Pruned Fuzzy Hypersphere Neural Network (PFHSNN) for Lung Cancer Classification

D. N. Sonar
Dept. of Comp. Sci.,
College of Comp. Sci. & Info. Tech.,
Latur-413531, (M.S.) India

U. V. Kulkarni
Professor, Dept. of Comp. Sci. & Engg.,
SGGS Institute of Engg. & Tech.,
Vishnupuri, Nanded-431606, (M.S.), India

ABSTRACT
In this paper Pruned Fuzzy Hypersphere Neural Network (PFHSNN) is proposed which is an extension of Fuzzy Hypersphere Neural Network (FHSNN). A pruning procedure is incorporated into FHSNN after its leaning phase to reduce the network size. The experimental results for JSRT database show that PFHSNN is considerably superior in terms of training and recall time. It yields 91.66% recognition rate.

Keywords
Artificial Neural Network, Lung Nodule, Fuzzy Hypersphere Neural Network (FHSNN), Pattern Classification

1. INTRODUCTION
The fuzzy neural networks are among the popular and widely used approaches for undertaking pattern classification and clustering. Quteishat A., Lim C.P., and Tan K.S. [1] have proposed a modified fuzzy min-max neural network with genetic algorithm based rule extractor. Carpenter G. and Tan A. [2] proposed two stage technique for rule extraction. In first stage, pruning the network structure by removing excessive recognition categories is done and in second, quantizing the continuous network weights and translating the final system states into a set of descriptive rules is carried out. Simpson P.K. [3] proposed supervised fuzzy min-max neural network (FMN) that utilizes fuzzy sets as pattern classes. Kwan and Cai [4] have proposed four-layer feed forward unsupervised fuzzy neural network. Belacel N. et al. [5] have proposed a decision support system for leukemia diagnosis based on supervised learning algorithms. Kulkarni et al. have proposed two separate algorithms for classification [6] and clustering [7]. Gabrys B. and Bargiela A. [8] have presented general fuzzy min-max neural network (GFMM) for clustering and classification which is a mixture of supervised and unsupervised learning. Lin D. and Yan C. [9] stated a neural fuzzy model to formulate the diagnosis rules for identifying the pulmonary nodules. Kulkarni, Doye and Sontakke [10] have proposed a general fuzzy hypersphere neural network classifier which is an extension of fuzzy hypersphere neural network. It utilizes fuzzy sets as pattern classes and each fuzzy set is a union of fuzzy set hyperspheres. In FHSNN algorithm proposed by Kulkarni and Sontakke [11], a hypersphere is defined by its center point and radius, and characterized by hypersphere membership function. All patterns contained within the hypersphere have full class membership.

In this paper, a pruning procedure is incorporated into FHSNN after its leaning phase. The objective of pruning is to reduce the network complexity on the basis of confidence factor of each hypersphere and to get a high classification performance. The proposed algorithm achieves a considerable rate of overall identification when compared with FHSNN algorithm.

This paper is organized as follows. In Section 2, the proposed two-stage PFHSNN is explained in brief. In Section 3, the comparative results of two algorithms are presented. Finally in Section 4, Conclusions are drawn and several issues for further studies are indicated.

2. TWO-STAGE PFHSNN
2.1 Stage I: The Pruned FHSNN Pattern Classifier
In this stage, modification in the original FHSNN is done by introducing a pruning process after its learning phase. The PFHSNN uses two subsets (training and testing), the input-output data set. Then, the hyperspheres created in the learning phase undergo a pruning process based on user defined threshold which range in between [0-1]. The training data set is used for creating hyperspheres in FHSNN, while the testing data set is used for testing purpose. The details are given below.

2.1.1 Learning of FHSNN
The FHSNN [11] is built by using hypersphere fuzzy sets. A hypersphere is completely defined by its center point and radius, and characterized by hypersphere membership function. All patterns contained within the hypersphere have full class membership.

The structure of FHSNN consists of three layers of nodes as shown in Fig. 1(a). The first layer (F_R) is an input layer which accepts patterns and consists of n processing elements, one for each dimension of the pattern. Second layer (F_M) nodes are constructed during training and each node represents hypersphere (HS) characterized by its membership function. The processing of HS node is shown in Fig. 1(b). The weights between first (F_R) and second (F_M) layer represent center points (CPs) of the HSs. As shown in Fig. 1(b), $C_j = (C_{j1}, C_{j2}, ... C_{jn})$ represents CP of the HS m_j. The threshold input of HS denoted as T is set to 1 and it is weighted by ζ_j. ζ_j represents radius of the HS m_j, which is updated during training. Matrix C and vector ζ stores the CPs and radiuses of the HSs, respectively. The maximum size of HS is bounded by a user defined value λ, called as growth parameter where $0 \le \lambda \le 1$. λ puts maximum limit on the radius of the HS.
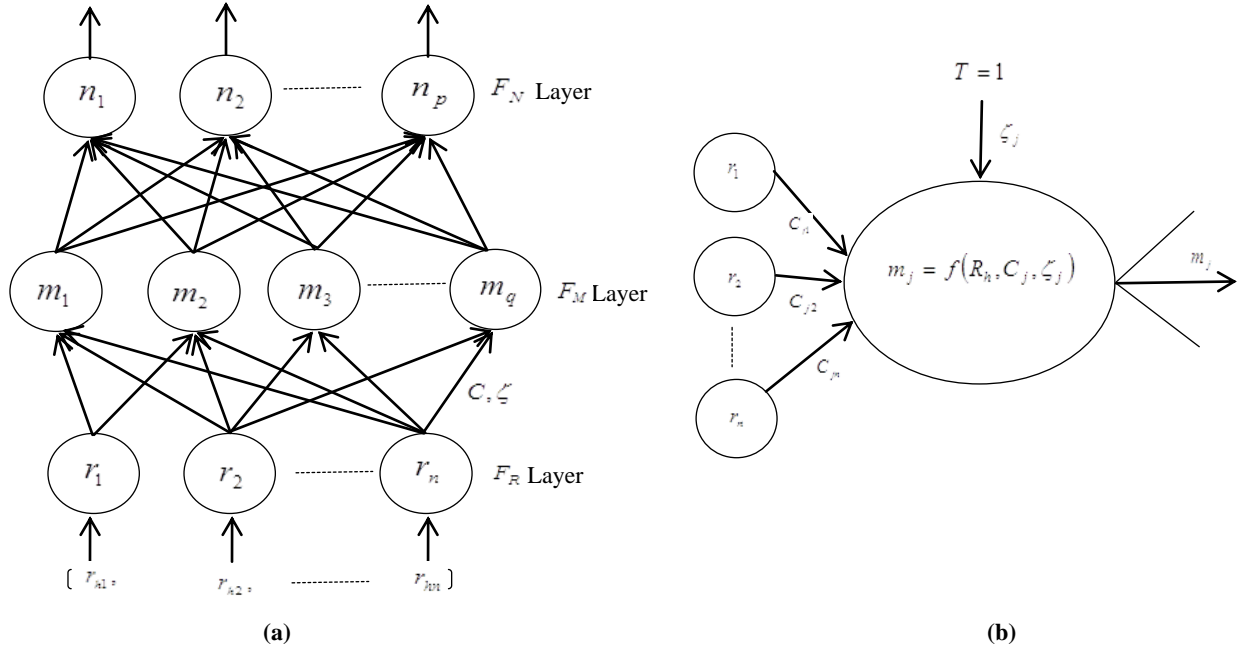
(a)   **(b)**

Fig 1:(a) Fuzzy hypersphere neural network, (b) implementation of fuzzy hypersphere

Consider the training set $R = \{R_h | h = 1,2,...,P\}$, where $R_h = (r_{h1}, r_{h2}, ... r_{hn})$ is the hth pattern, the membership function of the HS node m_j is defined as,

$$m_j(R_h, C_j, \zeta_j) = 1 - f(l, \zeta_j, \gamma), \qquad (1)$$

where $f(l, \zeta_j, \gamma) = \begin{cases} 0, & \text{if } 0 \le l \le \zeta_j \\ l\gamma, & \text{if } \zeta_j < l \le 1, \\ 1, & \text{if } l > 1 \end{cases}$

and the argument l is defined as,

$$l = \left(\sum_{i=1}^{n} (C_{ji} - r_{hi})^2 \right)^{1/2} \qquad (2)$$

The membership function returns $m_j = 1$, if HS includes the pattern R_h. The sensitivity parameter $\gamma, 0 \le \gamma \le 1$, governs how fast the membership value decreases outside the HS, as the distance between R_h and C_j increases.

Each node of F_M layer represents a class. The weights assigned to the connections between F_M and F_N layers are binary values that are stored in the matrix U. The equation for values of U is

$$u_{jk} = \begin{cases} 1 & \text{If } m_j \text{ is a hypersphere of class } n_k \\ 0 & \text{Otherwise} \end{cases}$$

where m_j is the jth node of second layer and n_k is the kth F_N node for $k = 1,2,...,p$ and $j = 1,2,...,q$. Each F_N node performs the union of fuzzy values returned by HSs is described as,

$$n_k = m_j * u_{jk} \qquad (3)$$

for $k = 1,2,...,p$ and $j = 1,2,...,q$.

Creating hyperspheres in the pattern space is a three-step process consisting of creation/expanding of hyperspheres, overlap test and contraction.
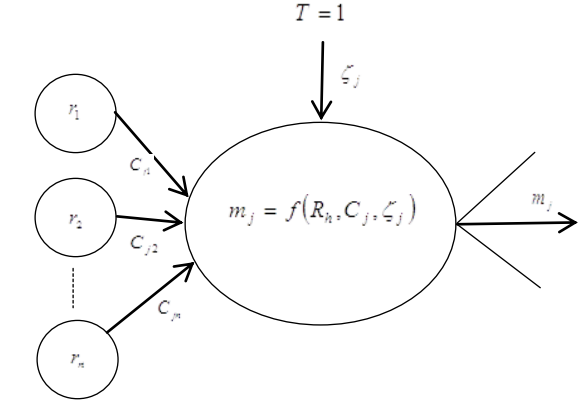
2.1.1.1. Creation of HSs:

Given the hth training pair (R_h, d_h), find all the HSs belonging to the class d_h. Following are the steps for possible inclusion of input pattern R_h.

Step 1: Determine whether the pattern R_h is contained by any one of the HSs by using (1). If R_h is included by any one of the HS then the remaining steps in the training process are skipped and the training continues with the next pair.

Step 2: If the pattern R_h falls outside the HS, then the HS is expanded to include the pattern if the expansion criterion is met. For the HS m_j to include R_h the constraint in (4) should be met.

$$\left(\sum_{i=1}^{n} (C_{ji} - r_{hi})^2 \right)^{1/2} \le \lambda. \qquad (4)$$

If the expansion criterion is satisfied then the pattern R_h is included as,

$$\zeta_j^{new} = \left(\sum_{i=1}^{n} (C_{ji} - r_{hi})^2 \right)^{1/2}. \qquad (5)$$

Step 3: If the pattern R_h is not included by any of the above steps then the new HS is created for that class described as,

$$C_{new} = R_h \text{ and } \zeta_{new} = 0. \qquad (6)$$

2.1.1.2. Overlap Tests

The learning algorithm allows overlap of HSs from the same class and eliminates the overlap between HSs from separate classes. Overlap test is performed as soon as the HS is expanded in step 2 or created by step 3.

(a) Overlap test for step 2: Let $C_u = [x_1, x_2, ... x_n]$ and ζ_u represents CP and the radius of the expanded HS, respectively.

Similarly $C_v = [x_1, x_2, ... x_n]$ and ζ_v represents respectively CP and the radius of the HS of other class. If

$$\left(\sum_{i=1}^{n}(c_{ui}-c_{vi})^2\right)^{1/2} \leq \zeta_u + \zeta_v \qquad (7)$$

means two HSs from separate classes are overlapping.

(b) Overlap test for step 3: If the created HS falls inside the HS of other class means there is an overlap. This can be verified using the HS membership function defined in (1).

2.1.1.3. Removing Overlap

If step 2 has created overlap of HSs from separate classes then overlap is removed by restoring the radius of just expanded HS. Let m_j be the expanded HS then it is contracted as,

$$\zeta_j^{new} = \zeta_j^{old}, \qquad (8)$$

and new HS is created for the pattern using (9).

If the step 3 creates overlap then it is removed by modifying HS of other class. Let C_p and ζ_p, represents CP and the radius of created HS and C_q and ζ_q are CP and radius of other class. Then overlap is removed as,

$$\zeta_j^{new} = \left(\sum_{i=1}^{n}\left(c_{pi}-c_{qi}\right)^2\right)^{1/2} - \delta \qquad (9)$$

where δ is a small number selected just enough to remove overlap. In our experiments $\delta = 0.0001$.

2.2 Stage II: Pruning the FHSN Network

After the learning of FHSNN, a pruning procedure is incorporated to reduce the number of hyperspheres (HS). The pruning approach removes the hyperspheres that are not needed. The main objective behind this is to reduce the network complexity with a considerable classification performance. Pruning method is based on confidence factor of each hypersphere. The HSs with low confidence factor are pruned using a user defined threshold value to reduce number of hyperspheres and hence the network complexity.

Confidence factor CF_j for each hypersphere m_j is calculated in terms of its usage frequency and accuracy. It is calculated as,

$$CF_j = (1-\gamma)U_j + \gamma A_j \qquad (10)$$

Where U_j and A_j are the usage and accuracy indices of hypersphere m_j and $\gamma \in [0, 1]$ is a weighing factor. The usage index of a hypersphere is defined to be the number of training set patterns classified by any hypersphere m_j divided by the maximum number of training set patterns classified by any hypersphere with the same classification class. Usage is thus scaled to lie between 0 and 1, with U_j equal to 1 for at least one hypersphere. On the other hand, the accuracy index of a hypersphere is defined to be the number of correctly classified training set patterns by any hypersphere m_j divided by the maximum correctly classified patterns classified with the same classification class. Accuracy is also scaled to lie between 0 and 1, with A_j equal to 1 for at least one hypersphere.

The confidence factor identifies hyperspheres that are frequently used and generally accurate in prediction as well as hyperspheres that are rarely used but yet highly accurate. Hyperspheres with a confidence factor smaller than a user-defined threshold (setting is 0.5 in this experiment) are pruned. After removing hyperspheres with low confidence factor, the remaining hyperspheres are used for testing purpose.

3. EXPERIMENTAL RESULTS

To evaluate the proposed algorithm, images are obtained from the public Standard Digital Image Database, JSRT [12]. All the data samples are divided into five subsets according to the subtlety of nodule. These images contain nodules within the range 5-60 mm. The pre-processing including normalization of size, filtering to enhance the contrast is done on images.

Table 1 shows the comparative performance of FHSNN and PFHSNN in terms of % recognition rate, number of hyperspheres and the λ value. λ represents the maximum radius limit of the hypersphere of fully trained neural network. To get better result, the values of λ are adjusted in the range of 0 and 1. The user defined pruning threshold was set to 0.6. This results in removing the hyperspheres with Confidence Factor less than 0.6. After removal of unneeded hyperspheres, the remaining hyperspheres are used for further processing, i.e., testing.

Thus the % recognition rate in Table 1 shows that the proposed PFHSNN gives superior classification performance than FHSNN with less number of hyperspheres.

Table 1. Comparative performance of FHSNN and PFHSNN for pattern classification of Lung Nodules

Neural Network	λ Value	No. of HS	Recognition rate (%)
FHSNN	0.041	60	83.33
PFHSNN	0.41	04	91.66

4. CONCLUSION

In this paper, a PFHSNN has been proposed. A pruning procedure is incorporated after learning phase of the neural network based on confidence factor of each hypersphere. The hyperspheres with a low confidence factor are pruned depending upon a user defined threshold value so that the network complexity gets reduced.

The performance of PFHSNN is evaluated using images of the public Standard Digital Image Database, JSRT. All the data samples are divided into five subsets according to the subtlety of nodule. The experimental result shows that the proposed pruning approach (PFHSNN) gives superior performance in terms of recognition rate and recall time with less number of hyperspheres.

There are some aspects that can be investigated in further studies. The proposed algorithm can be modified to work in unsupervised environment for classification. It will be interesting to use this algorithm for clustering also.

5. REFERENCES

[1] Quteishat A., Lim C.P. and Tan K.S. 2010. A Modified Fuzzy Min-Max Neural Network with a Genetic-Algorithm-Based Rule Extractor for Pattern Classification IEEE Trans. on Systems.

[2] Carpenter G. and Tan A. 1995. Rule extraction: From neural architecture to symbolic representation.

[3] Simpson P.K. 1992. Fuzzy min-max neural networks-Part 1: Classification. IEEE Trans. neural networks.

[4] Kwan H.K. and Cai Y. 1994. A fuzzy neural network and its applications to pattern recognition. IEEE Trans. on fuzzy systems.

[5] Belacel N., Vincke P., Scheiff J.M. and. Boulassel R. Acute leukemia diagnosis aid using multicriteria fuzzy assignment methodology.

[6] Kulkarni U.V., Sontakke T.R. and Randale G.D. 2001. Fuzzy hyperline segment neural network for rotation invariant handwritten character recognition.

[7] Kulkarni U.V., Sontakke T.R. and Kulkarni A.B. 2001. Fuzzy hyperline segment clustering neural network for rotation.

[8] Gabrys B. and Bargiela A. 2000. General Fuzzy Min-Max Neural Network for Clustering and Classification. IEEE Trans. neural networks.

[9] Lin D. and Yan C. 2002. Lung Nodules Identification Rules Extraction with Neural Fuzzy Network. IEEE Processing on Neural Information.

[10] Kulkarni U.V., Doye D.D. and Sontakke T.R. 2002. General Fuzzy Hypersphere Neural Network. IEEE, IJCNN.

[11] Kulkarni U.V. and Sontakke T.R. 2001. Fuzzy Hypersphere Neural Network Classifier. published in proceedings of 10th IEEE conference on Fuzzy Systems held at University of Melbourne, Australia,

[12] Japanese Society of Radiological Technology (JSRT), http://www.jsrt.or.jp.