Knowledge-Based Systems 246 (2022) 108620

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A compact fuzzy min max network with novel trimming strategy for pattern classification

Santhos Kumar A. [a,*], A. Kumar [a], V. Bajaj [a], G.K. Singh [b]

[a] PDPM Indian Institute of Information Technology Design and Manufacturing, Jabalpur 482005, India
[b] Department of Electrical Engineering, Indian Institute of Technology Roorkee, Uttarakhand 247667, India

ARTICLE INFO

Article history:
Received 4 September 2020
Received in revised form 17 March 2022
Accepted 18 March 2022
Available online 24 March 2022

Keywords:
Hyperbox classifier
Fuzzy min–max (FMM) network
Proposed FMM (PFMM) model
Pattern classification
Histopathology image
Breast cancer

ABSTRACT

Hyperbox classifier has large contribution to the field of pattern classification, because of its efficiency and transparency. Hyperbox classifier is efficiently implemented by using fuzzy min–max (FMM) neural network. FMM was modified many times to improve the classification accuracy. Moreover, there still exists a space for increasing the accuracy of hyperbox based classifiers. In this paper, four modifications are proposed to FMM network for increasing the classification accuracy rate. First, centroid and K-highest (CCK) based criteria to select the expandable hyperbox. Second, a new set of overlap test cases to consider all types of overlapping regions. Third, a new set of contraction rules to settle the overlapped regions. Fourth, novel hyperbox trimming strategy to reduce the system complexity. The proposed method is compared with FMM, enhanced FMM (EFMM) and Kn_FMM using five datasets. Experimental results clearly reflect the improved efficiency of proposed method. Proposed FMM (PFMM) network is also used to classify the histopathological images for knowing the best magnifying factor.

the tendency of an artificial neural network to completely and abruptly forget previously learned information upon learning new information.

1. Introduction

Artificial Neural Network (ANN) is a technique, which is motivated by the biological networks of the brain [1]. ANNs have been successfully applied in many fields like health care [2], data classification [3,4], diagnosis [5], arrhythmias classification [6], human recognition [7–9], and fault detection [10] etc. Learning strategies of most ANN based methods are related to offline or batch learning [11]. In ANN models, catastrophic forgetting is the main problem related to batch learning, such as multi-layer perceptron (MLP) and radial basis function (RBF) [12,13]. A number of ANN models are proposed to overcome the problem of catastrophic forgetting, such as fuzzy min–max (FMM) [14,15], and adaptive resonance theory (ART) networks [16,17]. Among different ANN models, many were investigated on FMM, and its variants due to its many significant features. Simpson introduced two hybrid ANN models to overcome the catastrophic forgetting problem, these are supervised classification FMM network [14] and unsupervised clustering FMM [15] network. Over the years, FMM and its variants are being used in many applications such as color recognition [18], fault classification [19], ECG beat analysis [20], medical diagnosis [21] and medical image classification [22] etc.

Zadeh [23] has introduced the fuzzy set theory, which is widely applied to the problem of pattern recognition and classification. Since the intelligent systems were developed, many researchers have exploited the inference properties of fuzzy systems and ANN learning ability [24]. Several hybrid models have been designed by combining ANNs and fuzzy sets, and few can be found in [25–28].

In recent years, many FMM variants have been introduced which are mostly based on two conventional FMM networks, such as FMM classification model [14] and FMM clustering model [15]. In [29], GFMM neural network was introduced which is an extension of FMM network, in terms of expansion and contraction rules. In [30], a general reflex FMM (GRFMM) network was introduced which combines the classification and clustering methods of FMM and human reflex mechanism. In [31], a stochastic FMM neural network was introduced that is an extension of reinforcement learning [27]. In [32], an adaptive resolution min max neural network was presented that overcomes the undesired properties of convolutional FMM network. Moreover, during expansion process, no constraint is applied to the dimension of hyperbox. It has relatively less expansion network structure than the FMM network.

An inclusion/exclusion fuzzy system was introduced in [33] which creates inclusion and exclusion hyperbox. Exclusion hyperbox contains the overlapping patterns, whereas inclusion hyperbox contains input patterns of the same class. In [34], a weighted FMM (WFMM) neural network was introduced, in which the

* Corresponding author.
E-mail addresses: hymavathisanthu@gmail.com (Santhos Kumar A.), anilkdee@gmail.com (A. Kumar), bajajvarun056@yahoo.co.in (V. Bajaj), gksnfee@gmail.com (G.K. Singh).

https://doi.org/10.1016/j.knosys.2022.108620
0950-7051/

A radial basis function (RBF) is a real-valued function whose value depends only on the distance between the input and some fixed point, either the origin, so that , or some other fixed point , called a center, so that . Any function that satisfies the property is a radial function.

A common strategy is to grow the tree until each node contains a small number of instances then use pruning to remove nodes that do not provide additional information. Pruning should reduce the size of a learning tree without reducing predictive accuracy as measured by a cross-validation set.

hyperbox size is not maintained by the expansion and overlap test rules. In [35], a supervised FMM classifier network with compensatory neurons model was proposed. In [36], data-core-based fuzzy min max network (DCFMN) model was presented for data classification problem, which uses a novel fuzzy membership function for overlapping and classifying neurons.

A modified FMM network was introduced in [37] which increases the classification accuracy of FMM model. In [38], a k-nearest expansion criteria was introduced which avoids creation of too many smaller sized hyperboxes, and reduces the network complexity. In [39], an enhanced FMM (EFMM) was presented that decreases the overlapping of hyperboxes during expansion. Hyperbox pruning strategy was introduced in [40] which is an extension of EFMM network. In [41], an improved FMM network was presented that uses semi-perimeter of the hyperbox along with k-nearest mechanism to select the expandable hyperbox, and also modifies all the contraction rules of conventional FMM and its variants. A fuzzy clustering method was presented in [42] which creates hyperbox entropy (HE) to evaluate the hyperbox performance. In recent years, FMM and EFMM models were applied to medical image of different magnifications [43,44] to find best magnification factor for classification.

Motivation and Contributions

FMM, EFMM and Kn_FMM have several limitations. Expansion procedure used by FMM and Kn_FMM increases the overlapped areas between different class hyperboxes. Moreover, the expansion process used by EFMM generates more small sized hyperboxes. The overlap test rules considered by FMM and Kn_FMM, avoid some of the overlapping cases. On the contrary, EFMM has a large number of overlapping and contraction cases that can be further minimized without affecting the system functionality. During the expansion process, FMM, EFMM and Kn_FMM arbitrarily select the expandable hyperbox to fix the equal membership value problem. Moreover, contraction mechanisms used by FMM, EFMM and Kn_FMM are often biased towards a particular hyperbox, and evenly partition the overlapped area irrespective of the hyperbox size, which is inappropriate. An inappropriate training set may force a FMM based methods to generate some inefficient hyperboxes, which are unable to accurately represent the data of different class labels. Simultaneously, an ambiguous training set may sometimes badly influence the creation of hyperboxes, where data of one class are covered by the hyperboxes of different class. The above mentioned restrictions have inspired the authors to propose a compact FMM with a novel trimming strategy by fixing the various short comings of FMM based methods.

This paper presents new techniques to reduce the network complexity and misclassification rate. The main contributions of the paper are as follows:

1. A new rule is proposed, it combines the centroid and k-highest rule to select a candidate hyperbox, when two or more hyperboxes contain equal highest fuzzy membership function (MF) value,
2. A new set of compact overlap test cases are introduced to identify different types of overlapped regions between two hyperboxes of different class,
3. A new set of prioritized contraction cases are introduced to fix the problem of partially biased disorder and fully biased disorder, which unevenly assigns priority to different hyperboxes.
4. A novel hyperbox trimming strategy is introduced to revise the number and size of hyperboxes created during training phase. This reduces the unfavorable effect of an inappropriate training set.
5. The proposed fuzzy min–max neural network is also applied to BreaKHis database for each magnifying factor to classify the malignant and benign breast tumors, and then the accuracy rate and list of generated hyperboxes are calculated. The obtained results for each magnifying factor are compared for knowing the best magnification factor for classification.
6. Experimental results on benchmark databases (i.e. hepatitis, Lung cancer, WBC, Mammographic Mass and Multiple Features (mfeat-fac) Dataset) demonstrate that the proposed fuzzy min–max network achieves higher accuracy rate and less number of hyperboxes than the state-of-the-art methods.

It is important to mention that the proposed technique is an improvement of FMM, EFMM and Kn_FMM. In order to include a new training sample, above mentioned approaches arbitrarily select a probable hyperbox for expansion, when multiple hyperboxes have equal maximum fuzzy membership function value. However, such arbitrary selection is inappropriate. Therefore, the proposed method uses centroid and k-highest mechanism motivated rule to break the tie. The overlap test rules considered by FMM and Kn_FMM, avoid some of the overlapping cases. On the contrary, EFMM has a larger number of overlapping cases. Hence, in case of proposed method, a new set of overlap test rules are introduced that tackle the missed overlap test rules of FMM, Kn_FMM and unnecessary overlap test rules of EFMM. The contraction rules considered by the existing techniques inappropriately resolve the overlapped regions. In a view to accurately perform contraction, the proposed method incorporates side length of hyperbox to resolve the partially biased disorder and fully biased disorder. Generally, FMM based methods generate some inefficient hyperboxes during training phase. Furthermore, a hyperbox created during training phase may become inefficient due to a redundant training set. In order to solve the above problems, a novel class reversal mechanism is proposed in this paper which increases the accuracy rate and reduces the network complexity. In this paper, five benchmark datasets such as hepatitis, Lung cancer, Wisconsin Breast cancer (WBC), Mammographic Mass Dataset and Multiple Features (mfeat-fac) Dataset are used to compare the performance of proposed method with the state-of-the-art techniques. It is concluded, from experimental outcomes, that the proposed technique provides higher accuracy rate and generates very less hyperboxes than the existing state-of-the-art techniques.

The rest of this article is organized as follows, Section 2 presents an overview of FMM, Kn_FMM and EFMM neural networks, Section 3 describes the proposed methodology, Section 4 shows the experimental results with detailed discussion, and finally, Section 5 concludes this article.

2. Back ground study

This segment describes the traditional FMM [14], EFMM [39] and Kn_FMM [38] neural networks. There exist a number of modifications to FMM, however in this work, EFMM, Kn_FMM are considered due to its supremacy.

2.1. Fuzzy Min Max (FMM) network

Simpson have proposed the concept of hyperbox for data classification problems [14]. Each hyperbox is defined by a pairs of min–max points in an n-dimensional unit cube (I^n) along with an equivalent fuzzy membership function, where n indicates the list of features. FMM contains expansion, overlap, and contraction rules. The training process of FMM begins with a set of input

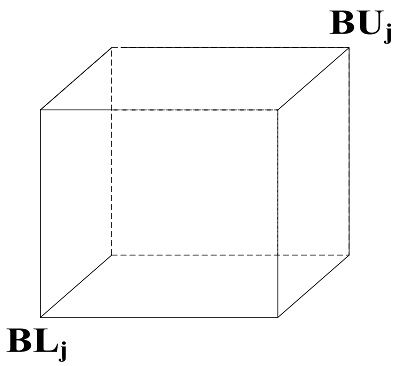

Fig. 1. 3-D hyperbox.

patterns BA_h and a set of target classes C_i where, h = 1… M and i = 1… k. Here, M indicates the total number of data samples and k represents the total number of classes. A graphical portrayal of 3-dimensional hyperbox is shown in Fig. 1, where its uppermost end is given by the maximum point (BU_j) and lowermost end is the minimum point (BL_j).

Each hyperbox can be indicated by using fuzzy logic HB_j as follows:

$$HB_j = \left(\left\{BA_h, BL_j, BU_j, f\left(\left(BA_h, BL_j, BU_j\right)\right)\right\}\right), \ \forall\ BA_h \in I^n \tag{1}$$

where, $BA_h = (ba_{h1}, ba_{h2}, \ldots, ba_{hn})$ indicates the hth input pattern and $BL_j = \left(bl_{j1}, bl_{j2}, \ldots, bl_{jn}\right)$, $BU_j = \left(bu_{j1}, bu_{j2}, \ldots, bu_{jn}\right)$ represent the jth hyperbox (HB_j), minimum and maximum points respectively. Hyperbox size is directly proportional to the expansion coefficient (θ). Number of generated hyperboxes is inversely proportional to the expansion coefficient (θ). The fuzzy MF value of an input sample corresponding to every hyperbox can be calculated using Eq. no. (2)

$$HB_j\left(BA_h\right) = \frac{1}{2n}\sum_{i=1}^{n}\left[\begin{array}{l} \max\left(0, 1-\max\left(0, \gamma \min\left(1, \mathrm{Ba}_{hi} - BU_{ji}\right)\right)\right) \\ +\max\left(0, 1-\max\left(0, \gamma \min\left(1, \mathrm{BL}_{ji} - Ba_{hi}\right)\right)\right) \end{array}\right] \tag{2}$$

where, HB_j represents MF of the jth hyperbox, BA_h represents the hth input pattern, BL_{ji} is the min point of HB_j and BU_{ji} is the max points of HB_j, and γ designates the sensitivity parameter, which regulates the downward speed of the fuzzy membership function.

FMM mainly contains 3 step process, which are expansion criteria, overlap cases, and contraction rules.

Expansion: Simpson developed Eq. (3) to test the expandability of a hyperbox to include a new input pattern BA_h.

$$n\theta \geq \sum_{i=1}^{n}\left(\max\left(BU_{ji}, Ba_{hi}\right) - \min\left(BL_{ji}, Ba_{hi}\right)\right) \tag{3}$$

where, Ba_{hi} denotes the input pattern, θ indicates the expansion coefficient, where $0 \leq \theta \leq 1. BU_{ji}$, BL_{ji} represents the maximum and minimum point of hyperbox respectively. The hyperbox HB_j can be expanded to include the input pattern BA_h if Eq. (1) is satisfied by HB_j for BA_h.

Overlap rules: The following four overlap cases are used to find the existence of overlap between two different class hyperboxes:
Case 1:

$$BL_{ji} < BL_{ki} < BU_{ji} < BU_{ki},\ \delta n = \min(BU_{ji} - BL_{ki}, \delta o), \tag{4}$$

Case 2:

$$BL_{ki} < BL_{ji} < BU_{ki} < BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o), \tag{5}$$

Case 3:

$$\begin{aligned} &BL_{ji} < BL_{ki} < BU_{ki} < BU_{ji}, \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{6}$$

Case 4:

$$\begin{aligned} &BL_{ki} < BL_{ji} < BU_{ji} < BU_{ki} \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o). \end{aligned} \tag{7}$$

Contraction rule: The contraction rules to resolve the overlap cases as shown in Eq. (4) to (7) are represented by Eq. (8) to (13) respectively.
Case 1:

$$BL_{j\Delta} < BL_{k\Delta} < BU_{j\Delta} < BU_{k\Delta}, BU_{j\Delta}^{new} = BL_{k\Delta}^{new} = \frac{BU_{j\Delta}^{old} + BL_{k\Delta}^{old}}{2} \tag{8}$$

Case 2:

$$BL_{k\Delta} < BL_{j\Delta} < BU_{k\Delta} < BU_{j\Delta}, BU_{k\Delta}^{new} = BL_{j\Delta}^{new} = \frac{BU_{k\Delta}^{old} + BL_{j\Delta}^{old}}{2}, \tag{9}$$

Case 3(a):

$$\begin{aligned} &BL_{j\Delta} < BL_{k\Delta} < BU_{k\Delta} < BU_{j\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) < (BU_{j\Delta} - BL_{k\Delta}),\ BL_{j\Delta}^{new} = BU_{k\Delta}^{old}, \end{aligned} \tag{10}$$

Case 3(b):

$$\begin{aligned} &BL_{j\Delta} < BL_{k\Delta} < BU_{k\Delta} < BU_{j\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) > (BU_{j\Delta} - BL_{k\Delta}),\ BU_{j\Delta}^{new} = BL_{k\Delta}^{old}, \end{aligned} \tag{11}$$

Case 4(a):

$$\begin{aligned} &BL_{k\Delta} < BL_{j\Delta} < BU_{j\Delta} < BU_{k\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) < (BU_{j\Delta} - BL_{k\Delta}),\ BU_{k\Delta}^{new} = BL_{j\Delta}^{old}, \end{aligned} \tag{12}$$

Case 4(b):

$$\begin{aligned} &BL_{k\Delta} < BL_{j\Delta} < BU_{j\Delta} < BU_{k\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) > (BU_{j\Delta} - BL_{k\Delta}),\ BL_{k\Delta}^{new} = BU_{j\Delta}^{old}. \end{aligned} \tag{13}$$

Classical FMM model [14] suffers from various drawbacks. In order to resolve various shortcomings of FMM network, many researchers modified it to obtain a better accuracy rate. Among all the modifications, EFMM [39] is robust and provides better accuracy rate. Hence, in this paper, EFMM is thoroughly studied.

After the completion of above mentioned process, a set of hyperboxes is generated to indicate the FMM model.

2.2. Enhanced FMM (EFMM) model

Mohammed et al. [39] have proposed an enhanced FMM to resolve different problems of FMM neural network. EFMM uses a new expansion criteria, a new set of overlap rules and contraction rules, which are discussed as follows.

Expansion Criteria: Simpson developed Eq. (3) to test the expandability of a hyperbox to include a new input pattern. Eq. (3) creates some overlapping hyperboxes, due to the comparison of sum of all dimensions with $n\theta$. To overcome the above mentioned drawback, Eq. (14) is used in EFMM to test whether the hyperbox is expandable or not.

$$\mathrm{Max}_n(BU_{ji}, Ba_{hi}) - \mathrm{Min}_n(BL_{ji}, Ba_{hi}) \leq \theta. \tag{14}$$

Eq. (14) compares all hyperbox dimensions individually with the expansion criteria (θ). Here, the expansion procedure is applied only if all the dimensions of hyperbox do not surpass the expansion criteria. Hence, it minimizes the overlapped region between two different class hyperboxes.

Overlap test cases: The overlap cases adopted by traditional FMM are not sufficient to identify all types of overlappings. Therefore, EFMM have used a new set of overlap test cases to identify

all types of overlapping. Overlap test rules used by EFMM are described as follows:
Case 1:

$$BL_{ji} < BL_{ki} < BU_{ji} < BU_{ki},\ \delta n = \min(BU_{ji} - BL_{ki}, \delta o), \tag{15}$$

Case 2:

$$BL_{ki} < BL_{ji} < BU_{ki} < BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o), \tag{16}$$

Case 3:

$$\begin{aligned} &BL_{ji} = BL_{ki} < BU_{ji} < BU_{ki}, \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{17}$$

Case 4:

$$\begin{aligned} &BL_{ji} < BL_{ki} < BU_{ji} = BU_{ki}, \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{18}$$

Case 5:

$$\begin{aligned} &BL_{ki} = BL_{ji} < BU_{ki} < BU_{ji}, \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{19}$$

Case 6:

$$\begin{aligned} &BL_{ki} < BL_{ji} < BU_{ki} = BU_{ji} \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{20}$$

Case 7:

$$\begin{aligned} &BL_{ji} < BL_{ki} \le BU_{ki} < BU_{ji}, \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{21}$$

Case 8:

$$\begin{aligned} &BL_{ki} < BL_{ji} \le BU_{ji} < BU_{ki} \\ &\delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \end{aligned} \tag{22}$$

Case 9:

$$BL_{ki} = BL_{ji} < BU_{ki} = BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o). \tag{23}$$

Contraction Rule: The contraction rules to solve the overlapping cases as shown in Eq. (15) to (23) are represented by Eq. (24) to (35) respectively.
Case 1:

$$BL_{j\Delta} < BL_{k\Delta} < BU_{j\Delta} < BU_{k\Delta}, BU_{j\Delta}^{new} = BL_{k\Delta}^{new} = \frac{BU_{j\Delta}^{old} + BL_{k\Delta}^{old}}{2}, \tag{24}$$

Case 2:

$$\begin{aligned} &BL_{k\Delta} < BL_{j\Delta} < BU_{k\Delta} < BU_{j\Delta}, \\ &BU_{k\Delta}^{new} = BL_{j\Delta}^{new} = \tfrac{BU_{k\Delta}^{old} + BL_{j\Delta}^{old}}{2}, \end{aligned} \tag{25}$$

Case 3:

$$BL_{j\Delta} = BL_{k\Delta} < BU_{j\Delta} < BU_{k\Delta}, BL_{k\Delta}^{new} = BU_{j\Delta}^{old}, \tag{26}$$

Case 4:

$$BL_{j\Delta} < BL_{k\Delta} < BU_{j\Delta} = BU_{k\Delta}, BU_{j\Delta}^{new} = BL_{k\Delta}^{old}, \tag{27}$$

Case 5:

$$BL_{k\Delta} = BL_{j\Delta} < BU_{k\Delta} < BU_{j\Delta}, BL_{j\Delta}^{new} = BU_{k\Delta}^{old}, \tag{28}$$

Case 6:

$$BV_{k\Delta} < BV_{j\Delta} < BW_{k\Delta} = BW_{j\Delta}, BW_{k\Delta}^{new} = BV_{j\Delta}^{new}, \tag{29}$$

Case 7(a):

$$\begin{aligned} &BL_{j\Delta} < BL_{k\Delta} \le BU_{k\Delta} < BU_{j\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) < (BU_{j\Delta} - BL_{k\Delta}),\ BL_{j\Delta}^{new} = BU_{k\Delta}^{old}, \end{aligned} \tag{30}$$

Case 7(b):

$$\begin{aligned} &BL_{j\Delta} < BL_{k\Delta} \le BU_{k\Delta} < BU_{j\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) > (BU_{j\Delta} - BL_{k\Delta}),\ BU_{j\Delta}^{new} = BL_{k\Delta}^{old}, \end{aligned} \tag{31}$$

Case 8(a):

$$\begin{aligned} &BL_{k\Delta} < BL_{j\Delta} \le BU_{j\Delta} < BU_{k\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) < (BU_{j\Delta} - BL_{k\Delta}),\ BU_{k\Delta}^{new} = BL_{j\Delta}^{old}, \end{aligned} \tag{32}$$

Case 8(b):

$$\begin{aligned} &BL_{k\Delta} < BL_{j\Delta} \le BU_{j\Delta} < BU_{k\Delta} \text{ and} \\ &(BU_{k\Delta} - BL_{j\Delta}) > (BU_{j\Delta} - BL_{k\Delta}),\ BL_{k\Delta}^{new} = BU_{j\Delta}^{old}, \end{aligned} \tag{33}$$

Case 9(a):

$$\begin{aligned} &BL_{j\Delta} = BL_{k\Delta} < BU_{j\Delta} = BU_{k\Delta}, \\ &BU_{j\Delta}^{new} = BL_{k\Delta}^{new} = \tfrac{BU_{j\Delta}^{old} + BL_{k\Delta}^{old}}{2}, \end{aligned} \tag{34}$$

Case 9(b):

$$\begin{aligned} &BL_{k\Delta} = BL_{j\Delta} < BU_{k\Delta} = BU_{j\Delta}, \\ &BU_{k\Delta}^{new} = BL_{j\Delta}^{new} = \tfrac{BU_{k\Delta}^{old} + BL_{j\Delta}^{old}}{2}. \end{aligned} \tag{35}$$

2.3. K-nearest FMM (Kn-FMM) network

Mohammed et al. [38] introduced a variant of traditional FMM neural network, in which K-nearest rule is applied to traditional FMM model. FMM considers the hyperbox with maximum fuzzy MF value for expansion. However, the proposed method has k hyperboxes with k highest fuzzy membership function values. Classical FMM neural network generates a new hyperbox, when the hyperbox with biggest fuzzy MF value fails the expansion criteria to add the current pattern. However, in case of Kn_FMM neural network, hyperbox with next maximum fuzzy MF value is tested for expansion, if the hyperbox with largest fuzzy membership function value fails the expandability criteria. This procedure is continued up to k hyperboxes corresponding to k largest fuzzy membership function values. A fresh hyperbox is generated to add the present training sample only if all the k number of hyperboxes fail to satisfy the expansion test.

Partially biased disorder: The contraction rules of FMM corresponding to overlap test case I and case II give equal priority to both the hyperboxes of different class i.e. after the contraction, irrespective of the size of hyperboxes, equal priority is given to the overlapped hyperboxes. The above mechanism is also considered by EFMM for its contraction case of I and II. This type of contraction is inappropriate, since different hyperboxes have different sizes and contains a different number of training samples. This type of problem can be called as partially biased disorder. In this paper, authors have applied a reasonable priority to each hyperbox that balances the size of hyperboxes according to their size, which is clearly described in Section 3.

Fully biased disorder: The contraction rules of FMM corresponding to overlap test case III, and case IV, give full priority to inner (smaller) hyperbox i.e. after the contraction, the size of inner hyperbox does not change. However, relatively less priority is assigned to the outer hyperbox, whose size is reduced to some extent. The above mechanism is also used by EFMM in the contraction rules i.e. case III–VIII. This type of contraction is not appropriate, since size of the outer hyperbox changes, but size of inner hyperbox remains unchanged. This type of problem can be called as fully biased disorder. In this paper, authors have applied a reasonable priority to each hyperbox that balances the size of inner and outer hyperbox, which is clearly described in the following section.

3. Proposed Fuzzy min max network

Fig. 2 shows the system diagram of proposed FMM (PFMM) neural network classifier. In order to accomplish the classification task, each database is divided into three sets, such as primary

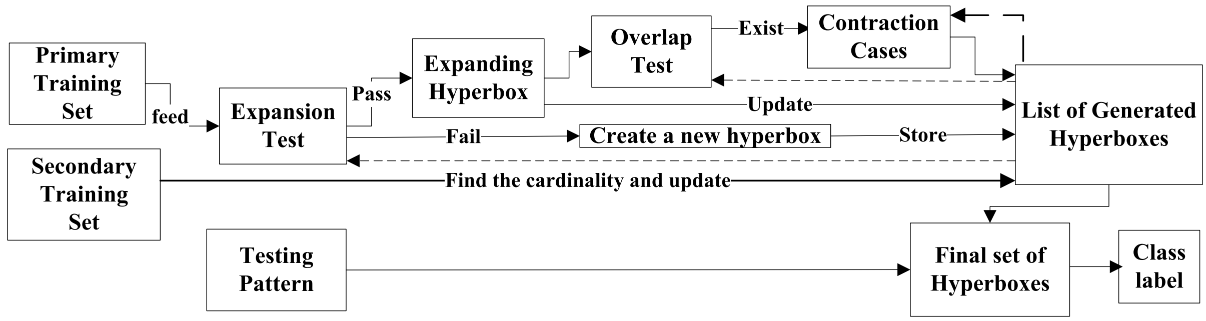

Fig. 2. System Diagram of Proposed FMM (PFMM) Neural Network classifier.

training, secondary training and testing sets. Initially, each training sample of primary training set is checked for the expansion test. In order to include a new training pattern, FMM, EFMM and Kn_FMM approaches arbitrarily select a probable hyperbox for expansion, when multiple hyperboxes have same maximum membership function value. However, such arbitrary selection is inappropriate. Therefore, the proposed method uses centroid and k-highest mechanism based rule to break the tie. If any of the primary training sample fails to satisfy the expansion test, then it is treated as a new point hyperbox. If the training sample passes the expansion test to include in the same class hyperbox, then that hyperbox should be expanded. The expanded hyperbox is checked for the overlap test. In order to know the existence of any overlap between the hyperboxes of different classes, overlap test is performed. In case of proposed method, a new set of overlap test rules are introduced that tackle the missed overlap test rules of FMM, Kn_FMM and unnecessary overlap test rules of EFMM. If any overlap exists between two different class hyperboxes, it is eliminated using the appropriate proposed contraction mechanism. The contraction rules considered by the existing techniques inappropriately resolve the overlapped regions. In a view to accurately perform contraction, the proposed method incorporates side length of hyperboxes to resolve the partially biased disorder and fully biased disorder. After completing all of the three processes such as expansion test, overlap test cases, and contraction cases, a set of hyperboxes is generated. Secondary training pattern is used in this work to generate a final set of efficient hyperboxes by refining the hyperboxes generated using primary training set. Generally, FMM based methods generate some inefficient hyperboxes during training phase. Furthermore, a hyperbox created during training phase may become inefficient due to a redundant training set. In order to solve the above problems, a novel class reversal mechanism is proposed in this paper which increases the accuracy rate and reduces the network complexity. After completing all these processes, a refined set of hyperboxes is generated to classify the testing patterns. The four modifications of the proposed method are clearly described in this Section:

3.1. A new (centroid + K-highest) rule to find the expandable hyperbox

In order to include a new input pattern, FMM and EFMM always considers the hyperbox with maximum fuzzy membership function value to be the probable expandable hyperbox. However, a new hyperbox is created to include the input pattern, if the hyperbox with largest fuzzy MF value fails the expansion criteria. FMM and EFMM use Eqs. (3) and (14) to check the hyperbox expandability respectively. Moreover, if multiple hyperboxes have equal maximum fuzzy MF value and all satisfy the expansion test simultaneously, then FMM and EFMM arbitrarily select a hyperbox to break the tie.

However, such an arbitrary selection of hyperbox leads to an inaccurate set of hyperboxes, since the hyperboxes with same fuzzy membership value may have different size; and hence, different impact to hyperbox selection system. In order to break the tie, the proposed method considers the centroid of all the hyperboxes that constitutes highest fuzzy membership value. The hyperbox whose centroid is nearest to the input pattern is selected for expansion.

$$C_{ji} = \frac{1}{N_j}\sum_{i=1}^{k} Ba_{hi} \tag{36}$$

$$D\left(BA_h, C_j\right) = \sqrt{\sum_{i=1}^{n}\left(Ba_{hi} - C_{ji}\right)^2} \tag{37}$$

where, C_{ji} represents the centroid of *jth* hyperbox, Ba_{hi} represents the *ith* dimension in *hth* input pattern, N_j represents the number of input patterns contained in *jth* hyperbox. $D\left(BA_h, C_j\right)$ is the distance between the input BA_h and centroid C_j of each winner hyperbox.

Moreover, to reduce the network complexity arising due to larger number of small-sized hyperboxes, "K-highest fuzzy membership value" is used. FMM and EFMM only considers the hyperboxe(s) with biggest fuzzy membership value. But this leads to unwanted creation of hyperboxes. In order to solve the above problem, instead of testing the hyperboxe(s) with highest fuzzy membership value, k number of hyperboxes with k-highest fuzzy membership value are considered, and k-highest rule is applied. A new hyperbox is created only, when k number of hyperboxes fail to satisfy the expansion test.

Note 1 (*k-highest Rule*)**.** First consider the hyperbox with maximum fuzzy MF value. Centroid is used to break the tie as described above. If this hyperbox satisfies the expandability test, then it is expanded to include the input pattern. If the above hyperbox is unable to satisfy the expandability criteria, then the hyperbox with next maximum fuzzy MF value is considered. This procedure is continued up to k hyperboxes. If the *k*th hyperbox also failed to satisfy Eq. (3), then a new hyperbox is created to include the input pattern.

3.2. A new compact set of overlapping rules

The overlap test rules considered by FMM, skips some of the overlapping cases such as (a) $BL_{ji} = BL_{ki} < BU_{ji} < BU_{ki}$, (b) $BL_{ji} < BL_{ki} < BU_{ji} = BU_{ki}$, (c) $BL_{ki} = BL_{ji} < BU_{ki} < BU_{ji}$, and (d) $BL_{ki} < BL_{ji} < BU_{ki} = BU_{ji}$ which are depicted in Fig. 3.

EFMM considers the overlapping scenario, which were not considered by FMM. However, few overlapping cases as proposed by EFMM are similar and ambiguous. Hence, same contraction mechanism can be applied to more than one overlap test rules

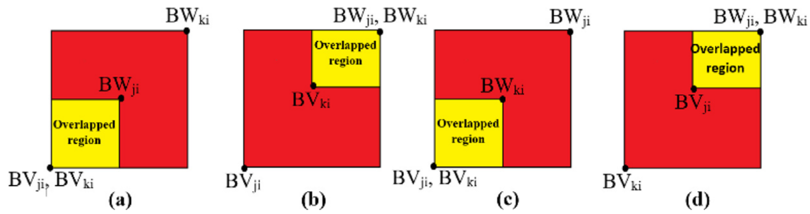

Fig. 3. Different overlapping cases skipped by FMM.

of EFMM. Therefore, these overlapping test rules can be grouped into a single type. By thoroughly investigating the overlapping test rules proposed by EFMM, it is noticed that:

(1) Overlap test cases 3, 6, and 8 of EFMM are applied to the conditions

$$BL_{ji} = BL_{ki} < BU_{ji} < BU_{ki},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e. Case 3, Eq. number (17)),

$$BL_{ki} < BL_{ji} < BU_{ki} = BU_{ji}\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e Case 6, Eq. number (20)), and

$$BL_{ki} < BL_{ji} \leq BU_{ji} < BU_{ki}\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e. Case 8, Eq. number (22)), respectively.

Hence, these can be grouped into a single overlap test rule (overlap case 3 of the proposed method i.e Eq. number (40),

$$BL_{ki} \leq BL_{ji} < BU_{ji} \leq BU_{ki},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

).

(2) Overlap test cases 4, 5 and 7 of EFMM are applied to the conditions

$$BL_{ji} < BL_{ki} < BU_{ji} = BU_{ki},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e. Case 4: Eq. number (18)),

$$BL_{ki} = BL_{ji} < BU_{ki} < BU_{ji},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e. Case 5: Eq. number (19)), and

$$BL_{ji} < BL_{ki} \leq BU_{ki} < BU_{ji},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

(i.e. Case 7: Eq. number (21)) respectively.

Hence, these can be grouped into a single overlap test rule (overlap case 4 of the proposed FMM model i.e. Eq. no. (41),

$$BL_{ji} \leq BL_{ki} < BU_{ki} \leq BU_{ji}\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o),$$

).

(3) Overlap test cases 2, and 9 of EFMM are applied to the conditions

$$BL_{ki} < BL_{ji} < BU_{ki} < BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o),$$

(i.e. Case 2, Eq. number (16)), and

$$BL_{ki} = BL_{ji} < BU_{ki} = BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o).$$

(i.e. Case 9, Eq. number (23)) respectively.

Hence, these two cases can be grouped into a single overlap test rule (overlap case 2 of the proposed FMM model i.e. Eq. no. (39),

$BL_{ki} \leq BL_{ji} < BU_{ki} \leq BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o),$)

In order to take a compact set of overlapping cases, the first case of FMM was considered, and three new overlapping test cases have been proposed to include the important overlapping cases between two hyperboxes of different class. The overlapping cases of the proposed FMM network is as follows.

Case 1:

$$BL_{ji} < BL_{ki} < BU_{ji} < BU_{ki},\ \delta n = \min(BU_{ji} - BL_{ki}, \delta o), \tag{38}$$

Case 2:

$$BL_{ki} \leq BL_{ji} < BU_{ki} \leq BU_{ji},\ \delta n = \min(BU_{ki} - BL_{ji}, \delta o), \tag{39}$$

Case 3:

$$BL_{ki} \leq BL_{ji} < BU_{ji} \leq BU_{ki},\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \tag{40}$$

Case 4:

$$BL_{ji} \leq BL_{ki} < BU_{ki} \leq BU_{ji}\\ \delta n = \min(\min(BU_{ji} - BL_{ki}, BU_{ki} - BL_{ji}),\ \delta o), \tag{41}$$

Firstly, set $\delta o = 1$. If $\delta o - \delta n > 0$, set $\Delta = i$, $\delta o = \delta n$ and continue to test the presence of an overlapped region for next dimension. Else, set $\Delta = -1$ and the testing stops.

3.3. Novel prioritized contraction rules

In order to fix the first and second overlapping cases, the contraction mechanism of FMM and EFMM divides the overlapped regions into four equal sized areas. Out of the above equal sized areas, down-left and up-right parts are contained in two separate hyperboxes, which are shown in Fig. 4(b). It can be seen in Fig. 4(b) that two different sized hyperboxes are settled with the same amount of overlapped regions. This type of contraction mechanism allocates the overlapping region by breaking it along its center point. It is not appropriate, since breaking along the center point allocates equal amount of overlapped regions to left and right hyperboxes, which are of different sizes. Furthermore, the FMM and EFMM give entire priority to inner (smaller) hyperbox for the overlapping cases of III, IV and III to VIII respectively. Giving entire priority to inner (smaller) hyperbox is not appropriate, since this scenario eliminates more region from the outer (larger) hyperbox which will degrade the system performance. Hence, in this paper, to resolve the above mentioned drawbacks, a new set of contraction case is introduced which assigns priority to minimum and maximum points of hyperboxes to get the separation point. Assigned priorities depend on the side-length of respective hyperbox.

Side-length of *ith* hyperbox S_i can be calculated using Eq. (42)

$$S_i = \sum_{x=1}^{n} |BU_{ix} - BL_{ix}| \tag{42}$$

The contraction rules of the proposed FMM network is as follows:

Case 1:

$$BL_{j\Delta} < BL_{k\Delta} < BU_{j\Delta} < BU_{k\Delta},\\ BU_{j\Delta}^{new} = BL_{k\Delta}^{new} = (\in)\, BU_{j\Delta}^{old} + (1-\in)\, BL_{k\Delta}^{old}, \tag{43}$$

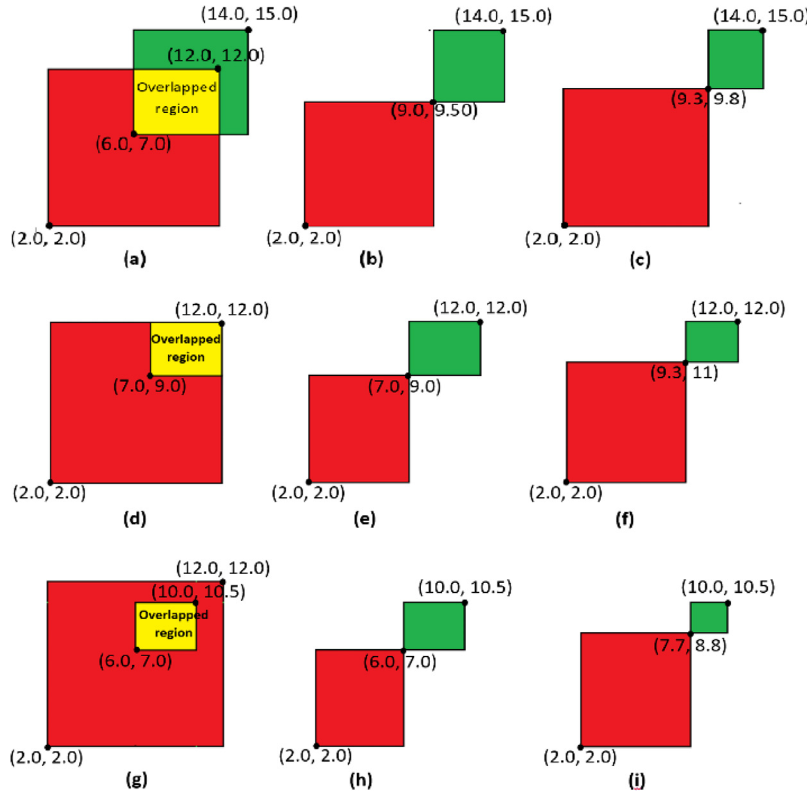

Fig. 4. Usefulness of the PFMM contraction mechanisms than the FMM [14], Kn_FMM, and EFMM [39]. (a), (d) and (g) indicates 3 types of overlapped rules. (b), (e) and (h) designates the corresponding contraction rules considered by state-of-the-art models. (c), (f) and (i) shows the corresponding contraction rules introduced by the PFMM model.

where $\in = \frac{S_j}{S_j + S_k}$, S_j and S_k represent Side-Length of jth and kth hyperbox respectively

Case 2:

$$BL_{k\Delta} \le BL_{j\Delta} < BU_{k\Delta} \le BU_{j\Delta}, \quad BU^{new}_{k\Delta} = BL^{new}_{j\Delta} = (\in) BU^{old}_{k\Delta} + (1- \in) BL^{old}_{j\Delta}, \tag{44}$$

Case 3(a):

$$BL_{k\Delta} \le BL_{j\Delta} < BU_{j\Delta} \le BU_{k\Delta} \text{ and } \left(BL_{j\Delta} - BL_{k\Delta}\right) \le \left(BU_{k\Delta} - BU_{j\Delta}\right)$$
$$BL_{k\Delta} = BU_{j\Delta} = BU_{j\Delta} - BU_{j\Delta}\left(\frac{BU_{j\Delta} - BL_{j\Delta}}{\left(BU_{k\Delta} - BL_{k\Delta}\right) + \left(BU_{j\Delta} - BL_{j\Delta}\right)}\right) \tag{45}$$

Case 3(b):

$$BL_{k\Delta} \le BL_{j\Delta} < BU_{j\Delta} \le BU_{k\Delta} \text{ and } \left(BL_{j\Delta} - BL_{k\Delta}\right) > \left(BU_{k\Delta} - BU_{j\Delta}\right)$$
$$BU_{k\Delta} = BL_{j\Delta} = BL_{j\Delta} + BL_{j\Delta}\left(\frac{BU_{j\Delta} - BL_{j\Delta}}{\left(BU_{k\Delta} - BL_{k\Delta}\right) + \left(BU_{j\Delta} - BL_{j\Delta}\right)}\right) \tag{46}$$

Case 4(a):

$$BL_{j\Delta} \le BL_{k\Delta} < BU_{k\Delta} \le BU_{j\Delta} \text{ and } \left(BL_{k\Delta} - BL_{j\Delta}\right) \le \left(BU_{j\Delta} - BU_{k\Delta}\right)$$
$$BL_{j\Delta} = BU_{k\Delta} = BU_{k\Delta} - BU_{k\Delta}\left(\frac{BU_{k\Delta} - BL_{k\Delta}}{\left(BU_{j\Delta} - BL_{j\Delta}\right) + \left(BU_{k\Delta} - BL_{k\Delta}\right)}\right) \tag{47}$$

Case 4(b):

$$BL_{j\Delta} \le BL_{k\Delta} < BU_{k\Delta} \le BU_{j\Delta} \text{and } \left(BL_{k\Delta} - BL_{j\Delta}\right) > \left(BU_{j\Delta} - BU_{k\Delta}\right)$$
$$BU_{j\Delta} = BL_{k\Delta} = BL_{k\Delta} + BL_{k\Delta}\left(\frac{BL_{k\Delta} - BU_{k\Delta}}{\left(BL_{j\Delta} - BU_{j\Delta}\right) + \left(BL_{k\Delta} - BU_{k\Delta}\right)}\right) \tag{48}$$

The Fig. 4 depicts the usefulness of proposed contraction mechanism. From Fig. 4, the following conclusions can be drawn:

1. In Fig. 4(a), smaller hyperbox area is 64, whereas the area of larger hyperbox is 100. If FMM, Kn_FMM, and EFMM based contraction mechanism is used to Fig. 4(a), then the total amount of area loss occurred by these techniques is union of hyperboxes depicted in Fig. 4(a) minus union of hyperboxes depicted in Fig. 4(b) i.e. (64+100-30)-(27.5+52.5) = 54. However, the total loss of area occurred due to the proposed FMM method is (64+100-30)-(24.44+56.94) = 52.5. Moreover, in case of proposed FMM network, the size of hyperbox before and after contraction are proportional to each other. On the other hand, start-of-the-art methods lose sizeable regions from larger hyperboxes.
2. The loss of total information occurred due to applying FMM, Kn_FMM, and enhanced FMM networks for Fig. 4(d) is (15+100-15)-(15+35) = 50, whereas the loss of total amount information occurred due to applying proposed FMM method is (15+100-15)-(2.7+65.7) = 31.6
3. The total information lost by FMM, Kn_FMM, and EFMM for Fig. 4(g) is (14+100-14)-(14+20) = 66, whereas the loss of total information occurred by applying the proposed FMM is (14+100-14)-(38.76+3.91) = 57.33.
4. Thus, the larger hyperbox relatively loses more overlapping regions, when the contraction mechanisms of FMM, Kn_FMM, and EFMM are applied. Hence, classification accuracy rate is reduced. However, the overlapping region is evenly partitioned by the proposed FMM contraction mechanisms by using some priority, where priority corresponds to hyperbox size.

Note 2. When a single point of a class is enclosed in a hyperbox of different class i.e. ($BL_{j\Delta} < BL_{k\Delta} = BU_{k\Delta} < BU_{j\Delta}$ and

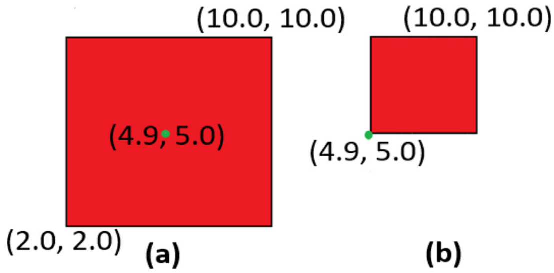

Fig. 5. Contraction rule followed by EFMM when $BL_{j\Delta} < BL_{k\Delta} = BU_{k\Delta} < BU_{j\Delta}$ and $BL_{k\Delta} < BL_{j\Delta} = BU_{j\Delta} < BU_{k\Delta}$ (a) represents overlapped case, (b) designates the corresponding contraction case considered by enhanced FMM.

Table 1
A summary of FMM variants.

Model	Missed overlap rules	Missed contraction cases	Partially biased disorder	Fully biased disorder	Arbitrary selection
FMM	Y	Y	Y	Y	Y
AFMN	Y	Y	Y	Y	Y
FMM-GA	Y	Y	Y	Y	Y
MFMM	Y	Y	Y	Y	Y
WFMM	Y	Y	Y	Y	Y
GFMM	Y	Y	Y	Y	Y
EFMM	N	N	Y	Y	Y
Kn_FMM	Y	Y	Y	Y	Y
PFMM	N	N	N	N	N

Y= Yes, limitation exists, N= No, limitations not exists.

$BL_{k\Delta} < BL_{j\Delta} = BU_{j\Delta} < BU_{k\Delta}$) as defined in Fig. 5(a), the contraction performed by the enhanced FMM is shown in Fig. 5(b). This contraction mechanism is inappropriate due to the following reason: If contraction as defined in Fig. 5(b) is performed, then larger hyperbox loses the sizable amount of information, which is inappropriate.

Note 3. Moreover, FMM and enhanced FMM networks have not used the overlap test rules as mentioned in Fig. 6(a) and (b) for contraction, it happens when $(BU_{k\Delta} - BL_{j\Delta}) = (BU_{j\Delta} - BL_{k\Delta})$ and $(BU_{j\Delta} - BL_{k\Delta}) = (BU_{k\Delta} - BL_{j\Delta})$ respectively. However, the proposed FMM neural network considers the above overlapped cases for contraction.

3.4. Novel trimming strategy

Sometimes, due to an inappropriate training set, hyperboxes created using FMM and EFMM neural network are inefficient to represent each class of data. Hence, the class label of various hyperboxes are to be revised using a second set of training patterns, secondary training set. So, in the first phase, hyperboxes are created using the primary training set. In the second phase, class labels of some/many hyperboxes are allotted through a secondary training set using an appropriate methodology. The proposed trimming strategy is as follows:

Step 1: First apply the proposed FMM neural network to a primary training set and obtain a temporary set of hyperboxes.

Step 2: For each hyperbox of the temporary set

(1) Find the number of training patterns of the secondary training set lying within it.
(2) Compute the probability of each class label using the cardinality of each class label.
(3) Using the share of each class label, either delete the hyperbox or change the class label or retain the class label of hyperbox using the following cases respectively.

Case1: Delete *hth* hyperbox

$$\frac{\max\left|{}^{h}cy_{y}^{x}\right|}{n_{s}^{h}} \leq 0.3 \tag{49}$$

Case 2: Change the class label to x

$$0.3 < \frac{\max\left|{}^{h}cy_{y}^{x}\right|}{n_{s}^{h}} \text{ and } x \neq y \tag{50}$$

Case 3: No change of the class label

$$0.3 < \frac{\max\left|{}^{h}cy_{y}^{x}\right|}{n_{s}^{h}} \text{ and } x = y \tag{51}$$

where, y is the actual class label of a hyperbox. $\left|{}^{h}cy_{y}^{x}\right|$ represents the number of secondary training patterns with class label x lying in *hth* hyperbox with class label y. $n_s h$ is the number of secondary training patterns lying in *hth* hyperbox and n_s is the total number of secondary training patterns.

Step 3: Apply normal procedure of the proposed FMM to final set of hyperboxes obtained in step 2 to include the current input pattern.

After completion of all the above process, a final set of hyperboxes is obtained.

Table 1 depicts the important properties of FMM and related limitations. Various limitation of FMM are missed overlap rules, missed contraction cases, partially biased disorder, arbitrary selection, and fully biased disorder. From Table 1, it is clear that except PFMM, all other FMM variants suffer from at least two limitations in the training stage.

4. Performance evaluation

4.1. Database

In order to measure the expansion coefficient effects on the accuracy of FMM, EFMM, Kn_FMM and PFMM, expansion coefficient is varied appropriately to classify five benchmark datasets such as hepatitis, Lung cancer, Wisconsin Breast cancer (WBC) [45], Mammographic Mass Dataset and Multiple Features (mfeat-fac) Data Set. WBC dataset contains a total of 194 instances with 33 input features. Lung Cancer dataset contains a total of 27 instances with 56 input features, hepatitis dataset contains a total of 80 instances with 19 input features and Mammographic Mass Dataset contains a total of 830 instances with 5 input features. WBC, Hepatitis and Mammographic Mass Dataset contain a total of two classes, whereas lung cancer database contains three classes. Hence, for the hepatitis, lung cancer, WBC, and Mammographic Mass databases, the instances consisting values for all the attributes, are considered in the experiment. However, all the instances of Multiple Features (mfeat-fac) dataset contains values for all the attributes, but the dataset has larger amount instances. Thus, a fewer amount of instances (500 instances with 216 input features for 10 classes) are used for simulation. Table 2 depicts the simulation statistics for different databases. For all databases, eighty percent of the data samples (arbitrarily selected) are chosen for training and remainder of the data is used for testing.

4.2. Performance comparison

The performance of classification problems can be evaluated using two aspects: Qualitative and Quantitative. In this work, "classification accuracy rate" is considered for the qualitative evaluation, whereas "number of generated hyperboxes" is used

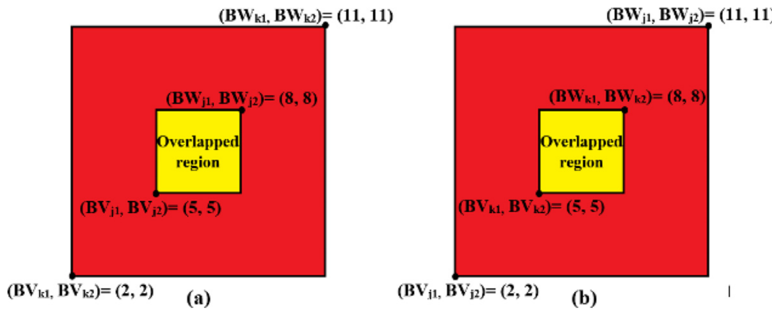

Fig. 6. (a) and (b) are the two different overlapped rules. FMM and EFMM networks are not considered the above overlapping cases for contraction.

Table 2
Databases description.

Databases	Instances		Attributes	Classes
	Total	Used		
Hepatitis	88	80	19	02
Lung cancer	32	27	56	03
WBC	198	194	33	02
Mammographic Mass	960	830	5	02
Multiple Features (mfeat-fac)	2000	500	216	10

to quantitatively evaluate the different techniques. The performance of proposed method is compared with FMM, EFMM and Kn_FMM using hepatitis, lung cancer, WBC, Mammographic Mass and Multiple Features (mfeat-fac) Datasets. The mean accuracy rates of various approaches, while classifying the Hepatitis, lung cancer, WBC, Mammographic Mass and Multiple Features (mfeat-fac) databases are shown in Fig. 7. At the same time, Fig. 8 depicts the mean list of hyperboxes created for various techniques, while classifying the above mentioned databases. The mean accuracy rate and mean number of hyperboxes created for each technique corresponding to various databases are obtained as follows: For each database, use 80% of arbitrarily selected samples for training, and accordingly train the system for each technique. Then, remainder of the samples are used for testing and got the accuracy rate and list of hyperboxes created for each technique respectively for various θ values, where $\theta = 0.05,\ 0.1x/x = 1, 2, \ldots, 9$. Execute the above mentioned procedure for six times (i.e. 6 times for each setting) using various arbitrarily chosen training and testing data. Then, the mean accuracy rates of six trails are calculated, which is the mean accuracy rate of the technique. In the same way, mean list of generated hyperboxes are calculated. Fig. 7 concludes that PFMM provides higher accuracy rate than FMM, EFMM and Kn_FMM. In the same way, From Fig. 8, it is clear that the PFMM generates fewer hyperboxes than FMM, EFMM and Kn_FMM for all the databases. Thus, the proposed technique offers improved performance than FMM, EFMM and Kn_FMM.

The proposed FMM (PFMM) creates fewer number of hyperboxes to classify the databases, when compared with FMM, EFMM and Kn_FMM, and thus it requires very less maintenance. Number of operations of an FMM based methods mainly depends on the number of overlap test cases and corresponding contraction rules. Proposed FMM network uses 4 overlap test rules to efficiently represent all the overlap test cases. Moreover, it uses 4 contraction rules corresponding to four overlap test cases to remove overlap between two different class hyperboxes. Furthermore, a novel class reversal mechanism is proposed in this paper, which increases the accuracy rate and reduces the network complexity. The overlap rules, contraction rules and novel trimming strategy are clearly explained in Section 3. In order to find the computational complexity of various FMM based methods, core i7 processor and 8 GB ram of Lenovo laptop is used. Table 3 shows the computational complexity (in seconds) of different FMM networks for various values of expansion co-efficient. From Table 3, it is clear that computational complexity of the proposed PFMM network for all the databases such as Hepatitis, lung, WBC, Mammogram, and FAC at each value of expansion coefficient is very low, when compared with other state-of-the-art-methods due to its compact set of overlap rules and novel trimming strategy. Computational complexity of different methods is sorted as PFMM <FMM<Kn_FMM<EFMM for all the datasets at each value of expansion co-efficient.

Table 3
Computational complexity of different FMM networks for various values of expansion co-efficient.

Database	Method	Value of expansion co-efficient			
		0.2	0.4	0.6	0.8
WBC	FMM	4.5327	2.8156	2.0170	2.0160
	EFMM	9.9316	8.4615	9.3757	6.3236
	Kn_ FMM	4.6244	4.7767	4.5177	2.6217
	PFMM	**0.9224**	**0.8354**	**0.8578**	**0.8225**
"	FMM	0.0687	0.0713	0.0536	0.0633
	EFMM	0.1190	0.1073	0.1336	0.1165
	Kn_ FMM	0.0611	0.0686	0.0567	0.0795
	PFMM	**0.0121**	**0.0126**	**0.0129**	**0.0119**
Hepatitis	FMM	0.6309	0.1312	0.1487	0.1297
	EFMM	1.1280	0.1458	0.1850	0.1515
	Kn_ FMM	0.6172	0.1377	0.1424	0.1417
	PFMM	**0.0494**	**0.0515**	**0.0625**	**0.0601**
Mammo-gram	FMM	0.6896	0.6942	0.6769	0.6717
	EFMM	44.2924	44.2835	44.2001	42.7230
	Kn_FMM	0.8453	0.7075	0.8561	0.6056
	PFMM	**0.1716**	**0.1867**	**0.1639**	**0.1609**
FAC	FMM	458.548	461.284	469.788	470.569
	EFMM	1256.8	1202.4	1194.3	1251.4
	Kn_ FMM	551.866	553.566	549.953	552.018
	PFMM	**53.8659**	**53.7247**	**52.3762**	**54.1045**

4.3. Application of proposed Fuzzy Min–Max network: finding the best Magnifying factor for classification of BreaKHis database

Fig. 9 shows the example sample of Breast cancer histopathology (BreaKHis) database.

In order to obtain the best magnification factor, the PFMM technique was applied to classify the images of BreaKHis database [46], which is standardized using [47]. The database contains a total of 7909 histopathology images at various magnifying factors such as 40×, 100×, 200×, and 400×. From each magnifying factor, 1000 images are arbitrarily selected for experiment. Out of the selected samples, five hundred samples are of malignant breast tumor, and remainder of the samples belong to benign breast tumor. For each magnifying factor, 80% of the samples are used for training and rest of the 20% samples are used for

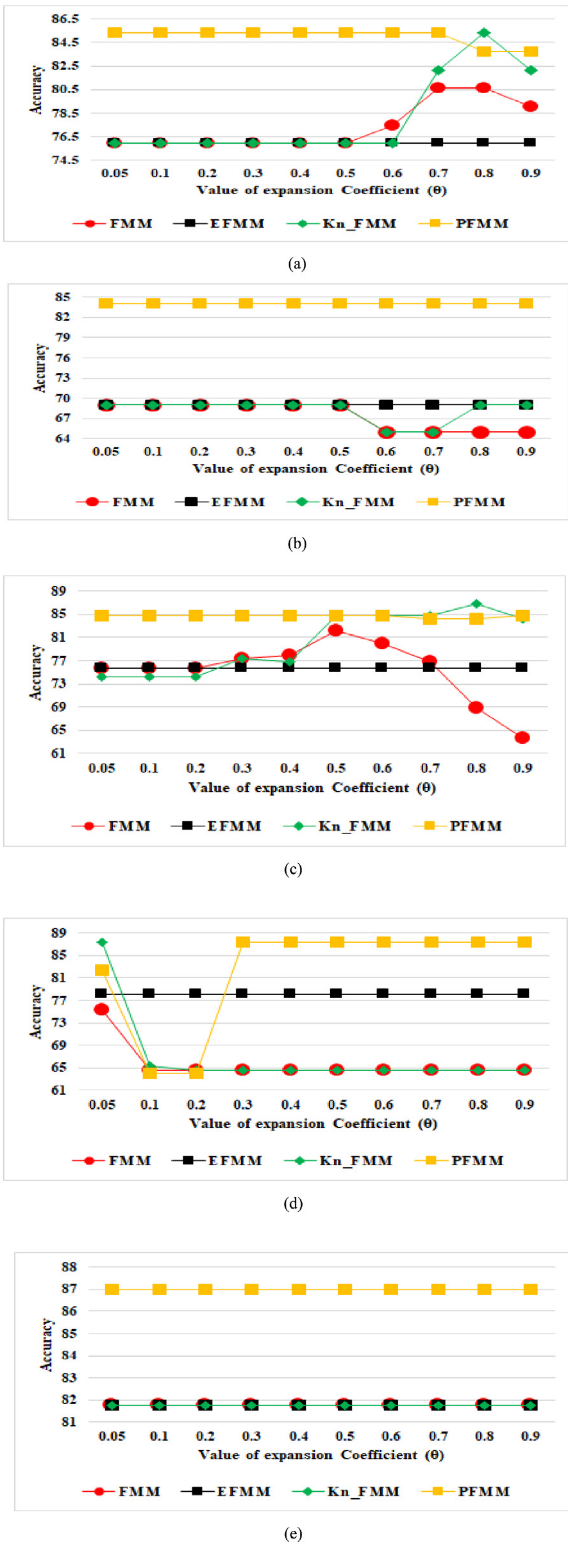

Fig. 7. Accuracy rate of various FMM models for (a) Hepatitis (b) Lung Cancer (c) Wisconsin Breast Cancer (Diagnostic), (d) Mammographic Mass and (e) Multiple Features (mfeat-fac) Dataset..

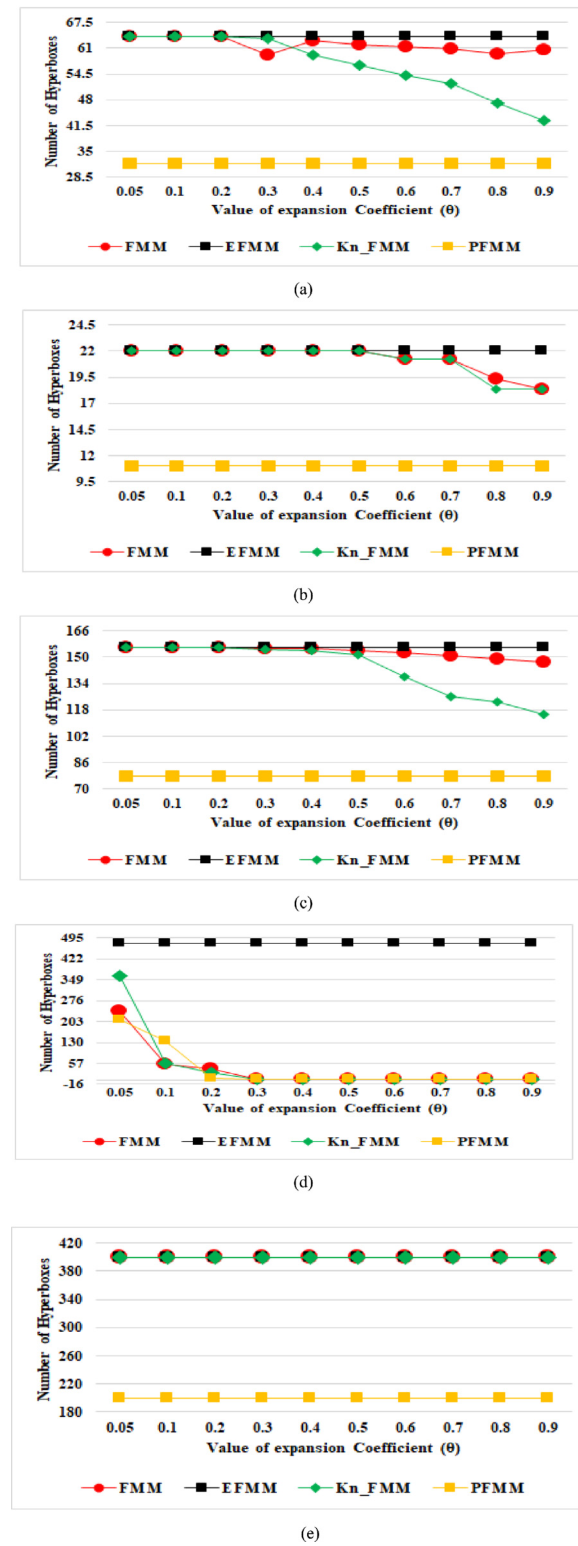

Fig. 8. List of hyperboxes generated during classification of various databases using different FMM networks. (a) Hepatitis (b) Lung Cancer (c) Wisconsin Breast Cancer (d) Mammographic Mass and (e) Multiple Features (mfeat-fac) Dataset..

testing. In order to extract the texture features from each image, gray level co-occurrence matrix (GLCM) is used. These extracted features are then fed to proposed FMM model to classify the database. In order to calculate the characteristics of GLCM, adjacency direction of 0° is considered. In [48], all the features of haralick are computed. At last, thirteen dimensional vector can be achieved by taking average values of haralick features for 3 channels of a color image. In this paper, the mean of eight experiments are reported.

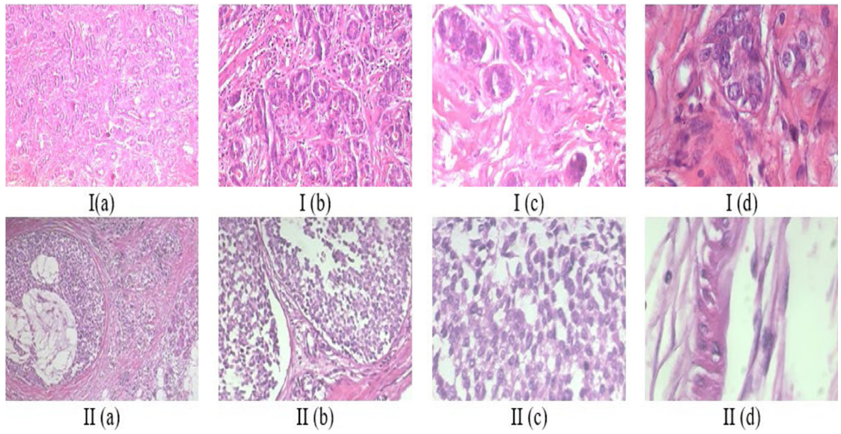

Fig. 9. Example of breast tumor histopathological images acquired at four different magnifications: I(a)-I(d) represent benign tumor samples with magnifying factor of 40×, 100×, 200×, and 400× respectively whereas II(a)-II (d) indicates malignant tumor samples acquired at 40×, 100×, 200×, and 400× magnifications respectively.

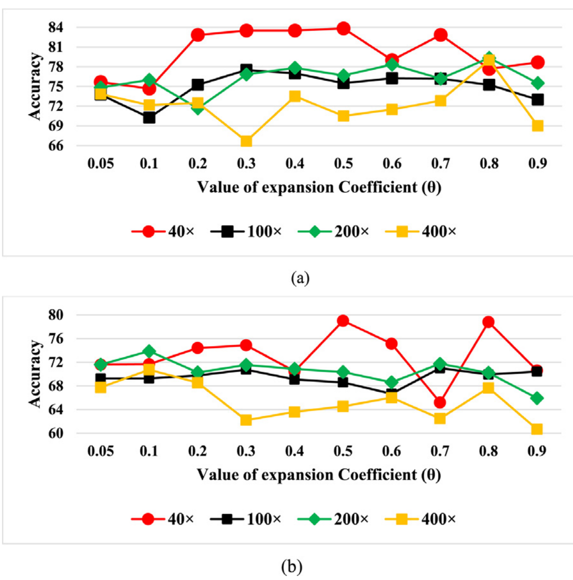

Fig. 10. Accuracy rate of different magnifying factors using PFMM neural network (a) Training/Testing=4 and (b) Training/Testing=2.33.

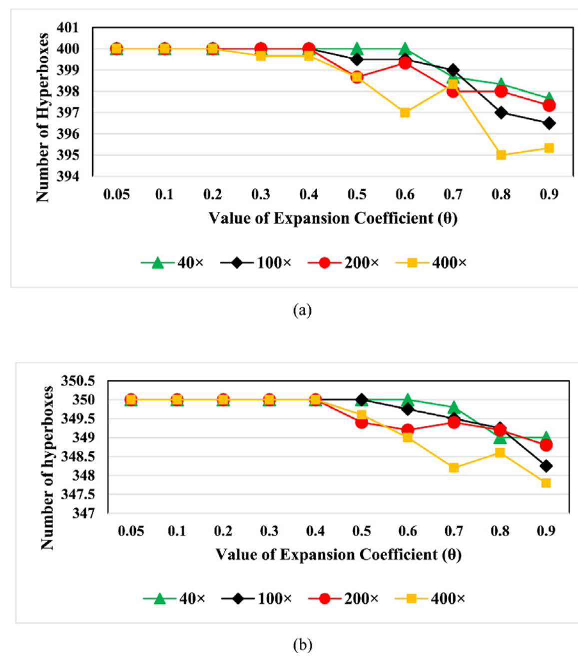

Fig. 11. List of hyperboxes created for different Magnifying factors using PFMM network (a) 80% of data is used for training remainder of the data is used for testing and (b) 70% of data is used for training rest of the data is used for testing.

Figs. 10(a) and 10(b) depicts the accuracy rates of four magnification factors obtained by PFMM. It is clear from Figs. 10(a) and 10(b) that the magnifying factor of 40× provides best accuracy rate for PFMM network. Figs. 11(a) and 11(b) show the list of hyperboxes created for different magnifying factors of BreaKHis database using PFMM neural network. From Figs. 11(a) and 11 (b) clearly present the list of generated hyperboxes for magnification factor of 40× > 100× > 200× > 400×. Table 4 depicts the mean accuracy rate for different magnifying factors of BreaKHis database. It is clear from Table 4 that the magnifying factor of 40× provides superior accuracy rate for both cases as shown in Fig. 10

Table 4
Mean accuracy rate for different magnifying factors of BreaKHis database.

Magnification	Mean accuracy rates	
	Fig. 10(a)	Fig. 10(b)
40×	**80.21667**	**73.1737**
100×	74.99167	69.4634
200×	76.31665	70.4937
400×	72.14933	65.4133

5. Conclusion

In this paper, FMM, EFMM and Kn_FMM are thoroughly studied and a number of limitations were noticed. To resolve the arbitrary selection problem of hyperboxes, when two or more hyperboxes have equal largest fuzzy MF values, the proposed FMM uses centroid to break the tie. Moreover, to resolve the erroneous expansion criteria of FMM and EFMM, k-highest mechanism is used by the proposed method. A new set of overlap test rules are introduced to tackle the skipped overlap test rules of FMM and Kn_FMM and unnecessary overlap test rules of EFMM. Correspondingly, contraction rules are approximately proposed using the side length of hyperbox to fix the partially biased disorder and fully biased disorder. Novel hyperbox trimming strategy is also employed by the proposed FMM network to increase the accuracy rate and to decrease the network complexity.

The proposed FMM network is compared with the FMM, EFMM and Kn_FMM using five benchmark datasets, namely hepatitis, Lung Cancer, Wisconsin Breast Cancer (Diagnostic), Mammographic Mass and Multiple Features (mfeat-fac) Dataset. It was found that the proposed FMM more accurately classifies the datasets than FMM, EFMM and Kn_FMM. Moreover, PFMM creates fewer number of hyperboxes to classify the databases, when compared with FMM, EFMM and Kn_FMM, and thus requires very less maintenance. Furthermore, Due to its compact set of overlap rules and Novel Trimming Strategy, the proposed FMM network has lesser network complexity. In order to find the best magnifying factor for BreaKHis database, proposed FMM network is considered to classify different images of BreaKHis database at different magnifying factors. It was concluded that the magnifying factor of 40× provides good classification performance, while classifying the histopathological images. Future work lies in applying the proposed FMM neural network to various real applications in industry, diagnosis, medicine, business, bioinformatics, remote sensing, text categorization, Face detection, gesture recognition and fault detection etc.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. Graupe, Principles of Artificial Neural Networks, World Scientific, Singapore, 1997.
[2] P.J.G. Lisboa, A review of evidence of health benefit from artificial neural networks in medical intervention, Neural Netw. 15 (1) (2002) 11–39.
[3] J.H.D.G.P. Zhang, Neural networks for classification: A survey, IEEE Trans. Syst. Man, Cybern. C, Appl. Rev. 30 (4) (2000) 451–462.
[4] R. Isola, R. Carvalho, A.K. Tripathy, Knowledge discovery in medical systems using differential diagnosis, LAMSTAR, and K-NN, IEEE Trans. Inf. Technol. Biomed. 16 (6) (2012) 1287–1295.
[5] A. Quteishat, C.P. Lim, J. Tweedale, L.C. Jain, A neural networkbased multi-agent classifier system, Neurocomputing 72 (7) (2009) 1639–1647.
[6] P. Melin, J. Amezcua, F. Valdez, O. Castillo, A new neural network model based on the LVQ algorithm for multi-class classification of arrhythmias, Inform. Sci. 279 (2014) 483–497.
[7] D. Sánchez, P. Melin, O. Castillo, A grey wolf optimizer for modular granular neural networks for human recognition, Comput. Intell. Neurosci. (2017).
[8] D. Sánchez, P. Melin, O. Castillo, Optimization of modular granular neural networks using a firefly algorithm for human recognition, Eng. Appl. Artif. Intell. 64 (2017) 172–186.
[9] D. Sánchez, P. Melin, O. Castillo, Optimization of modular granular neural networks using a hierarchical genetic algorithm based on the database complexity applied to human recognition, Inf. Sci. 309 (2015) 73–101.
[10] M.-Y. Chow, S.O. Yee, Methodology for on-line incipient fault detection in single-phase squirrel-cage induction motors using artificial neural networks, IEEE Trans. Energy Convers. 6 (3) (1991) 536–545.
[11] V.R. Puttige, S.G. Anavatti, Comparison of real-time online and offline neural network models for a UAV, in: International Joint Conference on Neural Networks, IJCNN, 2007, pp. 412–417.
[12] Roger Ratcliff, Connectionist models of recognition memory: constraints imposed by learning and forgetting functions, Psychol. Rev. 97 (2) (1990) 285.
[13] M. McCloskey, N.J. Cohen, Catastrophic interference in connectionist networks: The sequential learning problem, in: G.H. Bower (Ed.), The Psychology of Learning and Motivation, Vol. 24, Academic Press, New York, NY, USA, 1989, pp. 109–165.
[14] P.K. Simpson, Fuzzy min–max neural networks, I. classification, IEEE Trans. Neural Netw. 3 (5) (1992) 776–786.
[15] P.K. Simpson, Fuzzy min–max neural networks—Part 2: Clustering, IEEE Trans. Fuzzy Syst. 1 (1) (1993) 32–45.
[16] S. Grossberg, Adaptive pattern classification and universal recording: I. Parallel development and coding of neural feature detectors, Biol. Cybern. 23 (3) (1976) 121–134.
[17] S. Grossberg, Adaptive pattern classification and universal recording: II. Feedback, expectation, olfaction, illusions, Eiol, Biol. Cybern. 23 (4) (1976) 187–202.
[18] J. Kondala Rao, M. Okade, Role of hyperbox classifiers for color recognition, 2016.
[19] A.M. Quteishat, C.P. Lim, A modified fuzzy min–max neural network and its application to fault classification, in: Soft Computing in Industrial Applications, Springer, Berlin, Heidelberg, 2007, pp. 179–188.
[20] G. Bortolan, I.I. Christov, W. Pedrycz, Hyperbox classifiers for ECG beat analysis, in: Computers in Cardiology, 2007, IEEE, 2007, pp. 145–148.
[21] Anas Quteishat, Chee Peng Lim, Application of the fuzzy min–max neural networks to medical diagnosis, in: International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Springer, Berlin, Heidelberg, 2008, pp. 548–555.
[22] S. Kumar, A. Kumar, V. Bajaj, G.K. Singh, K-highest fuzzy min-max network to classify histopathological images, in: International Conference on Communication and Signal Processing, ICCSP, IEEE, 2019, pp. 0240–0244.
[23] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353.
[24] S. Yilmaz, Y. Oysal, Fuzzy wavelet neural network models for prediction and identification of dynamical systems, IEEE Trans. Neural Netw. 21 (10) (2010) 1599–1609.
[25] C.-F. Juang, T.-C. Chen, W.-Y. Cheng, Speedup of implementing fuzzy neural networks with high-dimensional inputs through parallel processing on graphic processing units, IEEE Trans. Fuzzy Syst. 19 (4) (2011) 717–728.
[26] K.S. Yap, C.P. Lim, M.T. Au, Improved GART neural network model for pattern classification and rule extraction with application to power systems, IEEE Trans. Neural Netw. 22 (12) (2011) 2310–2323.
[27] M. Davanipoor, M. Zekri, F. Sheikholeslam, Fuzzy wavelet neural network with an accelerated hybrid learning algorithm, IEEE Trans. Fuzzy Syst. 20 (3) (2012) 463–470.
[28] M. Pratama, S.G. Anavatti, E. Lughofer, GENFIS: Towards an effective localist network, IEEE Trans. Fuzzy Syst. (2013) http://dx.doi.org/10.1109/TFUZZ.2013.2264938.
[29] B. Gabrys, A. Bargiela, General fuzzy min–max neural network for clustering and classification, IEEE Trans. Neural Netw. 11 (3) (2000) 769–783.
[30] A.V. Nandedkar, P.K. Biswas, A general reflex fuzzy min–max neural network, Eng. Lett. 14 (1) (2007) 195–205.
[31] A. Likas, Reinforcement learning using the stochastic fuzzy min–max neural network, Neural Process. Lett. 13 (3) (2001) 213–220.
[32] A. Rizzi, M. Panella, F.M.F. Mascioli, Adaptive resolution min–max classifiers, IEEE Trans. Neural Netw. 13 (2) (2002) 402–414.
[33] A. Bargiela, W. Pedrycz, M. Tanaka, An inclusion/exclusion fuzzy hyperbox classifier, Int. J. Knowl.-Based Intell. Eng. Syst. 8 (2) (2004) 91–98.
[34] H.J. Kim, H.S. Yang, A weighted fuzzy min–max neural network and its application to feature analysis, in: L. Wang, K. Chen, Y. Ong (Eds.), Advances in Natural Computation (Lecture Notes in Computer Science), Vol. 3612, Springer-Verlag, Berlin, Germany, 2005, pp. 1178–1181.
[35] A.V. Nandedkar, P.K. Biswas, A fuzzy min–max neural network classifier with compensatory neuron architecture, IEEE Trans. Neural Netw. 18 (1) (2007) 42–54.
[36] H. Zhang, J. Liu, D. Ma, Z. Wang, Data-core-based fuzzy min–max neural network for pattern classification, IEEE Trans. Neural Netw. 22 (12) (2011) 2339–2352.
[37] A. Quteishat, C.P. Lim, A modified fuzzy min–max neural network with rule extraction and its application to fault detection and classification, Appl. Soft Comput. 8 (2) (2008) 985–995.
[38] M.F. Mohammed, C.P. Lim, Improving the fuzzy min-max neural network with a K-nearest hyperbox expansion rule for pattern classification, Appl. Soft Comput. 52 (2017) 135–145.
[39] M.F. Mohammed, C.P. Lim, An enhanced fuzzy min–max neural network for pattern classification, IEEE Trans. Neural Networks Learn. Syst. 26 (3) (2015) 417–429.
[40] M.F. Mohammed, C.P. Lim, A new hyperbox selection rule and a pruning strategy for the enhanced fuzzy min–max neural network, Neural Netw. 86 (2017) 69–79.
[41] S. Kumar, A. Kumar, V. Bajaj, G.K. Singh, An improved fuzzy min-max neural network for data classification, IEEE Trans. Fuzzy Syst. (2019) http://dx.doi.org/10.1109/TFUZZ.2019.2924396.
[42] J. Liu, Y. Ma, H. Zhang, H. Su, G. Xiao, A modified fuzzy min–max neural network for data clustering and its application on pipeline internal inspection data, Neurocomputing 238 (2017) 56–66.
[43] S. Kumar, A. Kumar, V. Bajaj, G.K. Singh, Histopathology image classification using enhanced fuzzy min-max network, in: International Conference on Communication and Signal Processing, ICCSP, IEEE, 2020, pp. 0488–0492.

[44] S. Kumar, A. Kumar, V. Bajaj, G.K. Singh, B. Kuldeep, A fuzzy min-max neural network based classification of histopathology images, in: International Conference on Signal Processing and Communication, ICSC, IEEE, 2019, pp. 143–146.
[45] K. Bache, M. Lichman, UCI Machine Learning Repository, School Inf. Comput. Sci. Univ. California, Irvine, CA, USA, 2013, [Online]. Available: http://archive.ics.uci.edu/ml.
[46] F. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, IEEE Trans. Biomed. Eng. 63 (7) (2016) 1455–1462.
[47] A.L. Mescher, Junqueiras Basic Histology: Text and Atlas, McGraw-Hill, NewYork, NY, USA, 2013.
[48] R. Haralick, et al., Textural features for image classification, IEEE Trans. Syst. Man Cybern. SMC-3 (6) (1973) 610–621.