World Applied Sciences Journal 9 (8): 898-906, 2010
ISSN 1818-4952

Improvement in Performance of Neural Network for Persian Handwritten Digits Recognition Using FCM Clustering

[1]S. Mohammad Razavi, [2]Mehran Taghipour and [3]E. Kabir

[1]Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran
[2]Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran
[3]Faculty of Electrical and Computer Engineering, Tarbiat Modarres University, Tehran, Iran

Abstract: In this paper, the approach has been proposed to optimize performance of MLP neural network in Farsi handwritten digits recognition. In proposed approach, data of Farsi handwritten digits have been clustered using Fuzzy C-Means (FCM) and also, membership value of each digit belonged to each cluster has been used in neural network learn and then in test step, data of new and unknown pattern are applied to trained neural network and then input new pattern will be assigned to a class that amount of corresponded neuron to that class is maximum in network's output. Obtained results of using introduced method show that with help of this approach we can reduce the rate of misclassification with respect to other common approaches. Also, by using proposed method, successful rate of recognition becomes 97.2 %.

Key words: Multi Layer Perceptron (MLP) neural network · Fuzzy C-Means (FCM) clustering · Recognition of Farsi handwritten digits · Loci feature

INTRODUCTION

Nowadays, one of the most important problems that may occur in handwritten documents, is error in read them that most of the times is due to similarity between two or more letters to each other or between numbers. Sometimes it can make big mistake for people and may confuse them. Therefore, correct and accurate recognition of this kind of numbers and letters seems essential.

In this paper, recognition of Farsi handwritten digits (as one of the common language in the world) has been taken into consideration and according to much similarity between Farsi and Arabic digits, it can increase worth of this paper.

Nowadays, neural networks and Fuzzy approach and their combination are used in pattern recognition and other fields of engineering such as; estimation of saturation thermodynamic properties, solving Different Pattern Classification Problems, reduction of the Codebook Search Time, fault detection and etcetera [1-4] in order to pattern recognize in the field of handwritten digits different approach can be used. For instance, in [5] some results of handwritten Bangla and Farsi numeral/digit recognition on binary and gray-scale images have been presented. Introduced approach in this paper is based on gradient direction histograms: it works on both binary and gray-scale images. Some advanced classifiers are used for classification of the feature vectors. In [6], at first the proposed engine extracts the numbers' features through the holistic approach which compares the unknown character's features with the features of the existing characters that itself is characterized through Mamdani inference engine on fuzzy rules which is largely enhanced with a multi layer perceptron neural network's learning on features of the different fonts' characters which leads to more comprehensive recognition of Farsi numeral characters in the proposed system. In this paper, a neuro-fuzzy system to recognize the printed Farsi numeral characters has been utilized, considering 33 different Farsi fonts. Also, a new technique for the recognition of optical off-line handwritten Arabic (Indian) numerals using hidden Markov models (HMM) has been proposed in [7]. The character recognition system used in this work is based on three general feature categories. Gradient, structural and concavity features at the sub-regions level are extracted and used as the features for the Arabic (Indian) numeral. The features were chosen because they are some- what orthogonal and are at different scales to each other.

Corresponding Author: Mehran Taghipour, Department of Electrical & Computer Engineering, University of Birjand, Birjand, Iran.
Tel: +98 561 2227795; Fax: +98 561 2227795; E-mail: mtaghipour@birjand.ac.ir.

In [8], a method for recognition of Persian handwritten digits based on characterization loci and mixture of experts is proposed. This method utilizes the characterization loci, as the main feature. In this project we propose a method for recognition of Persian handwritten digits based on characterization loci and mixture of experts. This method utilizes the characterization loci, as the main feature and the mixture of experts as classification stage. Three approaches of implementing and constructing the fuzzy neural networks, neural networks based fuzzy logic, have been discussed in [9] and an improved genetic algorithm, which is a special neural networks learning algorithm, has been proposed. In [10] introduces the artificial neural network group-based adaptive tolerance (GAT) tree model for translation-invariant face recognition has been proposed which is suitable for use in an airport security system. GAT trees use a two-stage divide-and-conquer tree-type approach.

Also, in [11], a Time-Delay Neural Network (TDNN) approach has been introduced to phoneme recognition in this paper characterization is fulfilled by two properties: one is using a 3 layer arrangement of simple computing units, second is that arrangement of time-delay enables the network to discover acoustic-phonetic features and the temporal relationships between them in depend of position in time. The ability of learning networks to generalize can be greatly enhanced by providing constraints from the task domain. In [12] it is demonstrated that how such constraints can be integrated into a backpropagation network through the architecture of the network. A single network learns the entire recognition operation, going from the normalized image of the character to the final classification. Also other implementation of ANN can be seen in [13-15]. Beside mentioned digit recognition approaches in [16], in order to handwritten word recognition two hybrids fuzzy neural systems are developed and applied. Each hybrid is a cascaded system. The first stage of both is a self-organizing feature map (SOFM). The second stages map distances into membership values. The third stage of one system is a multilayer perceptron (MLP). The third stage of the other is a bank of Choquet fuzzy integrals (FI). The two systems are compared individually and as a combination to the baseline system.

The approach that used more in pattern recognition is Multilayer perceptron (MLP) neural network which consists of some layers; each layer is composed of nodes and in the fully connected network considered here each node connects to every node in subsequent layers. Each MLP is composed of a minimum of three layers consisting of an input layer, one or more hidden layers and an output

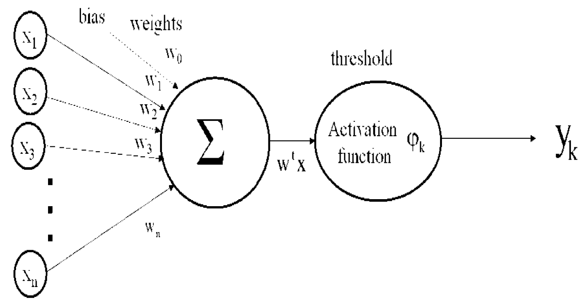


Fig. 1: One Neuron in MLP

layer [17].the inputs data are distributed to subsequent layers by the input layer. Each hidden unit node and output node have thresholds associated with them in addition to the weights. Also, the hidden unit nodes have nonlinear activation functions and the outputs have linear activation functions that its equation can be seen in equation (1) and figure of one neuron in MLP has been shown in Figure (1).

$$y_k = \varphi_k \sum_{i=1}^{m} w_i x_i \tag{1}$$

Where, X_i is input data to NN or output data of previous layer, w_is are synaptic weights between i^{th} neuron of previous layer and k^{th} neuron of next layer. φ_k is activation function that exist in different form.

The way that we can use MLP networks is that: the extracted features are applied to input layers neuron as input and in output for each class is determined one neuron in external layer. Addition to input layers and external layers middle layers can be existed.

In training step of network, each pattern that is applied to network, one target vector is assigned for it. In target vector the element corresponded to neuron of input class has a value equal to one and others are equal to zero. Output of network is calculated for input pattern and then difference between output of network and target vector is used to verify network's weights (the Error Back Propagation algorithm). In test step, features of unknown pattern are applied to neural network and then this pattern will be assigned to a class that amount of corresponded neuron to that class is maximum in network's output.

In each pattern recognition problem involved scattered samples in one class, we can consider sub-classes for samples of this class and we can do classification based on them. For example;

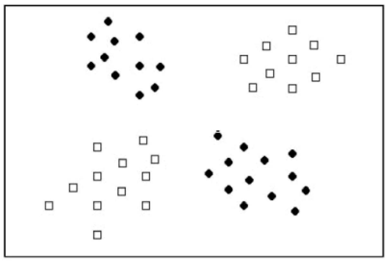


Fig. 2: A Two-class problem

in Figure (2), a Two-class problem has been illustrated, suppose that black samples belong to class 1 and the blank samples belong to class 2. If for each class, we consider two sub-classes, we will have 4-class problem and recognition of unknown pattern will be done easily.

Now, if we want to solve above problem with neural network, we can use a neural network with 4 neurons in output layer and according to what is said before, we can train network and to recognize unknown pattern use it.

But our proposed approach in this paper is that don't consider sub-classes related to the main class as a completely independent class. For this purpose, we can use Fuzzy clustering. To do proposed method, MATLAB software has been used, in order to cluster data, FCM function has been utilized and also, structure of neural network has been made in MATLAB programming.

It is mentionable that in section 2, the proposed procedure is discussed in detail and in section 3, we use this procedure for recognition of Farsi handwritten digits and at the end in section 4, we present conclusion and our recommendations.

FCM Clustering for Optimizing Network's Operation

Fuzzy C-Means (FCM) Clustering: C-Means algorithm is most widely used algorithm for data clustering, Fuzzy C-Means (FCM) is one of these kinds of algorithm which is based on fuzzy approach and rules [18]. In FCM sample data are divided into c clusters, so that, number of clusters are determined before. In FCM, objective function is as follow:

$$J = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} d_{ik}^{2} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} \| x_k - v_i \|^2 \quad (1)$$

Where, m is real number which is bigger than unit, in most of the time m is chosen equal to 2. Also, x_k is k^{th} sample, v_i is center of i^{th} cluster and n is number of sample data. In above equation, u_{ik} shows membership degree of i^{th} sample in k^{th} cluster. ||.|| is similarity (distance) value of sample with (from) center of cluster.

$$\sum_{i=1}^{c} u_{ik} = 1 \; \forall \; k = 1, \ldots, n \quad (2)$$

Concept of above constraint is that summation of membership degree to c clusters of each sample should be equal to unit. In order to obtain equations related to u_{ik} and v_i, defined objective function should be minimized. By using above constraint and putting derivative of objective function equal to zero, we will have:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}} \quad (4)$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} \quad (5)$$

By using two calculated formulas, steps of FCM are as follow:

- Step1) determining initial value for m, C and U^0, Then, initial clusters are determined randomly.
- Step2) determining center of clusters (calculation of v_is).
- Step3) calculating membership matrix based on calculated clusters in step 2.
- Step4) if $\|U^{l+1}-U^{l}\| \le \varepsilon$ algorithm will be stopped else go step 2.

Proposed Approach to Use FCM: In the pattern recognition problem that one or more class of it have samples with high scattering, we can divide training samples related to above classes into two or more clusters and consider each one of cluster like a sub-class, for example in Farsi, digit "6" may be written in 3 forms as can be seen in Table (1) and can be considered one sub-class for each form of this digit.

In order to link subclasses (related to one main class) to each other, Fuzzy clustering has been used. So, the membership value that Fuzzy clustering assign to each pattern at the time of neural network's training has been used. In Figure (3), neurons of output layer and target vector at the time of training in three parts have been

Table 1: Different form of digit "6" in Farsi handwritten digit

Digit "6"			

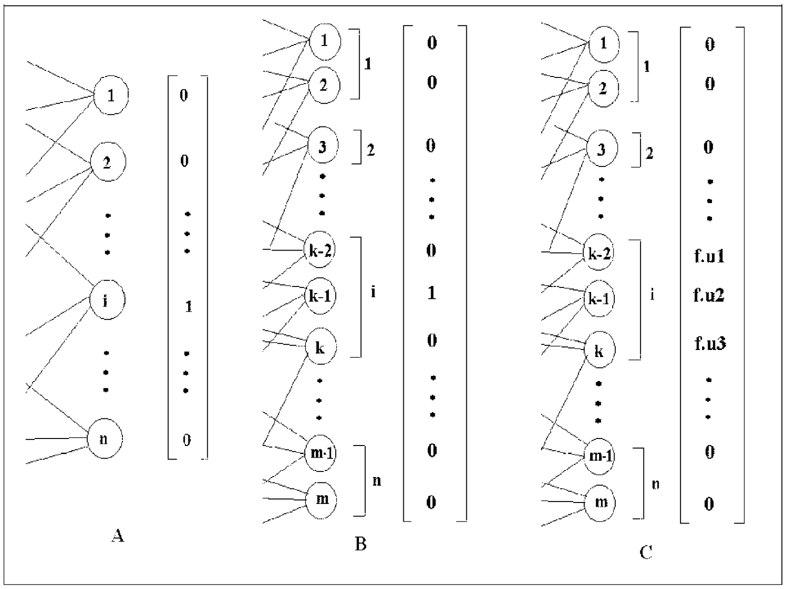


Fig. 3: The structure of last layer neurons of neural network and target vector, when one sample of i^{th} main class is applied to network in training step.
A) We have one neuron for each main class,
B) For each main class, we have some subclasses independent from each other,
C) For each main class, we have some subclasses related to each other.

illustrated. It should be paid attention that this target vector is related to the time that one train sample from i^{th} class has been applied to network. In part (A), for each class has been considering only one neuron, i.e. we have n main classes and n output neurons. In parts (B) and (C), for each class has been considered some sub-class. As we can see in Figure (3), for main class (1), we have two sub-classes. Class (2) isn't divided into subclasses. The i^{th} class is divided into 3 sub-classes and nth class is divided into 2 sub-classes. In pat (B), m sub-classes have been considered independent of each other and the sample that has been given to network is belong to second sub-class from i^{th} main class determined by k-l. In part (C), training samples of i^{th} main class have been divided into 3 fuzzy clusters by means of FCM and the sample that has been given to network, its dependence value to first, second and third subclasses of i^{th} main class are equal to u1, u2 and u3 respectively. Also, factor f is chosen as maximum value of elements of target vector is equal to unit and is obtained from equation $f = \frac{1}{\max(u1, u2, u3)}$.

Recognition of Farsi Handwritten Digits: In this section, recognition of Farsi handwritten digits using neural network and the way of use of proposed approach has been described. Also, flowchart of approach has been shown in Figure (4) and in following sections different parts of this flowchart have been explained in detail.

Test Data Group: In order to recognize Farsi handwritten digits, we gather the group of handwritten digits involving 1947 digits that get from different people. These digits have been written on white A4-sized paper and they have been imaged with precision of 300 dots per inch. Then the separated images of digits along with their labels from 0 to 9 and the number of image's column and row of each digit have been saved in a document. These samples are divided into two separated groups, testing and training groups. The training group involves 1000 digits and the testing group involves 947 digits. Some gathered samples are illustrated in Figure (5).

Feature Extraction: As we can see in Figure (5), similarity between some of the words can create problem. Therefore, finding a proper approach to extract feature of these digits is one of the most important section of each pattern recognition problem. So, in this paper, features of loci feature approach have been used.

Characterization Loci Feature: Characteristic loci features are usually defined in horizontal and vertical or 45° and 135° directions (Glucksman, 1967). In computing feature vectors, we assign a number to each background pixel as shown in Figure (6) the features are computed according to the number of intersections with the sub word body in right, upward, left and downward directions.

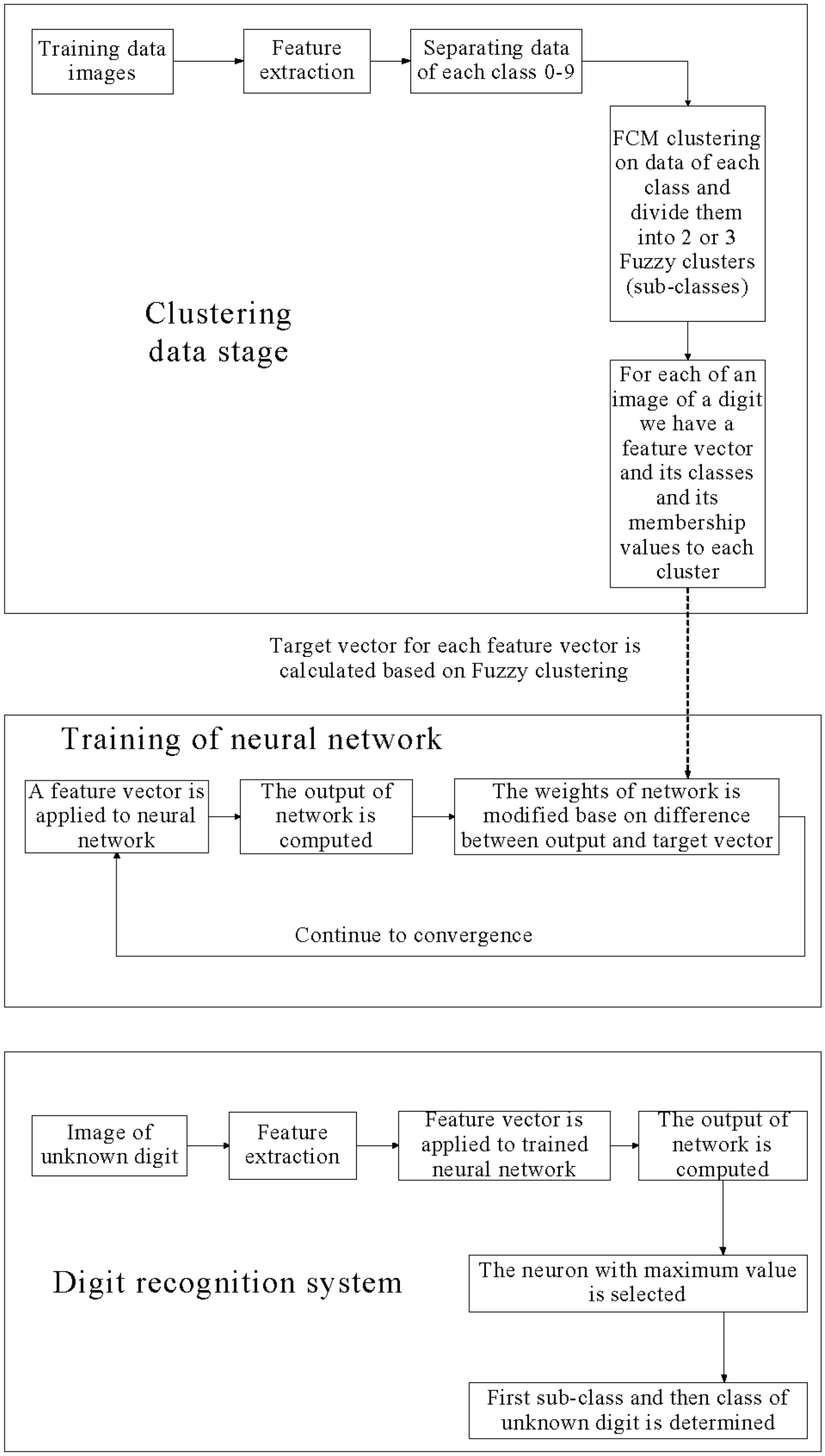


Fig. 4: Flowchart of proposed approach

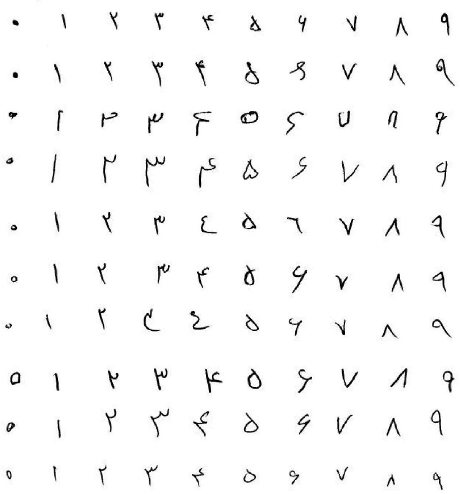


Fig. 5: Samples of used digits

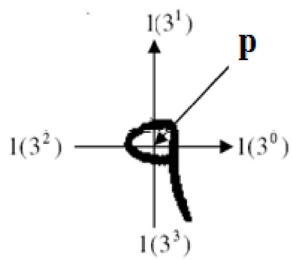


Fig. 6: Calculating characterization loci features

Then, for each background pixel, a four digit number of base 3 are obtained. For instance, the locus number of point P in Figure (6) is $(1111)_3= (40)_{10}$. The locus numbers are between 0 and 80. This is done for all background pixels. In this case, dimension of the feature vectors becomes 81. Each element of this vector represents the total number of background pixels that have locus number corresponding to that element. For example, 56th element of this vector represents the number of background pixels with locus number of 56. Features are normalized by dividing them by the total number of background pixels.

The proposed method is described as follow: First, we done feature extraction using Characteristic loci features then recognition of handwritten digits is done using mixture of experts.

$$\textit{Cutting Number} = \textit{Right Num of Cut} \times 3^0 + \textit{Top Num of Cut} \times 3^1 + \textit{Left Num of Cut} \times 3^2 + \textit{Down Num of Cut} \times 3^3 \rightarrow$$

$$(1111)_3 = 1\times 3^0 + 1\times 3^1 + 1\times 3^2 + 1\times 3^3 = (40)_{10}$$

Clustering of Test Samples: The number of test samples are 100 in each class that by using above mentioned features and by using Fuzzy C-Means, they have been divided into two clusters at first and another time into 3 clusters. In Figure (7) samples of numbers 1 and 4 that they have been divided into three clusters have been illustrated. It should be mentioned that samples don't belong to one cluster; in the case that each sample belongs to all three clusters in Fuzzy way. In Table (2) some samples of numbers 1 and 4 along with their membership value of each of three sub-clusters have been inserted.

Table 2: Images of samples of digits 1 and 4 along with their dependence value to each of three clusters

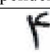
Image of digits						
Membership values to each one of clusters	0.9260	0.0398	0.0945	0.5409	0.3035	0.0980
	0.0328	0.8172	0.2088	0.4315	0.6697	0.1082
	0.0412	0.1429	0.6968	0.0276	0.0268	0.7939

Table 3: Results of recognition of numbers with the neural network without hidden layer

	Network A	Network B	Network C
Rate of recognition for training samples (%)	92.3	90.3	94.6
Rate of recognition for testing samples (%)	89.12	87.54	91.45

Table 4: Results of recognition of digits by neural network with one hidden layer and 3 subclasses for each digit

	Network A	Network B	Network C
Rate of recognition for training samples (%)	95.1	91.4	95.9
Rate of recognition for testing samples (%)	88.49	84.69	89.33

Table 5: Results of recognition of digits by neural network with one hidden layer and 2 subclasses for each digit

	Network A	Network B	Network C
Rate of recognition for training samples (%)	95.1	94.5	97.2
Rate of recognition for testing samples (%)	88.49	88.28	90.81

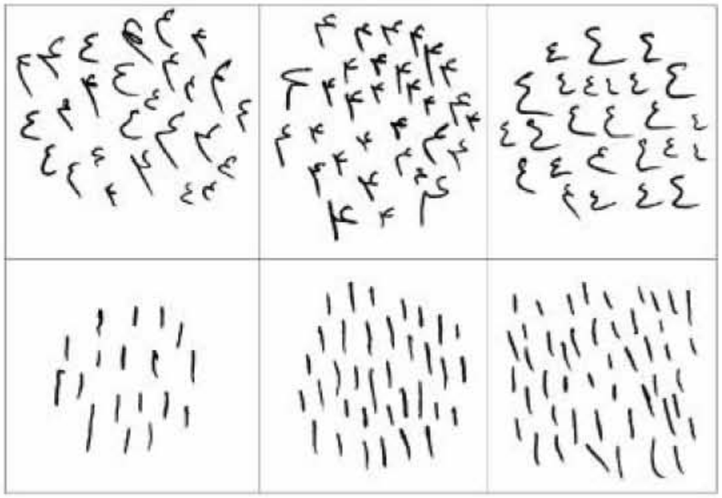


Fig. 7: Clusters of digits 1 and 4

Fig. 8: Samples of errors in network (A) that there aren't in network (C). Bottom of each digit's image have been presented two digits that the left number is input class and the right number is result of recognition of network (A).

Recognition of Digits Using Neural Network: The proposed approach has been tested on the network without hidden layer and another time on the network with hidden layer.

By using proposed method, successful rate of recognition for training data is 97.2 % and for testing data is 90.81%.

The Neural Network Without Hidden Layer: The used neural network has 82 neurons in input layer (81 neurons for features of loci feature and one bias neuron). The number of neurons in output layer in main network is 10 and in proposed network are 30. After finishing train of neural network, when rate of recognition becomes stable, recognition rate of test and train samples have been measured and their results have been inserted in Table (3). These results show performance of proposed approach (network C). Networks A, B and C are networks that their outputs and target vectors have been shown in Figures (3-A, 3-B and 3-C). Here, for each class has been considered 3 sub-classes. Network B has low performance in respect of other two networks.

Low performance of network (B) is due to this network, for samples of a main class that their features are similar to each other but they are different in subclass, cannot be trained well. For instance, suppose that P and Q are samples of a main class that they are similar to each other but they are in different subclasses, if we use network (A), for both of samples we will have same target vector. In network (B), when sample P has been applied to neural network, network modify own weights according to P and when Q has been applied to network, network will change its weights according to Q and this is due to forget sample P. but in network (C) training of sample P will lead to training of sample Q.

Samples of errors that are exist in network (A) and have been modified by network (C), have been illustrated in Figure (8). It is mentionable that in network (C), there are errors too that aren't in network (A), but we can say that total results of proposed approach are better.

In Figure (8), bottom of each digit's image have been presented two numbers that the left number is input class and the right number is result of recognition of network (A). As we said, in network (C) there aren't such errors. As we can see in Figure (8), most of errors of network (A) are related to assign numbers to class zero, the reason of this problem is that in this situation of output all neurons of last layer are became zero and because in this situation maximum algorithm chooses zero, samples are assigned to class zero, in such conditions, unknown pattern can be rejected.

The Network with One Hidden Layer and Three Subclasses for Each Digit: In this part, the network has been used is similar to previous network, but with one hidden layer involving 16 neurons. And results have been inserted in Table (4).

The Network with One Hidden Layer and Two Subclasses for Each Digit: In this part, in proposed network instead of three classes for each class, two classes have been used, i.e. networks (B) and (C) have 20 neurons in last layer. In Figure (6) obtained results have been presented.

CONCLUSION

Using the proposed approach in this paper, we can improve the rate of recognition of digits by neural network.

It is mentionable that the number of clusters for each class has an effect on rate of recognition. But in this paper, for all classes has been considered same cluster and it seems that if the number of cluster for each class is chosen based on dispersal rate of samples, performance of approach would be better.

At the end, it should be mentioned that features used for recognition are loci feature vertical [15]. And if other features such as feature of zoning [4] are used along with these features, we would have better results. Also by using preprocessing and after processing, we can reach better results too.

REFERENCES

1. Setareh Sheik, 2008. Morteza Bagherpour. Estimating the Saturation Thermodynamic Properties of Propene Using a Feed Forward Neural Network. World Appl. Sci. J., 4(2): 169-173.
2. Ibrahiem, M.M. El Emary and S. Ramakrishnan, 2008. On the Application of Various Probabilistic Neural Networks in Solving Different Pattern Classification Problems. World Appl. Sci. J., 4(6): 772-780.
3. Mansour Sheikhan and Amir Ali Sha'bani, 2009. Fast Neural Intrusion Detection System Based on Hidden Weight Optimization Algorithm and Feature Selection. World Appl. Sci. J. 7 (Special Issue of Computer and IT): 45-53.
4. Mansour Sheikhan and Sahar Garoucy, 2010. Reducing the Codebook Search Time in G.728 Speech Coder Using Fuzzy ARTMAP Neural Networks. World Appl. Sci. J., 8(10): 1260-1266.
5. Cheng-Lin Liu and Ching Y. Suen, 2009. A new benchmark on the recognition of handwritten Bangla and Farsi numeral characters. Pattern Recognition, 42: 3287-3295.
6. Gholam Ali Montazer, Hamed Qahri Saremi and Vahid Khatibi, 2010. A neuro-fuzzy inference engine for Farsi numeral characters recognition. Expert Systems with Applications, 37: 6327-6337.
7. Sameh, M. Awaidah and Sabri A. Mahmoud, 2009. A multiple feature/resolution scheme to Arabic (Indian) numerals recognition using hidden Markov models. Signal Processing, 89: 1176-1184.

8. Sadati, N. And B. Nazari, 1995. "Using Fuzzy Logic in Farsi handwritten digits Recognition", In proceeding of 3rd Electrical Engineering Conference of Iran, Tehran, pp: 241-247, (in Persian).
9. Yang, K., X. Xu and W. Zhang, 2000. Design Neural Networks Based Fuzzy Logic., Fuzzy Sets and Systems, pp: 325-328.
10. Zhang, M. and J. Fulcher, 1996. Face recognition using artificial neural network group-based adaptive tolerance (gat) trees. IEEE Transactions on Neural Networks, 7(3): 555-567.
11. Waibel, B. *et al.,* 1989. Phoneme recognition using time-delay neural networks. IEEE Transactions on acoustics, Speech and Signal Processing, 37(3): 328.339.
12. Le Cun, Y. *et al.,* 1989. Back-propagation Applied to Handwritten Zip Code Recognition. Neural Computation, 1: 541-551.
13. Widrow, B. and M. Lehr, 1990. 30 years of adaptive neural networks: Perceptron, Madaline and Backpropagation. Proceedings of the IEEE, 78: 1415-1442.
14. Xu, L., A. Krzy and C.Y. Suen, 1992. Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition. IEEE Transactions on Systems, Man and Cybernetics, 22(3): 418-435, May 1992.
15. Reza Ebrahimpour, Mohammad R. Moradian, Alireza Esmkhani and Farzad M. Jafarlou, 2009. Recognition of Persian handwritten digits using Characterization Loci and Mixture of Experts. Intl. J. Digital Content Technol. Its Applications, 3(3): 42-46, September 2009.
16. Chiang, J.H. and P.D. Gader, 1997. Hybrid Fuzzy-Neural Systems in Handwritten Word Recognition, IEEE Trans. On Fuzzy Systems, 5(4).
17. Jamal, M. Nazzal, Ibrahim M. El-Emary and Salam A. Najim, 2008. Multilayer Perceptron Neural Network (MLPs) For Analyzing the Properties of Jordan Oil Shale. World Appl. Sci. J., 5(5): 546-552.
18. Dulyakan, P. And Y. Rangasanseri, 2001. FUZZY C-MEANS CLUSTERING USING SPATIAL INFORMATION WITH APPLICATION TO REMOTE SENSING. 22nd Asian Conference on Remote Sensing, 5-9 November 2001, Singapore.
19. Kabir, A., K. Bahari and M. Ahmadzade, 1993. "Recognition of Typed Persian Text", In proceeding of 1st Electrical Engineering Conference of Iran, Tehran, pp: 285-294, (in Persian).
20. Joohari, V. Majd and S.M. Razavi, 2000. "Fuzzy recognition of Farsi handwritten digits", In proceeding of 1st Machine vision and Image processing of Iran, pp: 151-144, Birjand university, Birjand. (in Persian).