.e. 465 of input vectors with changing proportions of noise. Testing the application on the neural network is a foremost step to help ensure that the digit recognition is done properly.

A. *Validation:*

In simple validations we unknowingly split the set of labelled training specimens into two parts: one is used as the traditional set for fine tuning model parameters in the classifier. The other set is the validation set used to approximate the generalization error. Since our final goal is low generalization error, we train the classifier until we reach a lowest of this validation error. In general, validation is 100%, if the training data presented for training is completely representative of the feature of the data presented. But this is not the case because data is gathered using noisy measurements and training data can not be prototypical of all the features that will be presented. An easy generalization of the above method is m-fold cross-validation. Here almost 500 digits are utilized for validation.

B. *Accuracy:*

One thousand noisy digits are selected from the database for testing. Neural network is trained for over 7000 digits. Learning rate selected is 0.3. The network was trained only for one thousand digits initially. Then the validation was performed on a few digits and accuracy checked. The network was not giving proper accuracy, hence it was trained further. The network was trained till the error rate for recognizing digits is 0.001 and the convergence rate is 0.7. Convergence rate means the weights do not change significantly (or change by a very small value) even if trained for a longer time and the network is said to converge to a stable state. Testing set of 100 digits is divided into 10 groups, each of 100 digits. In each group of these hundred digits the total number of each numeral is counted, for example in range 0-100 the number of one's and is 12 or in the range 501-600 the total number of two's is 10 etc. After this the digits are tested for correct classification using both Neural Network and Statistical Method. Accuracy by both methods is then calculated. Bar graph for comparing the accuracy for each numeral in the range 0-1000 are then plotted. Average percentage of accuracy of each numeral is then calculated. Ultimately a graph for comparing the accuracy of both the methods is plotted using plot function. The digit selection, feature extraction, training, testing, validation, performance, regression and accuracy plots are shown in the following figures 2-4 and tables 2-3.

Range of Digits	Total No. of Ones	Classification by NN	Percentage of Accuracy	Statistical Method	Percentage of Accuracy
1-100	8	6	75	8	100
101-200	6	5	83	5	83
201-300	9	9	100	8	89
301-400	12	10	83	8	67
401-500	13	10	77	8	62
501-600	10	8	80	7	70
601-700	14	10	71	8	57
701-800	12	10	83	8	67
801-900	7	7	100	7	100
901-1000	11	10	91	9	82

Table 2: Accuracy of Classification for Numerical One

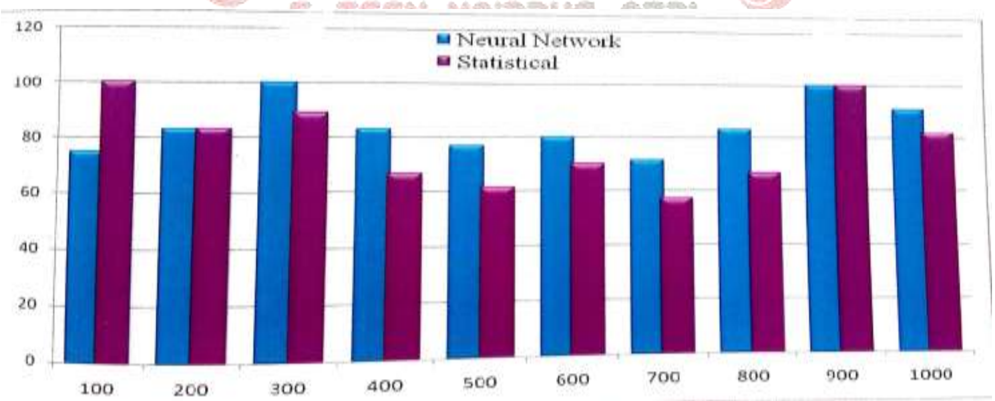

Fig 2. Comparison of Accuracy of Neural Network and Statistical Method for Numeric One.

Range of Digits	Total No. of Nine	Classification by NN	Percentage of Accuracy	Statistical Method	Percentage of Accuracy
1-100	15	10	67	12	80
101-200	8	8	100	7	88
201-300	9	7	78	9	100

301-400	14	11	79	12	86
401-500	6	6	100	6	100
501-600	15	12	80	10	67
601-700	6	6	100	5	83
701-800	5	5	100	4	80
801-900	10	9	90	8	80
901-1000	8	7	88	5	63

Table 3: Accuracy of Classification for Numerical Nine

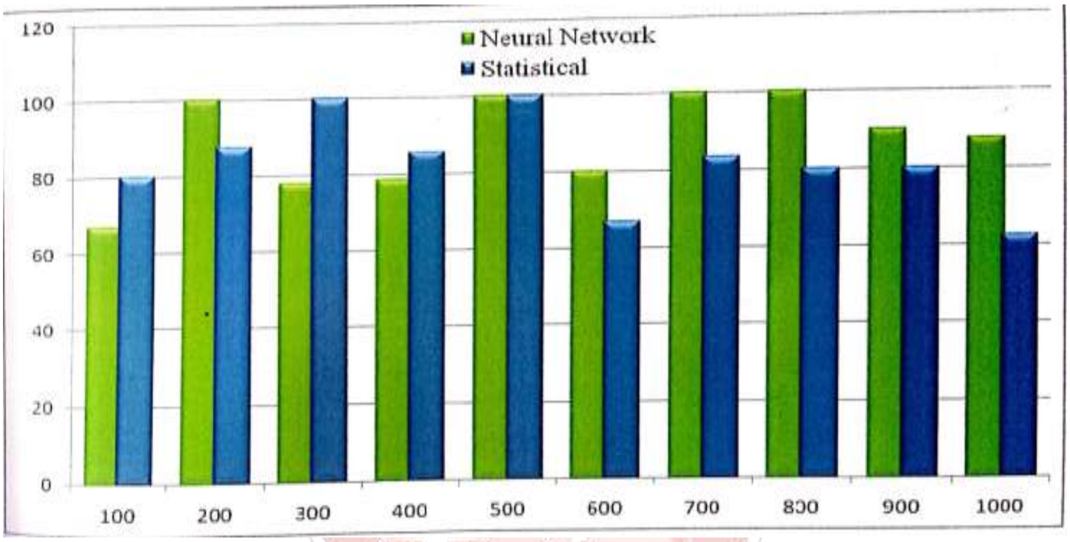

Fig 3. Comparison of Accuracy of Neural Network and Statistical Method for Numeric Nine.

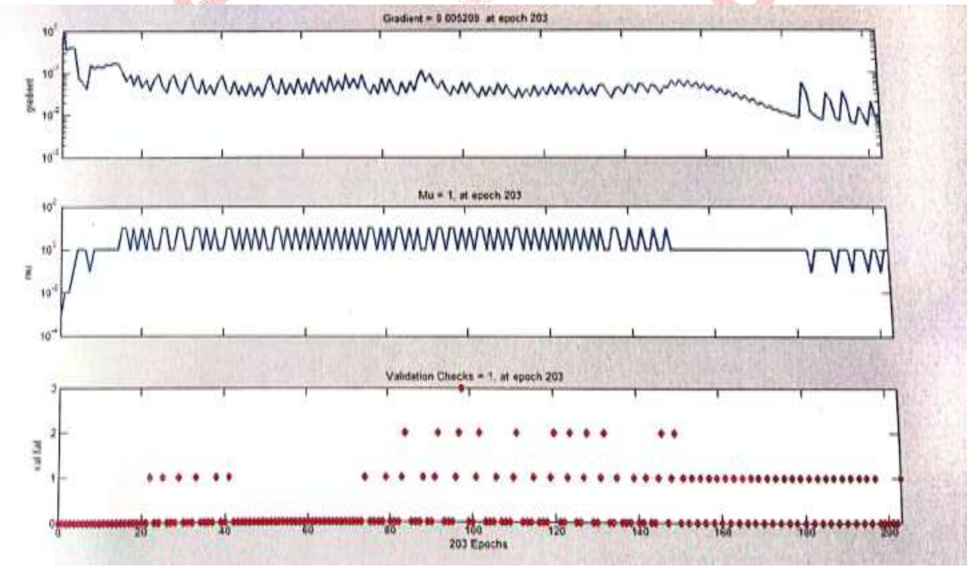

Fig 4. Plot Showing the Training State

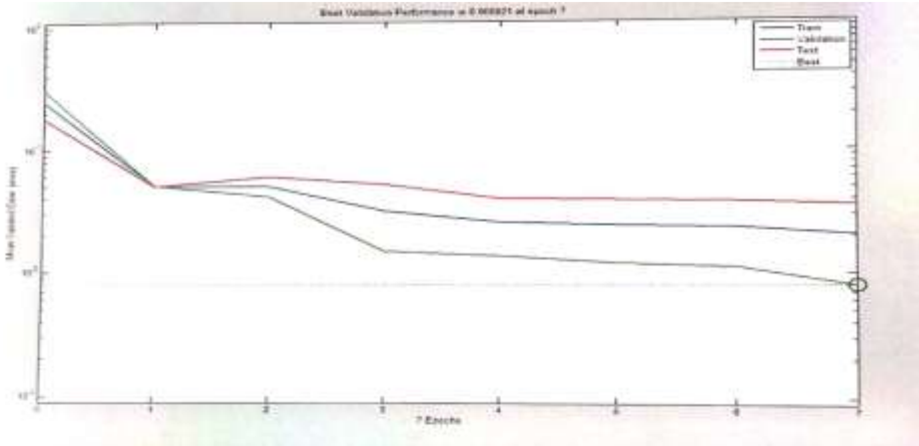

Fig 5. Performance Plot of Mean Square Error for Training, Testing, and Validation.

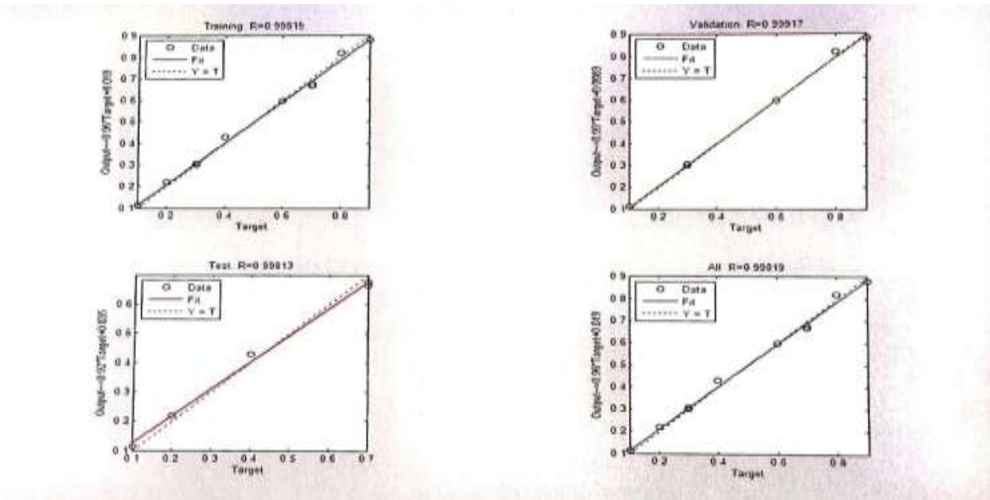

Fig 6. Regression Plot for Training, Testing and Validation.

V. CONCLUSION

In this paper, Neural Network has been successfully used to solve the problem of handwritten digit identification. The most popularly used multi-layer perceptron back propagation algorithm has been practically endeavoured. The network is instructed using the back propagation algorithm up-to the error rate for identifying digits is 0.01 and the convergence rate is 0.7. Three attributes were withdrawn initially i.e. skewness, trained mean and standard deviation, but the results were not so accurate, hence features were increased from three to six i.e. kurtosis, variance and entropy. The data file of 15,000 digits was developed out of which 7,000 are used for training, 1,000 are used for validation and remaining are used for testing.

If more accuracy is required, the network can be trained for a prolonged time or retrained with additional neurons in its concealed layer. The number of hidden layers can also be changed. Finally, handwritten digit recognition has also been done using statistical methods, Bayesian decision theory is selected for this purpose. Graphs are plotted for accuracy of each numeral (0 to 9) for neural network as well as statistical method and it is found that accuracy of each numeral is different. Finally, the graph is plotted for comparison of neural networks and statistical methods. The project is more focused on neural networks, a statistical method is used here for comparison and it has been found that accuracy of neural network is better than statistical method.

In future, the proposed system can be designed for character A to Z recognition, face recognition, and also signature verification on cheques.

REFERENCES

[1] Mikko, Koivuluoma, 'Statistical Pattern Classification – Recognition of Handwritten Digits", SGN-1600 Signaalinkäsittelyn ja multimedian työkurssi, 2003.

[2] Hui Yu, Mingjing Li, HongJiangZhang, Jufu Feng, "Color appearance Moments for Content Based Image Retrieval", Centre for Information Science, Peking University, Beijing, China, IEEE 2002.

[3] Yun Lan& Sean Lee, "Handwritten Digit Recognition Using Perceptron Neural Network", Department of Information and Computer Science, Irvine, California.

[4] H. El, Fadili, K.Zenkouar, H.Qjidaa "Evolving Neural Networks Using Moment Method for Handwritten Digit Recognition", PWASET Vol 6, June 2005.

[5] Firas Hamze, "Handwritten Neural Recognition Using Neural Network", IEEE 368, Stanford University, 27 May 2000.

[6] CS170 Program # 2: Handwritten Digit Recognition Using Neural Network", March 2003.

[7] Neural Network ToolBox in MATLAB 7.4.

[8] S.N. Sivanandam, S.Sumathi, S.N. Deepa, "Introduction to Neural Network Using MATLAB 6.0', McGraw-Hill Publication 2006.

[9] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Second Edition.