Received March 24, 2019, accepted April 11, 2019, date of publication April 18, 2019, date of current version May 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2911955

A Critical Review on Selected Fuzzy Min-Max Neural Networks and Their Significance and Challenges in Pattern Classification

ESSAM ALHROOB[1,2], MOHAMMED FALAH MOHAMMED[2], CHEE PENG LIM[3], AND HAI TAO[1]

[1]School of Computer Science, Baoji University of Arts and Sciences, Baoji 721000, China
[2]Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, Kuantan 26300, Malaysia
[3]Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, VIC 3220, Australia

Corresponding author: Mohammed Falah Mohammed (falah@ump.edu.my)

The authors gratefully acknowledge the financial support of the Universiti Malaysia Pahang grants (RDU180369, PGRS1803110) and the doctoral scientific research initial funding project of Baoji University of Arts and Sciences (ZK2018062) for this work.

ABSTRACT At present, pattern classification is one of the most important aspects of establishing machine intelligence systems for tackling decision-making processes. The fuzzy min–max (FMM) neural network combines the operations of an artificial neural network and fuzzy set theory into a common framework. FMM is considered one of the most useful neural networks for pattern classification. This paper aims to 1) analyze the FMM neural network in terms of its impact in addressing pattern classification problems; 2) examine models that are proposed based on the original FMM model (i.e., existing FMM-based variants); 3) identify the challenges associated with FMM and its variants, and; 4) discuss future trends and make recommendations for improvement. The review is conducted based on a methodical protocol. Through a rigorous searching and filtering process, the relevant studies are extracted and comprehensively analyzed to adequately address the defined research questions. The findings indicate that FMM plays a critical role in providing solutions to pattern classification issues. The FMM model and a number of FMM-based variants are identified and systematically analyzed with respect to their aims, improvements introduced and results achieved. In addition, FMM and its variants are critically analyzed with respect to their benefits and limitations. This paper shows that the existing FMM-based variants still encounter issues in terms of the learning process (expansion, overlap test, and contraction), which influence the classification performance. Based on the review findings, research opportunities are suggested to propose a new model to enhance the number of existing FMM models, particularly in terms of their learning process by minimizing hyperbox overlap pertaining to different classes as well as avoiding membership ambiguity of the overlapped region. In short, this review provides a comprehensive and critical reference for researchers and practitioners to leverage FMM and its variants for undertaking pattern classification tasks.

INDEX TERMS Fuzzy min–max, pattern classification, neural network, FMM models.

I. INTRODUCTION

Humans always explore a variety of methods to design and build intelligent machines [1], [2]. One effective strategy for achieving this goal is to study the way humans think and act [3]. Our human brain it consists of millions of inter-connected neurons. These neurons assist us in understanding patterns, performing inference, and making decisions [4], [5]. Over the years, researchers have attempted to simulate the human brain and develop intelligent machines that have the capabilities of our brain [6], [7].

Fuzzy set theory and neural networks are two complementary methods for modelling the capabilities of the human brain [8], [9]. Fuzzy set theory provides a means to represent higher-level human inference and reasoning processes [10], [11]. It is useful for modelling the psychological nature of the human mind [12]. On the other hand, neural networks attempt to model the information processing capabilities of the brain. Specifically, a neural network models the lower-level processes of the human brain [13]. It consists of

The associate editor coordinating the review of this manuscript and approving it for publication was Jeonghwan Gwak.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

millions of interconnected processing elements that simulate the biological neurons [14]. The properties of neurons and the connectivity between neurons are topologically simplified to resemble the computational capacity of the human brain [15], [16]. Moreover, a neural network can operate by approximating the cognitive and physiological properties of the human brain [17], [18].

Both fuzzy and neural computational methodologies have been used to provide a foundation of the essential behaviour pertaining to the human mind as an intelligent system [19]. In addition, both methodologies have been combined to form fuzzy neural networks, which is useful for pattern recognition and classification [20]. This function is related to the computational efficiency of neural networks in terms of the learning process and the ability of fuzzy set theory to represent complex decision boundaries, thereby making fuzzy neural networks highly effective for pattern analysis [21], [22].

On the basis of the advantages of synthesising fuzzy set and neural networks to solve pattern classification problems, the fuzzy min–max (FMM) neural networks were proposed by Simpson, one for supervised classification and another unsupervised clustering [23], [24]. This study is focused on the supervised learning, and hereafter FMM refers to the supervised version for pattern classification. In essence, the FMM network structure is constructed from hyperboxes. Each hyperbox is defined by its minimum and maximum points, which are encoded from the input patterns [23]. The FMM learning process consists of three steps: expansion, overlapping test, and contraction [25]. Through these processes, FMM can support online learning by creating new hyperboxes, associating them with new classes, and refining the existing classes without requiring the process of re-training [23], [26].

Various models have been introduced based on the original FMM network, with the aim to achieve better classification performance. Among the key variants of FMM include the general FMM (GFMM) neural network [27], inclusion/exclusion fuzzy hyperbox classifier (EFC) [28], Adaptive inclusion/exclusion fuzzy hyperbox classifier [29], neural network classifier (FMCN) [30], general reflex FMM (GRFMM) neural network [31], data core-based FMM neural network (DCFMN) [32], multilevel FMM (MLF) network classification [21], FMNWSM neural network [33], enhanced FMM (EFMM) neural network [34], modified FMM neural network for data with mixed attributes (MFMMN) [35] and k-nearest FMM (KnFMM) neural network [36]. However, most of these models still suffer from certain shortcomings in their learning algorithms [36]. Indeed, there are opportunities to further conduct investigations on FMM and it variants, in order to identify the challenges and suggest solutions that can improve the efficacy of FMM-based models for pattern classification. Therefore, this research aims to provide a comprehensive review on FMM and its variants, and determine their impact on pattern classification. A detailed examination on the objectives and, improvements introduced in each identified FMM variant is conducted. In this regard, each FMM variant is critically studied and analysed from both advantages and disadvantages perspectives. In addition, further research opportunities to undertake the challenges are identified.

This research makes contributions to researchers and practitioners in pattern classification by providing a clear evaluation on the advantages and disadvantages of FMM and its variants. Based on the analysis of this research future research directions to address the identified issues are presented. As a results, researchers and practitioners are able to enhance the quality of FMM classification by conducting new theoretical as well as empirical studies.

The rest of this paper is organised as follows. Section II discusses the existing studies related to this research. Section III elaborates the review protocol used in this study. Section IV presents the results with a detailed discussion on each defined research question. Section V discusses the findings. Section VI provides the limitations of this review, while Section VII concludes the study.

II. RELATED STUDIES

This study is motivated by a number of factors. Firstly, to the best of our knowledge, no study has yet been conducted on FMM in pattern classification with regard to analysing the impact of FMM in handling pattern classification problems, its variants, and the associated challenges. While several reviews on pattern classification are available, most of them are not specifically focused on FMM but on general pattern classification instead. Secondly, FMM is effective for pattern classification, which makes it an important model to be explored. Finally, this review acts as a crucial knowledge sources for researchers and practitioners by providing a clear analysis and evaluation on FMM-based classifiers. On the basis of the gathered literature, six review studies on FMM and/or related methods for pattern classification have been collected. Table 1 presents the related review studies with respect to the focus, difference, and similarity of findings as compared with those in our study.

García-Laencina *et al.* presented a review on pattern classification [37]. The focus was to provide a critical analysis of the missing data challenge in pattern classification tasks and present a synopsis for comparing pattern classification models with missing data. The review concluded that finding the appropriate treatment selection to address the missing data issue is a difficult and complicated task. Although their review provided an overview of pattern classification and a detailed description and analysis of the missing data problem, the FMM pattern classifier was not covered in the review because only the most essential and well-known missing data models were examined.

Another review of pattern classification was documented in [38]. This review focused on mislabelled data problems in pattern classification and the solutions with the existing methods models in the literature. The single-, ensemble- and local-learning methods were identified as the existing methods that could be used to measure and detect mislabelled training

TABLE 1. Summary of findings of related studies.

Study reference	Study focus	Similar findings compared with the present review	Uncovered findings added to the present review
Laencina et al. [37]	-Problem of missing data in pattern classification tasks.	-Gaps in pattern classification tasks in terms of misclassification issue.	-Detailed investigation of the significant effect of FMM on pattern classification. -Detailed explanations for other uncovered and recent FMM models. -Detailed investigation of each existing FMM model with respect to its objective (aim), improvements that have been introduced, desired result. -Detailed analysis of the limitations and benefits of each FMM model. - Comprehensive discussion on the main specified limitations in all (21) identified existing models.
Guan and Yuan [38]	-Issue of mislabelled training data in pattern classification and exploring existing methods used to handle mislabelled data.	-Overview of pattern classification.	
Sotoca et al. [39]	-Measures and applicability of data complexity to pattern classification issues.	-Overview of pattern classification.	
Jambhulkar [40]	-FMM for pattern classification. -Explanation for six existing FMM models.	-Overview of FMM pattern classification. -Description of six FMM models.	
Jain and Kolhe [41]	-FMM for pattern classification. -Explanation of seven existing FMM models.	-Overview of FMM pattern classification. -Description of seven FMM models.	
Sayaydeh et al. [42]	-FMM for pattern classification. -A general classifying for the existing FMM variants models based on models using contraction process and model without contraction. -Analysis the usability and applicability of the existing FMM models in different domains and applications using both benchmarks and real data sets.	-Overview of FMM pattern classification. -Descriptions on the existing (14) FMM models and their limitations.	-Detailed and specific overview of the impact of the FMM on pattern classification. -Detailed investigation of each existing FMM model with respect to its objective (aim), improvements that have been introduced, desired result. -Detailed analysis of the limitations and benefits of each FMM model. -Comprehensive discussion on the main specified limitations in all identified existing models covered in our review.

data in pattern classification tasks. An overview on pattern classification and the benefits and limitations of the identified existing methods were included in the review. However, this review did not provide an exhaustive review of FMM as the specific topic.

Sotoca *et al.* [39] conducted a review on the measures and applicability of data complexity to pattern classification issues. Their review explored and identified data complexity measures pertaining to the category of class separability, statistical, overlap, density and geometric measures. Moreover, the prototype selection, feature selection, and meta-analysis of classifiers were identified as three applications in which the measures of data complexity were implemented. However, FMM was not considered as a related model for complexity measures in their review.

A study on FMM-based models for pattern classification was conducted in [40]. The study aimed to provide an overview on six FMM-based variants, namely FMM, GFMM, EFC, FMNC, DCFMN and MLF. Similarly, an examination on for seven FMM-based models, namely FMM, GFMM, EFC, FMCN, GRFMN, DCFMN and MLF, was presented in [41]. However, these two studies did not provide a critical analysis with respect to investigate the capability of the reviewed models in tackling pattern classification problems and the associated limitations.

Recently, Sayaydeh *et al.* [42] conducted a survey on FMM. The survey concentrated on FMM models and their applications. The description was focused on the usage of FMM and fourteen FMM variants for pattern classification. In addition, the review classified the existing models into two categories (with and without contraction), which was based on contraction or otherwise in the learning process. Even though this study usefully emphasized on FMM models and applications as compared with previous review studies, it lacks a detailed analysis on the objectives and improvements introduced by each FMM variant, as well as structural complexity and expansion user defined parameter of each FMM variant.

Our study can be viewed as complementary to the previous review studies [37]–[42]. Comparatively, a detailed analysis on the advantages and disadvantages of each FMM variant covered in our review is presented. Furthermore, the effects of FMM on pattern classification are investigated in detail, in order to highlight the impacts of this model in the pattern classification domain, as highlighted in Section (IV, A). A total of 21 existing FMM models

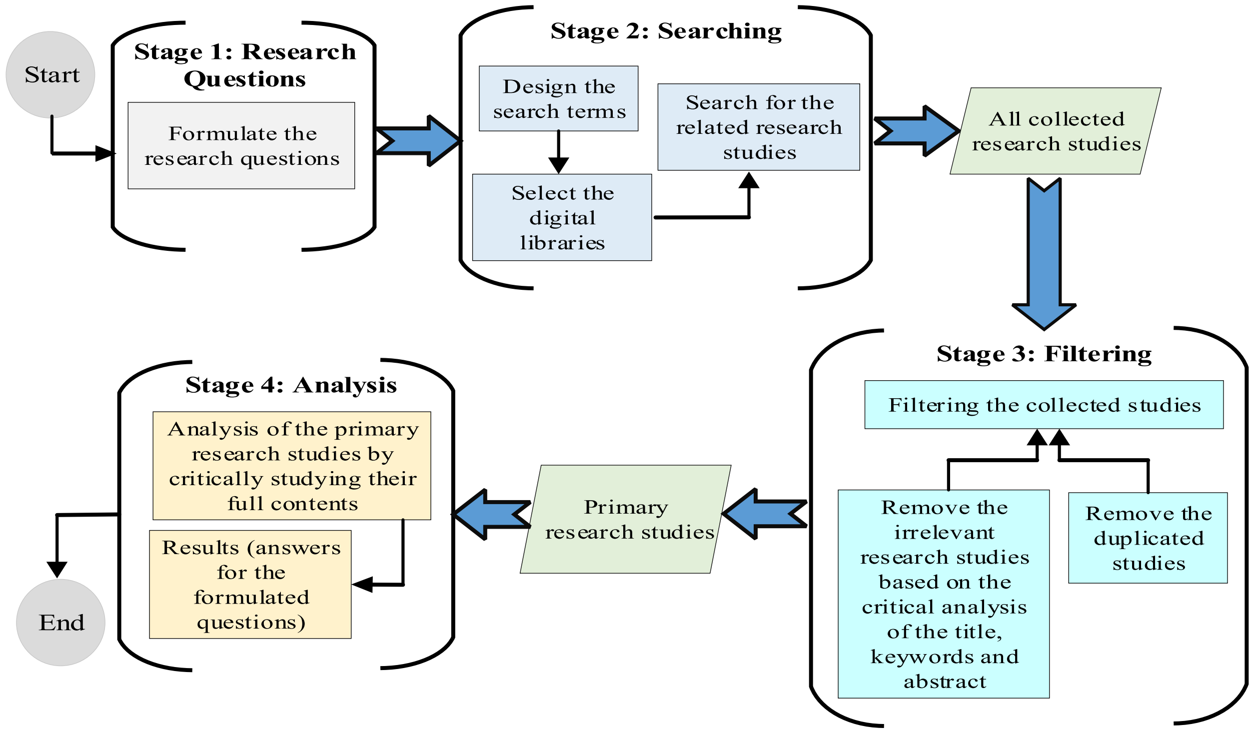

FIGURE 1. Review method.

are identified. The objectives, improvements, and results of each FMM variant are discussed comprehensively, which encompasses structural complexity, user defined parameter, benefits, and limitations. Furthermore, the challenges of FMM-based variants are identified and precisely elaborated, along with further research opportunities to tackle the identified challenges.

III. RESEARCH METHOD

This study is conducted by following a rigorous review protocol that has been designed based on [43], [44]. Figure 1 illustrates the four stages of the review protocol: research questions, searching, filtering, and analysis. The explanation of each stage is provided in Figure 1. In the first stage, the research questions are formulated based on the objectives of this study. The objectives are three-fold: (1) to provide a comprehensive review of FMM and its effects in handling the classification issues, (2) identify and analyse existing FMM variants in terms of their advantages and disadvantages; (3) suggest future research opportunities to address the specified challenges. To achieve these objectives, a list of research questions (denoted as RQ) has been formulated as follows:

- RQ1: What is the significant impact of FMM in the pattern classification?
- RQ2: What are the available FMM variants (derived from the original FMM model) and their objectives, improvements, and desired results?
- RQ3: What are the advantages and disadvantages of each identified FMM variant?
- RQ4: What are the recommended future research to address the specified limitations?

RQ1 aims to investigate the important influence of FMM on pattern classification. The purpose of this investigation is to specify the definite importance of FMM to solve pattern classification problems. This investigation helps provide an understanding of the reasons for using FMM in pattern classification. This objective can be achieved by highlighting the effects (i.e., online learning and handling the overlapped classes' issue) of implementing FMM in addressing pattern classification issues. In RQ2, we intend to specify the available FMM variants that have been proposed based on the original FMM model, and discuss their objectives, improvements that have been introduced, and the achieved results of each variant.

RQ3 is specifically formulated to critically analyse the benefits and limitations of each FMM variant. Such analysis provides a clear view of each variant, which can assist researchers and practitioners in selecting a suitable model to use in classification tasks. RQ4 is articulated to suggest possible future trends for research, which help solve challenges through possible solutions.

Related studies have been extracted in the second stage (searching). This process was performed by launching an online search in certain digital libraries, namely, IEEE Xplore digital library, Google Scholar, Web of Science and Scopus. These digital libraries provide wide-ranging and reliable research studies that cover existing computational intelligence papers relevant to FMM and pattern classification,

TABLE 2. Search terms.

• FMM neural network classification.
• Importance OR/AND impacts OR/AND benefits OR/AND salient properties of FMM neural network.
• FMM neural network OR variants OR models OR techniques OR methods OR frameworks OR approach.
• Limitations OR challenges OR issues OR drawbacks OR disadvantages OR/AND benefits OR advantages of FMM neural network OR variants OR models OR techniques OR methods OR frameworks OR approach.

and can customise a search based on publication year, type, and domain.

To ensure the quality performance of the search process, a list of search terms have been specified based on the following steps [43], [45], [46]:

1. Specifying the main terms based on the respective research questions.
2. Finding the alternative spelling and synonyms of the specified main terms.
3. Validating the search terms in any relevant study.
4. Combining these terms with Boolean operators (OR/AND).

Table 2 presents the list of search terms used in this study. The defined search terms are employed in the titles, abstracts, and keywords in the identified electronic libraries.

The collected papers have been filtered in Stage 3 (filtering) to select the most relevant studies. We included studies that were written in English. In case a study has multiple copies in different versions, we selected the most recent and complete one. The boundary for the publication year of was set from 1992 to 2019 because the original FMM model was introduced in 1992.

Moreover, unpublished or non-peer reviewed publications (e.g. studies published on websites or those that do not have bibliographic details, such as publication date or type) were regarded as grey studies and excluded in this review. Studies in progress were excluded as well. The title, keywords, and abstract were screened to include only studies that focused on FMM models for pattern classification and exclude irrelevant studies. Consequently, 21 key papers were selected as the primary research studies related to FMM for this review. To obtain the results, a critical analysis of the primary research studies has been performed in Stage 4 (analysis) by studying and analysing the full contents of each paper and then extracting information to address the defined research questions.

IV. RESULTS AND DISCUSSION

The primary papers reviewed in this study consisted of: 17 journal papers and 4 conference papers, all published between 1992 and 2019.

A. RQ1 IMPACT OF FMM NEURAL NETWORK ON PATTERN CLASSIFICATION

FMM is a supervised learning model encompassing both artificial neural network (ANN) and fuzzy set theory [23]. It exhibits useful learning properties for addressing pattern classification problems, viz. Learning online in one-pass through the data samples, overcoming overlapped classes, having nonlinear separability, short training time, and both soft and hard decisions.

Catastrophic forgetting is one of the main challenges in online training of neural networks, which is the limitation of standard multilayer perceptron and radial basis function models [47], [48]. This issue is related to the inability of a data-based learning model to recall what it has previously learned when new data samples are provided for incremental learning [44].

Catastrophic forgetting can also be referred to as the stability–plasticity dilemma [34], [49], [50]. This dilemma stipulates how a classifier can be plastic in absorbing new information from incoming data samples incrementally, while remain stable its existing knowledge base from being washed away by new information [51], [52]. It plays a critical role when an ANN is required to learn from data samples with a single-pass, online learning method [52], [53]. Many ANN models have been proposed to tackle the stability–plasticity dilemma. Among them are the Adaptive Resonance Theory (ART) family of networks [49], [54], [55]. On the other hand, Simpson proposed FMM [23], [24] that exhibited the capability of online learning. In other words, FMM can address the stability–plasticity dilemma without conducting a new and complete retraining for existing and new data [23].

Furthermore, the inability to execute nonlinear separation is regarded as one of the key impediments of ANN models. FMM is capable of building a nonlinear decision boundary of any shape to separate data samples from different target classes. On the other hand, the target classes tend to overlap with one another, thereby creating the issue of overlapping classes. FMM can build a nonlinear decision boundary to minimise the degree of misclassification by eliminating overlapping regions of different classes. Moreover, FMM requires a shorter training time than those of other ANN models, such as backpropagation, cascade correlation and Boltzmann machine. This is due to the learning algorithm of FMM that requires only a single-pass without iteration through the data samples.

To yield the output class of a given input sample, FMM provides both soft and hard decisions with respect to the predicted target classes. A soft decision offers a prediction that indicates the degree (between 0 and 1) to which an input pattern fits into the available target classes. By contrast, a hard decision yields a value of either 0 or 1, indicating that the input fits into only a target class.

The aforementioned FMM salient properties have motivated researchers to enhance the performance of FMM in pattern classification. Therefore, numerous FMM variants have been proposed. A comprehensive explanation and analysis of the original FMM model and its variants is provided in the following subsections.

B. RQ2 EXISTING FMM VARIANTS

RQ2 aims to identify and discuss the available FMM variants, and investigating their objectives, improvements that have been introduced, and the achieved results.

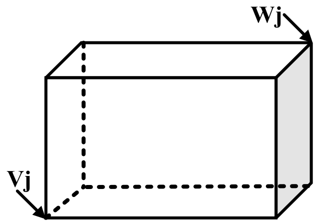

FIGURE 2. 3D hyperbox and its min–max points.

1) ORIGINAL FMM

FMM is constructed using hyperbox fuzzy sets. An n-dimensional hyperbox is defined by two vertices (V and W), i.e. the min–max points. Figure 2 shows a three-dimensional hyperbox. Each hyperbox is encoded by a nodes in the middle layer of FMM for storing knowledge, and is associated with one target class. Each target class can be associated with one or more hyperboxes. All data samples contained inside a hyperbox have a full fuzzy class membership; while the min–max points of each hyperbox are associated with the fuzzy membership function. This function is used to measure the degree to which a pattern (data sample) fits into a hyperbox with respect to the min–max points. The membership is decreased when the distance between a pattern and a hyperbox is increased. The membership, which ranges from 0 to 1, is calculated as follows:

$$B_j(A_h) = \frac{1}{2n}\sum_{i=1}^{n}[max(0, 1 - max(0, \gamma min(1, a_{hi} - w_{ji}))) + max(0, 1 - max(0, \gamma min(1, v_{ji} - a_{hi})))] \quad (1)$$

where B_j (A_h) denotes the membership function, $A_h = (a_{h1}, a_{h2}, a_{h3}, \ldots, a_{hn})\epsilon I^n$ is the A_h input pattern, $\gamma\epsilon[0, 1]$ represents the sensitivity parameter that regulates the decreasing rate of membership as the distance between A_h and B_j increases; and $V_j = (v_{j1}, v_{j2}, \ldots, v_{jn})$ is the minimum point of B_j, $W_j = (w_{j1}, w_{j2}, \ldots, w_{jn})$ is the maximum point of B_j.

Figure 3 shows the FMM structure that comprises three layers. The first layer holds the input patterns (F_A). It comprises input nodes equal in number to the dimensions of the input pattern. The second layer (F_B) contains hyperboxes. Each node in the second layer represents a hyperbox fuzzy set, which is created and adjusted during the learning phase. The connections between F_A, and F_B have weights corresponding to the min–max points of a hyperbox, which are stored in two matrices (V and W).

Each node in the third layer (F_C) encodes a target class. The connections between a hyperbox and class

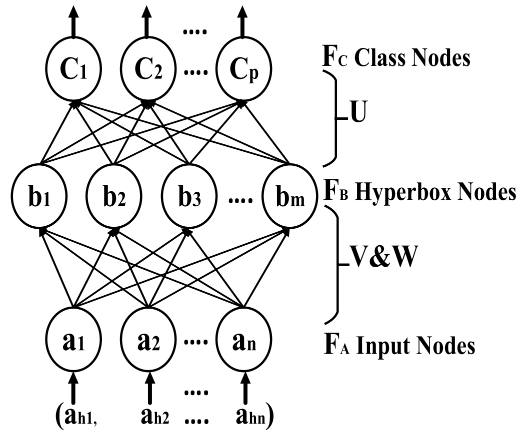

FIGURE 3. Three-layer FMM network.

nodes (F_B, F_C) are binary values stored in matrix U. The output of each class node (F_C) presents the degree to which input pattern A_h fits within target class C_k. The output of the class nodes can be a soft or hard decision. A soft decision measures the degree to which the input pattern belongs to a particular target class. By contrast, a hard decision yields an output of 0 or 1, indicating only one predicted target class for the input pattern.

The creation and adjustment processes of all the hyperboxes occur during the learning phase. Overlapping among hyperboxes of the same class is allowed, but not for different classes. Once an input pattern is presented, a procedure to check whether a hyperbox of the same class exists and whether the current input pattern is included in the identified hyperbox takes place. A winning hyperbox is selected to encode the input pattern based on the highest fuzzy membership value with Equation (1). Then, the expansion, overlap test, and contraction processes ensue. The detailed explanation of these processes is presented as follows.

Hyperbox Expansion: In this process, expansion is implemented to include the input pattern into one of the existing hyperboxes that belong to the same class, provided that the hyperbox size does not exceed the constraint in Equation. (2). The maximum hyperbox size is determined by a user-defined threshold. If none of the existing hyperboxes can be expanded to include the input pattern, a new hyperbox is created with the min-max points set equal to the input pattern.

$$n\theta \geq \sum_{i=1}^{n}(max(w_{ji}, a_{hi}) - min(v_{ji}, a_{hi})) \quad (2)$$

where the hyperbox size, $\theta\epsilon[0, 1]$, is a user-defined threshold, n is the input dimension; v_{ji}, w_{ji}, a_{hi} are the i-th element of the minimum point (V_j), maximum point (W_j) of the j-th hyperbox,, and input pattern (A_h), respectively. The minimum and maximum points of the selected hyperbox are updated based on equations 3 and 4, if the condition in (2) is satisfied.

$$v_{ji}^{new} = min(v_{ji}^{old}, a_{hi}) \quad (3)$$

$$w_{ji}^{new} = max(w_{ji}^{old}, a_{hi}) \quad (4)$$

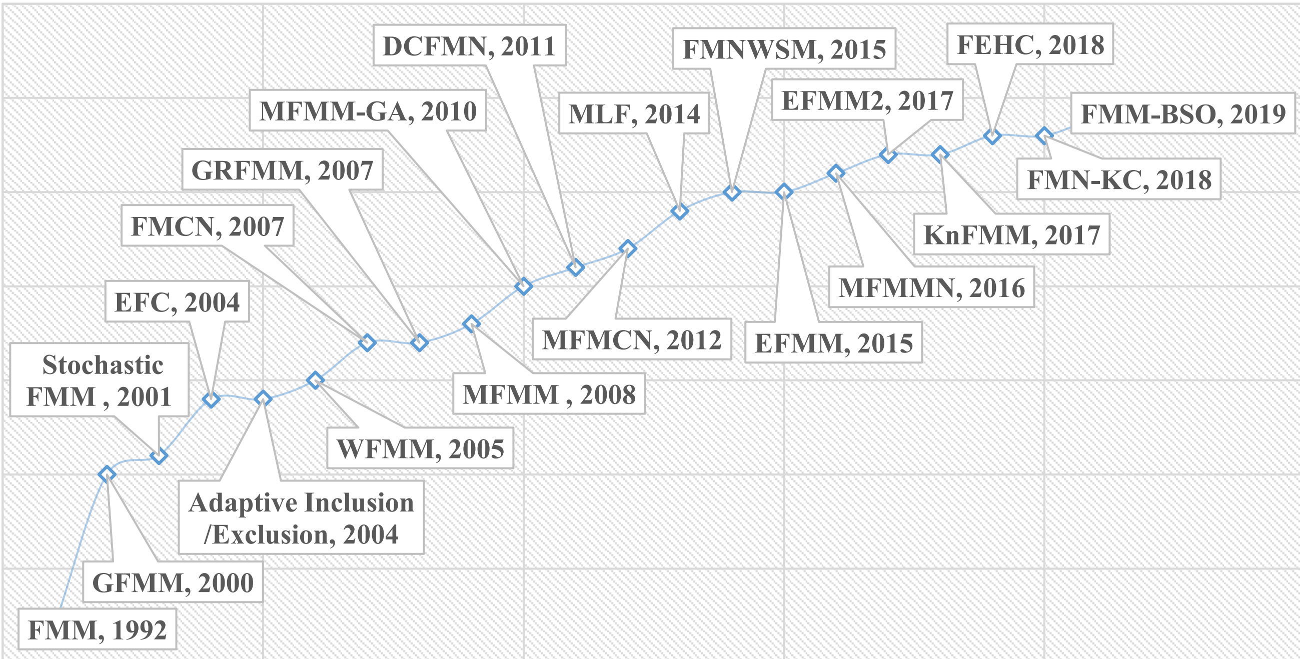

FIGURE 4. Development of FMM and its variants.

Hyperbox Overlap Test: The expansion process can lead to overlapping among hyperboxes of different classes. Consequently, the overlap test is implemented to identify any occurrence of overlapped regions. This test is conducted by checking on a dimension-by-dimension basis between the winning hyperbox with respect to those of different classes. In general, four cases are used to accomplish the overlap test, as follows:

$$\text{Case 1: } v_{ji} < v_{ki} < w_{ji} < w_{ki},\quad \delta^{new} = min(w_{ji} - v_{ki}, \delta^{old}), \tag{5}$$

$$\text{Case 2: } v_{ki} < v_{ji} < w_{ki} < w_{ji},\quad \delta^{new} = min(w_{ki} - v_{ji}, \delta^{old}), \tag{6}$$

$$\text{Case 3: } v_{ji} < v_{ki} < w_{ki} < w_{ji},\quad \delta^{new} = min(min(w_{ki} - v_{ji}, w_{ji} - v_{ki}), \delta^{old}), \tag{7}$$

$$\text{Case 4: } v_{ki} < v_{ji} < w_{ji} < w_{ki},\quad \delta^{new} = min(min(w_{ji} - v_{ki}, w_{ki} - v_{ji}), \delta^{old}). \tag{8}$$

Initially $\delta^{old} = 1$, and a dimension-by-dimension check for each hyperbox is conducted, from the first to the last one. An overlapped area is identified when $\delta^{old} - \delta^{new} < 1$. If no overlap is found, the contraction process is not required.

Hyperbox Contraction: If any of the four cases is detected, the contraction process ensues, with the aim to eliminate the overlapped regions between hyperboxes that belong to different classes. While overlap regions between hyperboxes of the same class are allowed, those between different classes are eliminated by adjusting the minimal overlapped dimension of hyperboxes. The contraction cases are as follows:

$$\text{Case 1: } v_{j\Delta} < v_{k\Delta} < w_{j\Delta} < w_{k\Delta},\quad w_{j\Delta}^{new} = v_{k\Delta}^{new} = \frac{w_{j\Delta}^{old} + v_{k\Delta}^{old}}{2}; \tag{9}$$

$$\text{Case 2: } v_{k\Delta} < v_{j\Delta} < w_{k\Delta} < w_{j\Delta},\quad w_{k\Delta}^{new} = v_{j\Delta}^{new} = \frac{w_{k\Delta}^{old} + v_{j\Delta}^{old}}{2}; \tag{10}$$

$$\text{Case 3a: } v_{j\Delta} < v_{k\Delta} < w_{k\Delta} < w_{j\Delta} \text{ and } (w_{k\Delta} - v_{j\Delta}) < (w_{j\Delta} - v_{k\Delta}),\quad v_{j\Delta}^{new} = w_{k\Delta}^{old}; \tag{11}$$

$$\text{Case 3b: } v_{j\Delta} < v_{k\Delta} < w_{k\Delta} < w_{j\Delta} \text{ and } (w_{k\Delta} - v_{j\Delta}) > (w_{j\Delta} - v_{k\Delta}),\quad w_{j\Delta}^{new} = v_{k\Delta}^{old}; \tag{12}$$

$$\text{Case 4a: } v_{k\Delta} < v_{j\Delta} < w_{j\Delta} < w_{k\Delta} \text{ and } (w_{k\Delta} - v_{j\Delta}) < (w_{j\Delta} - v_{k\Delta}),\quad w_{k\Delta}^{new} = v_{j\Delta}^{old}; \tag{13}$$

$$\text{Case4b: } v_{k\Delta} < v_{j\Delta} < w_{j\Delta} < w_{k\Delta} \text{ and } (w_{k\Delta} - v_{j\Delta}) > (w_{j\Delta} - v_{k\Delta}),\quad v_{k\Delta}^{new} = w_{j\Delta}^{old}. \tag{14}$$

Over the years, many variants have been proposed to improve the performance of FMM. Figure 4 presents the existing FMM variants. A total of 21 key variants are identified from the selected primary research studies. Table 4 in Appendix presents these identified variants and their acronyms. The following subsection discusses each variant in terms of its objectives, improvements introduced, and results achieved, which leads to analyse on the challenges faced by each variant.

2) GENERAL FUZZY MIN-MAX NEURAL NETWORK

General fuzzy min-max neural network (GFMM) appears to be the first variant to improve the performance of the original FMM network by addressing the following issues: the inability to distinguish between ignorance and equal interpretation of membership degrees, the inability to simultaneously address labelled and unlabelled data and interval analysis.

GFMM can simultaneously process labelled and unlabelled input patterns by combining supervised and unsupervised learning into a single algorithm. This feature enables the use of GFMM in three different modes: pure classification, clustering, and hybrid modes (partial supervision). Several changes are introduced to improve efficiency of FMM. The modified fuzzy membership function of GFMM differs from that of FMM, which is a new formulation to compute the membership values. Simultaneously, the sensitive parameter for regulating the maximum hyperbox size can be changed adaptively during the learning phase of GFMM. The input patterns can be fuzzy hyperboxes or crisp points in the pattern space.

Moreover, a change in hyperbox expansion is observed in GFMM, as compared with that in FMM. The GFMM algorithm defines a new constraint, which ensures that the differences between the minimum and maximum points of the individual dimension do not exceed a user-specified limit. GFMM has been compared with FMM in tackling classification and clustering tasks in a single-pass training scenario. Comparatively, GFMM produces fewer hyperboxes and exhibits lower misclassification rates [27]. GFMM uses the same contraction process as supervised learning FMM and assists in achieving the minimal overlapped dimensions of different classes. However, the application of this contraction process results in classification errors for labelled data [31].

3) REINFORCEMENT LEARNING USING STOCHASTIC FMM NEURAL NETWORK

Likas proposed a new pattern classification model in [56], which is called as reinforcement learning using stochastic FMM neural network (Stochastic FMM) to efficiently address reinforcement learning problems by extending the random hyperbox concept of FMM [57]. Unlike FMM, this variant utilises a stochastic automaton idea, instead of an action or class label. The probability vector of the stochastic automaton specifies the opposite class through random selection. The purpose of stochastic automaton is to control the degree of randomness in action selection. The location and boundaries of each hyperbox, along with the probability vector, are adjusted with reinforcement learning in the stochastic FMM variant. Its weakness is related to the utilisation of original FMM overlap test rules, which are unable to identify all overlapping cases, therefore affecting its classification performance [34].

4) AN INCLUSION/EXCLUSION FUZZY HYPERBOX CLASSIFIER

An inclusion/exclusion fuzzy hyperbox classifier (EFC) was developed by Bargiela *et al.* [28]. It introduces changes to the learning algorithm to address the overlap region problem in FMM. These changes are achieved by eliminating the contraction process in the learning algorithm. EFC uses two types of hyperboxes: inclusion (for patterns that belong to the same class) and exclusion (for patterns that belong to different classes). It contributes to solving the overlap region problem in FMM by combining the two types of hyperboxes to represent complex data topologies. The exclusion hyperboxes represent overlapping areas among different classes, with an exclusion node in the hidden layer of EFC.

EFC succeeds in minimising the learning algorithm into two steps (expansion and overlap test) instead of three steps (expansion, overlap test, and contraction) by introducing the exclusion hyperboxes. Although this variant can produce lower misclassification than that of FMM, it still fails to obtain reasonably good classification rates. This outcome is related to the inability of the overlap test cases in original FMM to identify all overlapped regions during the learning phase [34].

5) ADAPTIVE INCLUSION /EXCLUSION FUZZY HYPERBOX CLASSIFIER

The adaptive inclusion /exclusion fuzzy hyperbox classifier (Adaptive Inclusion /Exclusion) was subsequently developed by Bargiela *et al.* [29]. This variant serves as an extension of EFC to improve the classification performance. Similar to EFC, two types of hyperboxes (inclusion and exclusion) are used, while a change in the expansion parameter is introduced. The usage of the overlap test cases of original FMM compromises classification accuracy of this variant [42]. The empirical findings show that this adaptive Inclusion/Exclusion model outperforms EFC.

6) A WEIGHTED FUZZY MIN-MAX NEURAL NETWORK AND ITS APPLICATION TO FEATURE ANALYSIS

A weighted fuzzy min-max neural network and its application to feature analysis (WFMM) was developed by Kim and Yang [58] to introduce feature analysis capabilities to pattern classification. It presents a new type of membership function that considers the weight of each feature in a hyperbox during expansion. The weight indicates the frequency factor of feature values. Consequently, the weight factor effectively reflects the relationship between the feature range and its distribution.

Several changes are introduced to the learning algorithm for hyperbox creation (expansion, contraction, and weight update). A hyperbox can be expanded without the restrictions induced by the original overlap test and contraction processes. Although this variant can utilise the factors of weight and feature distribution in the learning stage, the contraction process used in WFMM is an issue, which leads to classification errors [31].

7) A FUZZY MIN-MAX NEURAL NETWORK CLASSIFIER WITH COMPENSATORY NEURON ARCHITECTURE

A fuzzy min-max neural network classifier with compensatory neuron architecture (FMCN) was introduced in [30] to address issues related to the FMM learning algorithm. It presents a new compensatory neuron architecture to classify samples that fall between hyperboxes that belong to different classes. The compensatory neuron concept is constructed to function in a similar way to that of the reflex system of the human brain in addressing the class overlap issue. In FMCN, the compensation neurons comprise:

- Overlap compensation neurons (OCNs) are used to address the normal overlapped region between two hyperboxes that belong to different classes, in which one hyperbox crosses another from a different class.
- Contentment compensation neurons (CCNs) are used to address the overlapped region between two hyperboxes that belong to different classes, in which one hyperbox is fully or partially inside another from a different class.

During the FMCN learning stage, hyperboxes are created to include class regions. These hyperboxes are stored in the classifying neurons (CLSs) of FMCN. A compensation neuron has two nodes: CCNs and OCNs. Both are added in the hidden layer of FMCN, where CCNs contain only one output and OCNs generate two outputs.

FMCN uses the compensatory nodes (OCNs and CCNs) to address the overlapped regions and improve the performance of original FMM. Comparatively, hyperbox contraction that retains the dimensions of overlapped hyperboxes through their min–max points is eliminated in FMCN. This variant exhibits higher efficiency in handling the class overlap issue than previous models [32]. However, FMCN does not use an appropriate membership function for the compensatory hyperboxes. Therefore, it cannot correctly classify samples that fall within the overlapped regions [21]. In addition, the FMCN structure is complex due to the use of compensatory neurons, which increases the number of nodes in its hidden layer [33].

8) A GENERAL REFLEX FUZZY MIN-MAX NEURAL NETWORK

A general reflex fuzzy min-max neural network (GRFMM) was proposed to improve the performance of GFMM [31]. It combines clustering and classification related to FMM, and applies the concept of the human reflex mechanism. Instead of the contraction process, the reflex mechanism uses compensatory neurons to address overlapping between labelled hyperboxes. Meanwhile, contraction is utilised to eliminate overlapping among unlabelled hyperboxes.

The GRFMM architecture consists of three types of neurons: classifying neurons (CLNs), OCNs and CCNs. CLNs act as the backbone of this variants. They are used to classify data into the target classes A hyperbox is created in CLNs if the existing ones from the same class cannot be expanded further to include the input pattern. The reflex section comprises OCNs and CCNs. Reflex is active whenever a test sample falls within the class overlapped region. As such, OCNs are used to handle an overlapped region in which a hyperbox from one class crosses another from a different class. By contrast, CCNs are utilised to handle an overlapped region in which a hyperbox from one class is fully or partially contained within another of a different class. For each overlapped region generated during the training phase, a node from one of the compensation sections is added in the hidden layer of GRFMM. The CCN and OCN nodes generate one and two outputs, respectively.

The empirical results indicate that GRFMM outperforms GFMM [31]. The use of compensatory neurons increases the number of nodes in the hidden layer of GRFMM, which induces the complexity issue its network structure [33].

9) A MODIFIED FUZZY MIN-MAX NEURAL NETWORK WITH RULE EXTRACTION AND ITS APPLICATION TO FAULT DETECTION AND CLASSIFICATION

Quteishat and Lim proposed a modified fuzzy min-max neural network with rule extraction and its application to fault detection and classification (MFMM) to improve the classification performance of FMM [59]. This variant aims to address the formation issue of a small number of large hyperboxes in original FMM. The membership function and Euclidean distance are used to classify a data sample during the test phase. The membership function is firstly employed, and the hyperboxes with the highest membership values are selected. Then, the Euclidean distance is computed between the centroids of the selected hyperboxes and the test sample. Based on the computed Euclidean distances, the hyperbox with the shortest one from the test sample is selected as the winning hyperbox.

MFMM has salient properties in line with those of FMM with regard to creating hyperboxes. The hyperboxes created by MFMM are fewer than those of FMM. Moreover, the classification rates of MFMM are better than those of FMM [59]. The major issue is the contraction process used in MFMM that leads to higher misclassification results [34].

10) A MODIFIED FUZZY MIN-MAX NEURAL NETWORK WITH GENETIC-ALGORITHM-BASED RULE EXTRACTOR

A modified fuzzy min-max neural network with genetic-algorithm-based rule extractor (MFMM–GA) was proposed to further improve MFMM by implementing several modifications [59]. In this variant, a data set is divided into three subsets: training, prediction, and test. MFMM–GA uses a pruning procedure to prune hyperboxes based on a confidence factor and then compute predictive accuracy using another prediction data set.

Each hyperbox has a confidence factor for identifying hyperboxes that are used frequently. This factor is generally accurate in making predictions. Moreover, the confidence factor aims to identify hyperboxes with a high accuracy rate and that are rarely used. This procedure is performed after the learning phase. The pruning procedure aims to reduce the network size with improved classification performance. After the pruning phase, the open hyperbox generation phase is executed to generate three types of hyperbox: closed, open, and don't care.

An open hyperbox refers to a hyperbox with dimensions that are not specified by its min-max points. By contrast, a hyperbox with dimensions that are specified by its points is denoted as a closed hyperbox. Other hyperboxes are referred to as 'don't care'.

A genetic algorithm (GA) is used to overcome the problem of a large number of antecedents that are associated with the extracted rules. The extracted rules select a set of hyperboxes that yield good test accuracy rates with a small number of features. The GA is also used to extract important features in the rules. Although this variant can produce lower misclassification rates than those of FMM [60], the use of the expansion, overlap test, and contraction process of the original FMM generates misclassification results [31]. It also defies the online learning property of FMM.

11) DATA-CORE-BASED FUZZY MIN-MAX NEURAL NETWORK CLASSIFICATION

Zhang *et al.* introduced data-core-based fuzzy min-max neural network classification (DCFMN) in [32]. It utilises the FMCN compensatory neurons to denote the overlapped region of hyperboxes from different classes. The learning algorithm has three steps: expansion, overlap test, and overlapping neurons (OLNs), if necessary. In contrast with FMCN, only one type of OLNs is required to handle all kinds of overlapped scenarios among hyperboxes. A new methodology for the learning algorithm is used in DCFMN. The overlap test starts after the creation and expansion of all hyperboxes for the training data. As such, expansion allows two hyperboxes from different classes to overlap, which reduces the number of hyperboxes in DCFMN. Two types of neuron class are available in DCFMN, as follows:

- Classifying neurons (CNs) – to classify data patterns.
- OLNs – to handle all types of overlapping scenarios in hyperboxes among different classes.

DCFMN uses two membership functions at the middle layer for the two types of neuron classes. The membership function of CNs uses certain parameters (e.g. geometric centre of a hyperbox, and data core) to consider data characteristics and noise in the data samples.

The membership function of OLNs is used to handle test data located in the overlapped regions among hyperboxes of different classes, in order to determine which class the test data sample belongs to. The number of hyperboxes and misclassification rates of DCFMN are lower than those of previous FMM variants [32]. However, the overlap test rules are insufficient to identify all overlapping cases, thereby compromising its classification performance [34].

12) MODIFIED FUZZY MIN-MAX CLASSIFIER USING COMPENSATORY NEURONS

Davtalab et al. proposed modified fuzzy min-max classifier using compensatory neurons (MFMCN) to solve the overlapped region problem of hyperboxes from different classes by using the FMCN compensatory nodes [61]. In contrast with FMCN, a new methodology is used for the learning algorithm. This variant handles the overlapped regions after the creation and expansion processes of all hyperboxes pertaining to the training data. This change is implemented to reduce time and space complexities. Furthermore, other modifications are introduced during the test stage by defining a new membership function for compensatory neurons.

On the basis of the conducted experiments in [61], MFMCN exhibits the ability to reduce complexity and time as compared with those of FCMN. However, this variant requires evaluation with real-world problems, in order to generalise and corroborate its findings. Moreover, the overlap test rules used are insufficient to discover all overlapping cases, thereby compromising is classification accuracy [34].

13) MULTI-LEVEL FUZZY MIN-MAX NEURAL NETWORK CLASSIFICATION

Multi-level fuzzy min-max neural network classification (MLF) [21] was developed to address the overlapped region problem of FMM by using smaller hyperboxes. This variant uses a multilevel tree structure, which comprises small hyperboxes within different levels of a network to classify data samples. Compared with original FMM, MLF does not use the contraction step to eliminate overlapped regions. Each node in MLF is used as an independent subnet and a separate classifier. Each node in MFL has two segments: hyperbox development (HBS) and overlap box segment (OLS), which are created and adjusted during the learning phase.

Hyperboxes are created in HBS when the existing hyperboxes in that class cannot be expanded to absorb the input pattern. OLS is created after HBS. Each overlapped hyperbox stored in OLS represents an overlapped region in HBS. The overlapped hyperboxes of OLS are called child subnets. The OLS step is implemented after the creation of HBS in the subnet. Therefore, OLS is used to classify data samples in the overlapped region between hyperboxes from different classes in the HBS of the subnet. Each subnet has a G node, which is used to determine the subnet output. The G node output is determined by the outputs of OLS and HBS. The output of each subnet is computed, and then the class with the largest number of G nodes among all other outputs is returned as the network output. The misclassification rate of MLF is lower than those in the previous variants [21]. However, the overlap test rules are insufficient to identify all overlapping cases as highlighted in [34]. This situation can affect classification accuracy of MLF [34].

14) FUZZY MIN–MAX NEURAL NETWORK FOR LEARNING A CLASSIFIER WITH SYMMETRIC MARGIN

Fuzzy min–max neural network for learning a classifier with symmetric margin (FMNWSM) was proposed by Forghani and Yazdi [33] with the aim of improving the time complexity and reducing misclassification rates. To reduce the learning time, FMNWSM avoids the contraction process and adding special nodes to handle the overlapped regions. As such, only the hyperbox expansion process is used during learning. Furthermore, FMNWSM is trained and tested with data samples from identical probability distributions, in order to reduce the misclassification rates.

The empirical results indicate the misclassification ratio of FMNWSM is often better than those of FMM, GFMN, FMCN, and DCFMN. However, FMNWSM inherits the drawback of the expansion process as in original FMM. This limitation generates more overlapped regions, as a result, reducing classification accuracy as indicated in [36] and [42].

15) AN ENHANCED FUZZY MIN–MAX NEURAL NETWORK

An enhanced fuzzy min–max neural network (EFMM) was proposed by Mohammed and Lim [34] to address the issues related to the learning process of FMM, and to enhance the classification performance. This variant points out that the limitations of the FMM learning process that can lead to misclassification. A solution to the limitations is offered by introducing three new heuristic rules: hyperbox expansion, overlap test and contraction rules.

A new constraint rule is used during expansion to minimise the overlapped regions of hyperboxes from different classes. It ensures that the difference between the min-max values for each dimension is not greater than the constraint. In addition, new rules for the hyperbox overlap test have been proposed to cover other possible overlapped regions.

Furthermore, a new hyperbox contraction process has been introduced to eliminate all overlapped dimensions among hyperboxes from different classes. These new heuristic rules make EFMM more efficient in pattern classification than the previous variants. Complexity is highlighted as the primary challenge of EFMM due to the use of the expansion rule. Indeed, the complexity of EFMM is higher than that of original FMM [34].

16) EXTRACTING CLASSIFICATION RULES FROM MODIFIED FUZZY MIN-MAX NEURAL NETWORK FOR DATA WITH MIXED ATTRIBUTES

Extracting classification rules from modified fuzzy min-max neural network for data with mixed attribute (MFMMN) was developed to address the issue related to discrete and continuous data [35]. It introduces several modifications to FMM by providing a new method for computing the membership values and implementing a change in the criteria of hyperbox expansion. The MFMMN overlap test is performed after creating a new hyperbox with the input pattern. To avoid ambiguous membership, 0.001 is set as the minimum distance among the boundaries of hyperboxes after the contraction process. To reduce network complexity, MFMMN uses a pruning strategy to reduce the number of less efficient hyperboxes in the network.

Although MFMMN can achieve good accuracy rates [35], it still inherits a few drawbacks (i.e. expansion, overlap test, contraction) of the original FMM as highlighted in [34], which affects the classification performance.

17) A NEW HYPERBOX SELECTION RULE AND A PRUNING STRATEGY FOR ENHANCED FMM NEURAL NETWORK

A new hyperbox selection rule and a pruning strategy for enhanced fmm neural network (EFMM2) aims to improve classification performance by solving the noise and complexity problems in EFMM [34], [62]. A new hyperbox selection rule is formulated to minimise the creation of an excessive number of small hyperboxes. This k-nearest hyperbox selection rule assists in reducing network complexity. EFMM2 selects a set of k-nearest hyperboxes from the same class label as the winning hyperboxes. All dimensions of the first k-nearest hyperbox are checked against the expansion coefficient. If a violation is detected, the second nearest hyperbox is selected to undergo the same steps. If all k-nearest hyperboxes fail to satisfy the expansion coefficient, a new hyperbox is created to include the input pattern.

Furthermore, EFMM uses a pruning strategy to minimise the number of hyperboxes and to extract rules. This strategy selects a set of hyperboxes created due to noise, outliers, and low accuracy rates. The pruning strategy eliminates the selected hyperboxes from EFMM2. The performance evaluation of EFMM2 indicates its ability to create fewer hyperboxes than those in EFMM [62]. However, EFMM2 uses a contraction process to remove the overlapped regions, which can lead to an increase in the misclassification rate [31].

18) IMPROVING THE FUZZY MIN-MAX NEURAL NETWORK WITH A k-NEAREST HYPERBOX EXPANSION RULE

Mohammed and Lim proposed improving the fuzzy min-max neural network with a k-nearest hyperbox expansion rule (KnFMM) in [62] to address the complexity issue of original FMM. FMM selects only one nearest hyperbox with the highest degree of membership as the winning hyperbox to encode the input pattern. This process can lead to an increase in the number of hyperboxes, which increases FMM complexity.

In KnFMM, a group of k-nearest hyperboxes from the same class is selected. If the first k-nearest hyperbox cannot satisfy the expansion coefficient, the second nearest hyperbox is used to undergo the same steps. If all k-nearest hyperboxes cannot satisfy the expansion coefficient, a new min–max hyperbox is created to include the input pattern. Although KnFMM can create fewer hyperboxes as compared with those in FMM [36], the contraction process can lead to increased misclassification rates [31].

19) IMPROVED DATA CLASSIFICATION USING FUZZY EUCLIDEAN HYPERBOX CLASSIFIER

Improved data classification using fuzzy euclidean hyperbox classifier (FEHC) was proposed by Azad *et al.* [63] with the aim of improving the classification performance of the FMM [23]. In contrast with FMM, a new way to compute the membership value for the hyperboxes based on the Euclidean distance is introduced. Thus, the process of calculating the membership value of each hyperbox is computed with consideration to the centroids of the hyperboxes. Although this variant can produce better results compared to FMM [63], it still inherit a few of the original FMM drawbacks, which can affect the classification performance as highlighted in [34] and [42].

20) OPTIMIZED FUZZY MIN-MAX NEURAL NETWORK: AN EFFICIENT APPROACH FOR SUPERVISED OUTLIER DETECTION

Optimized fuzzy min-max neural network: an efficient approach for supervised outlier detection (FMN-KC) presents a new network architecture to improve the recall time of the FMM [64]. The FMN-KC architecture added a new stage called as knowledge compaction to be executed after the training stage. The compaction stage represents the hyperboxes that are purely created (no overlap created between hyperboxes from different classes) during the training stage. This can helps to improve the recall time without decreasing classification rate. The empirical findings indicate that FMN-KC outperforms FMM [64]. However, FMN-KC inherits the drawbacks related to the expansion, overlap test, and contraction of the original FMM as indicated in [34] and [42], which leads to higher misclassification results.

21) A HYBRID MODEL OF FUZZY MIN–MAX AND BRAIN STORM OPTIMIZATION FOR FEATURE SELECTION AND DATA CLASSIFICATION

Proposed by Pourpanah et al, a hybrid model of fuzzy min–max and brain storm optimization for feature selection and data classification (FMM-BSO) aims to improve classification accuracy and reduce network complexity [65]. It has two phases: (1) learning phase; and (2) feature selection phase. In the first phase, FMM is employed as an incremental learning model to create hyperboxes for storing knowledge extracted from the training samples. BSO is used in the second phase to increase the accuracy rate and reduce the network complexity. BSO selects the most important features and eliminate irrelevant features. Although FMM-BSO can achieve better results as compared with those of FMM [65], it still inherits a few drawbacks related to the expansion, overlap test, and contraction processes of original FMM, as indicated in [42], which affects its classification performance.

C. RQ3: ADVANTAGES AND DISADVANTAGES OF FMM AND ITS VARIANTS

RQ3 aims to examine the advantages and disadvantages of each FMM variant. To answer this question, the identified 21 FMM-based variants are critically analysed. Table 3 presents the advantages and disadvantages of the analysed FMM variants. Although numerous investigations have been conducted, certain limitations that affect FMM classification performance remain. Some studies have concluded that original FMM experiences problems in its learning process [21], [32], [34]. This finding is related to the fact that the FMM learning process (expansion, overlap test, and contraction) causes misclassification. Most variants still exhibit some limitations, which can be summarized as follows.

1) HYPERBOX EXPANSION

This limitation is suffered by most FMM variants, which include FMM, EFC, FMCN, WFMM, MFMM, MFMM–GA, DCFMM, MFMCN, MLF, FMNWSM, MFMMN, M-FMMN, KnFMM and FMM-BSO. During the learning phase, data samples are presented in sequence. When a sample is presented for training, the network attempts to absorb the input pattern in one of the existing hyperboxes, in accordance with the highest degree of membership from the same class. This process is called hyperbox expansion.

The maximum hyperbox size is limited by parameter $0 \leq \theta \leq 1$, which must be pre-specified by the users. When a hyperbox is required to expand to include the input pattern, the expansion coefficient ($n\theta$) is calculated using Equation (2). These variants calculate the total difference between the min-max points of all dimensions of hyperbox expansion and compare the outcome with ($n\theta$). If the condition of the expansion coefficient (($n\theta$) is not satisfied, a new hyperbox is created to encode the input pattern. However, some dimensions of hyperbox expansion violate the expansion coefficient during the expansion process. Therefore, when certain dimensions exceed the expansion coefficient, the expansion process increases the number of overlapped regions among different classes. This adversely affects network performance and contributes towards inaccurate predictions.

2) HYPERBOX OVERLAP TEST

Hyperbox expansion can lead to overlapped regions among hyperboxes that belong to different classes. The overlap test process is implemented to determine whether an overlapped region occurs between the expanded hyperbox and the existing hyperboxes that belong to different classes. Most FMM variants, which include GFMM, stochastic FMM, EFC, adaptive inclusion/exclusion, FMCN, WFMM, MFMM, MFMM–GA, DCFMM, MFMCN, MLF, MFMMN, M-FMMN, KnFMM and FMM-BSO use the original test cases of FMM. These test cases are insufficient to detect all overlapped regions [34]. The inability to identify all overlapped areas reduces the classification rate, and subsequently affects the hyperbox contraction process used, which include FMM, GFMM, stochastic FMM, WFMM, MFMM, MFMM–GA, M-FMMN and KnFMM.

3) HYPERBOX CONTRACTION

Contraction is performed based on the overlap test cases. The contraction process of several existing variants, such as GFMM, WFMM, MFMM, MFMM–GA, EFMM, EFFM2 KnFMM and FMM-BSO are introduced to eliminate overlapped regions among hyperboxes that belong to different classes. However, the contraction process used affects network performance in terms of membership ambiguity among overlapped regions. As such, classification errors are produced during the learning phase. The patterns achieve a full membership grade when they fall within the overlapped boundaries of hyperboxes from different classes. Figure 5 provides an explanation for the issue of the contraction process used in the aforementioned models. The network is trained by using data samples shown in Figure 5(a). Two hyperboxes from different classes are created with an overlapped region, in which one of the hyperboxes crosses

TABLE 3. Advantages and limitations of existing FMM variants.

No	Model	Advantages	Limitations				
			Expansion process (increase overlap regions)	Overlap test (insufficient)	Contraction process (induces misclassification)	Expansion coefficient (user-defined)	Compensatory neurons (network structure complexity)
1.	FMM [23]	Combines ANN and fuzzy set theory in a unified framework to address pattern classification problems; supervises learning and supports online learning	H	H	H	H	N
2.	GFMM [27]	Simultaneously addresses labelled and unlabelled samples; learning algorithm can train GFMM in three modes: classification, clustering and hybrid modes (semi-supervised); combines unsupervised and supervised modes in a single training; defines a new type of membership function for hyperboxes and input samples	H	H	H	H	N
3.	Stochastic FMM [56]	Uses stochastic automaton instead of class or action labels, such as in traditional FMM; all hyperboxes are random and associated with stochastic learning automaton	H	H	H	H	N
4.	EFC [28]	Addresses the overlap region problem with two types of hyperbox: Inclusion, which includes patterns that belong to the same class, and exclusion, which includes patterns that belong to different classes; an exclusion box represents the overlap area in a network, which has nodes in the hidden layer of the network	H	H	N	H	H
5.	Adaptive exclusion inclusion [29]	Addresses overlapped areas by exclusion nodes, and adopts a new expansion parameter	H	H	N	H	H
6.	WFMM [58]	Defines a new type of membership function and modifies the learning algorithm for creating hyperboxes (expansion, contraction and weight updating)	H	H	H	H	N
7.	FMCN [30]	Uses compensatory neurons (OCNs and CCNs), instead of hyperbox contraction, to address the overlap area problem; preserves the dimensions of overlap hyperboxes by maintaining their min–max points; the compensatory neuron concept originates from the reflex system of the human brain, which requires control in difficult situations; defines the membership function for two types of neurons	H	H	N	H	H
8.	GRFMM [31]	Combines clustering and classification related to FMM and applies the concept of the human reflex mechanism to address the overlap problem of labelled hyperboxes; contraction is executed if overlap occurs between unlabelled hyperboxes	H	H	N	H	H
9.	MFMM [59]	Membership function and Euclidean distance are used in the test phase to classify test samples	H	H	H	H	N
10.	MFMM–GA [60]	Network pruning is performed to overcome hyperboxes with low confidence factors; membership function and Euclidean distance are used to predict test samples; GA is adopted for rule extraction; incorporates GA to select a small number of significant rules from a large pool of extracted rules to minimise the number of input features	H	H	H	H	N
11.	DCFMN [32]	Uses compensatory neurons to denote the overlap regions of hyperboxes from different classes; applies three learning algorithms: expansion, overlap test and OLNs (if necessary); adopts a new approach for checking class overlap after creating and expanding hyperboxes for all training data; uses a new membership function to consider data characteristics to suppress the influence of noise	H	H	N	H	H
12.	MFMCN [61]	Addresses overlap regions after creating and expanding hyperboxes for all training data to reduce time and space complexity; uses compensatory neurons instead of a contraction process to address the problem of overlap class	H	H	N	H	H
13.	MLF [21]	Uses a multilevel tree structure method to handle the overlap regions of hyperboxes from different classes; uses each node in MLF as an independent subnet and a separate classifier; learning for each node in the network has two segments (HBS and OLS), which are created and adjusted during the training phase	H	H	N	H	H
14.	FMNWSM [33]	The data is classified with symmetric margin by the use of a non-linear program, and only expansion process used in training phase	H	H	N	H	N
15.	EFMM [34]	Uses a new expansion rule to minimise the overlap regions of hyperboxes from different classes; extends the hyperbox overlap test rules of FMM by discovering other overlap case possibilities; formulates new hyperbox contraction rules	D	D	H	H	N
16.	MFMMN [35]	Processes discrete and continuous attributes; computes modifications of the membership values and the changes in the criterion under which hyperboxes are expanded; ensures a minimum distance of 0.001 to maintain boundaries between hyperboxes after contraction	H	H	D	H	N
17.	EFMM2 [62]	Uses a new hyperbox selection rule to minimise the creation of an excessive number of small hyperboxes and a new pruning strategy to address the network complexity of EFFM	D	D	H	H	N
18.	KnFMM [36]	Uses a new hyperbox selection rule to minimise the creation of an excessive number of small hyperboxes, which helps solve network complexity problems in FMM	H	H	H	H	N
19.	FEHC [63]	Introduces a new membership function for the hyperboxes based on the Euclidean distance	H	H	H	H	N
20.	FMN-KC [64]	Uses a new network architecture to improve the recall time of FMM	H	H	H	H	N
21.	FMM-BSO [65]	Uses BSO to increase the accuracy rate of FMM and reduce network complexity.	H	H	H	H	N

H: Have limitation, D: Do not have limitation, N: None covered

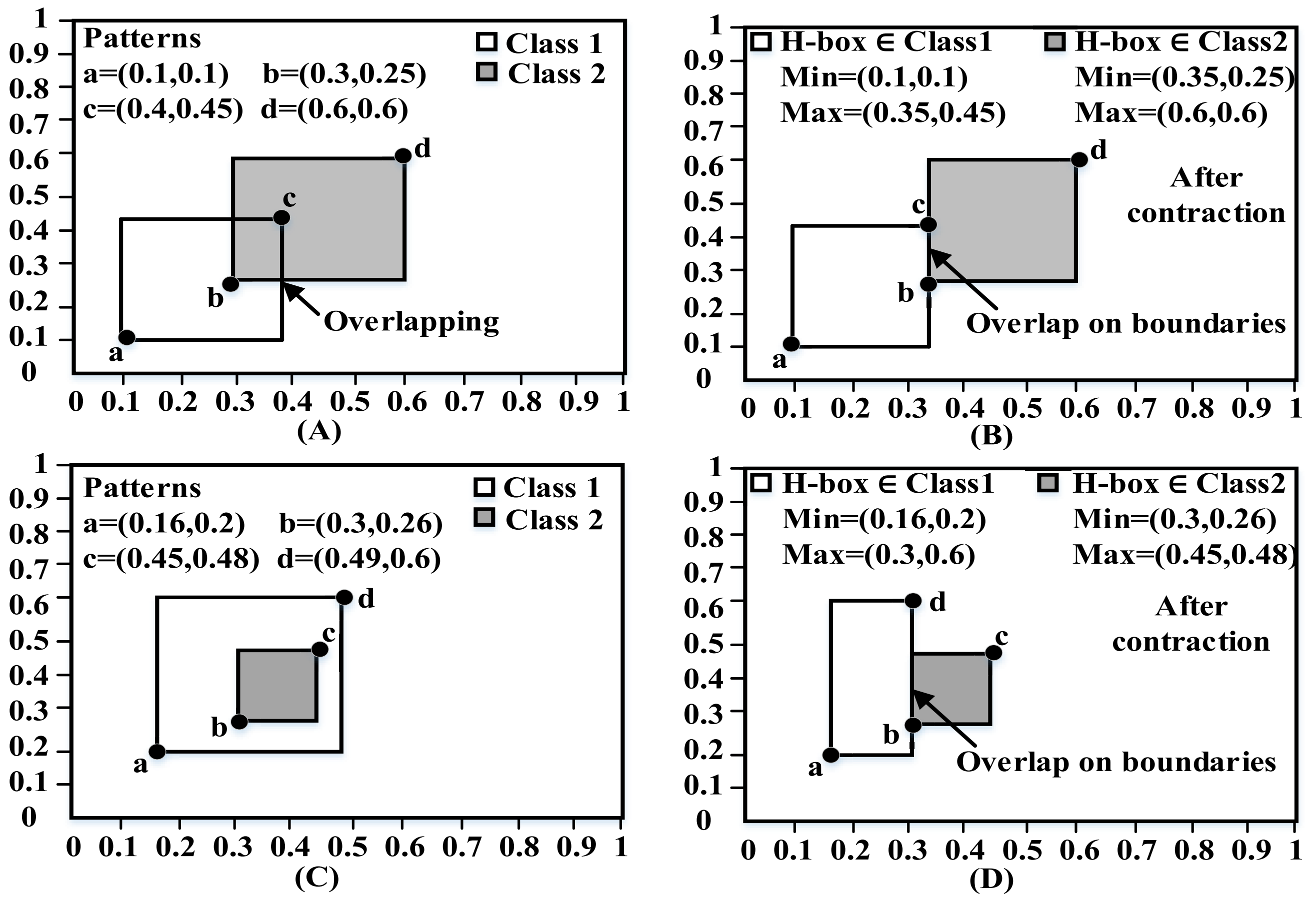

FIGURE 5. Contraction process.

the other. To eliminate the overlapped region, a contraction process is used, as shown in Figure 5(b). Figure 5(b) presents the outcome of the contraction process. Data samples b and c show that the overlapped region still occurs on the boundaries, which can lead to ambiguous membership calculation. Figure. 5(c) presents the hyperbox containment problem, in which a hyperbox of class 2 is fully inside a hyperbox of class 1. A contraction process is used to eliminate the overlapped region in both hyperboxes shown in Figure 5(d). After executing the contraction process as shown in Figure 5(d), the overlapped region between both hyperboxes still exist on the overlapping boundaries, thereby leading to ambiguous membership computation.

4) COMPENSATORY NEURONS

EFC, adaptive inclusion/exclusion, FMCN, DCFMM, and MFMCN use compensatory neurons, instead of hyperbox contraction, to denote the overlapped regions of hyperboxes from different classes. Compensatory neurons have been proposed to solve membership confusion by maintaining boundaries between overlapped hyperboxes to preserve their min–max points. However, the use of compensatory neurons increases the number of nodes in a network's hidden layer, which makes the network structure more complex than those that use the contraction process in original FMM.

5) EXPANSION COEFFICIENT

During the learning phase, whenever a new input pattern is provided for training, FMM and its variants attempt to accommodate the input pattern into one of the existing hyperboxes from the same class through the expansion process. The maximum hyperbox size is limited by a user-predetermined parameter (expansion coefficient). A hyperbox expands to include the input pattern. If its size exceeds the expansion coefficient, a new hyperbox with its min-max points equal to the corresponding the input pattern is created.

The hyperbox size is calculated using Equation. (2). As such, FMM and its variants are highly dependent on the expansion coefficient to perform hyperbox expansion. Such dependence can influence the formation of hyperboxes in the network structure. In addition, the learning algorithms of all FMM variants require users to identify the appropriate learning parameter, which affects the number of hyperboxes created. This poses another challenge that needs to be overcome.

D. RQ4: RECOMMENDED FUTURE RESEARCH OPPORTUNITIES OF FMM

FMM is a common neural network that combines ANN and fuzzy set to tackle pattern classification problems.

Given the number of useful properties of FMM (described in Section IV, A), many variants have been introduced. Based on this study, a clear view of the limitations of the key FMM variants has been identified. The main issues are related to hyperbox expansion, overlap test, and contraction processes of FMM. Accordingly, the performance of FMM variants can be enhanced by addressing the identified limitations. This goal can be achieved by proposing a new model with the following improvements.

Improvement 1: A new learning method to eliminate the expansion coefficient defined by users is required. As can be noticed from the analysis conducted in this study, FMM and its variants rely on the user-defined expansion coefficient to initiate the expansion process for including the input pattern. While increasing the expansion coefficient leads to a smaller network structure with fewer hyperboxes, it can increase the overlap regions between different classes, resulting in a low classification performance. On other hand, decreasing the expansion coefficient leads to a more complex network structure. Clearly, an inappropriate expansion coefficient influences the hyperbox formation process in the network structure, which directly affects the classification performance. Therefore, there is a need to eliminate the user defined coefficient during the learning process to generate accurate hyperboxes decision boundaries and decrease the misclassification rate. In order to overcome this limitation, a new learning method can be suggested, which avoids using the user-defined expansion coefficient, in order to enhance the decision boundaries of hyperboxes during the learning phase. The new learning method can adapt the expansion coefficient and expand the hyperboxes in accordance with the need to accommodate different sizes of classes in the data set. This will further enhance the decision boundary and increase classification accuracy.

Improvement 2: A new method for the overlap test to investigate the possibility of overlapped regions among hyperboxes from different classes and avoid their occurrence is required. Most of the expansion processes in FMM variants aim to increase the possibility of generating overlap regions among hyperbox of different classes. A new method for the overlap test to further improve detection of any overlapped regions can be developed by selecting a group of winning hyperboxes from the same class with high membership values. These winning hyperboxes are sorted in a descending order based on their achieved membership values. Then, all the dimensions of the first winning hyperbox (the highest membership value) are checked with the overlap test. If any of the dimensions of the first winning hyperbox induces overlapping with hyperboxes from different classes, the next winning hyperbox is selected to undergo the same procedure. If all the selected winning hyperboxes induce overlapping with hyperboxes from different classes, a new hyperbox is created to include the input pattern. This method can reduce overlapping among hyperboxes from different classes during the learning phase.

Improvement 3: A new contraction process to avoid membership ambiguity of overlapped regions is required. The majority of FMM variants adopt the original FMM contraction process to remove the overlap regions among hyperboxes of different classes during the training stage. The process suffers from two limitations: (i) inability to handle the boundary overlaps among hyperboxes of different classes (as described in Section C, 3); and (ii) the data distortion problem, which refers to loss of part of the contracted hyperbox information during the contraction procedure. Both limitations affect the performance of FMM variants negatively with respect to membership ambiguity, which leads to an increase in the misclassification rate. To solve the limitations, a new contraction process to avoid the membership ambiguity of overlapped regions, and to overcome the data distortion problem is necessary. In this aspect, both limitations can be eliminated by re-defining the min-max points based on the original information (created directly from an input sample) as well as subsequent information (created by previous expansion or contraction processes).

V. DISCUSSION

This review focuses on FMM research and provides a synopsis of its importance in pattern classification. Pattern classification is one of the primary components in the design and development of computerised intelligent systems. FMM offers a useful classification network with an incremental learning paradigm that requires a single-pass learning procedure through the data samples. It is a unique neuro-fuzzy model with some remarkable properties, i.e., online learning, nonlinear separability, non-overlapping classes, soft and hard classification decisions, and fast training time. In terms of online learning, FMM can learn new data samples without losing information extracted from previous data samples. In the case of nonlinear separability, FMM is able to build an arbitrary nonlinear decision boundary to an arbitrary degree of accuracy for separating data samples from different classes. Regarding the classification decision, unlike other classifiers, FMM provides both soft and hard decisions with respect to the target classes. Concerning the training time, FMM conducts it learning process with single pass through the data samples, as compared with other models (e.g. backpropagation, cascade correlation, and Boltzmann machine) that require repetitive training iterations.

Despite the salient properties of FMM, it has some issues related to the learning process, which include inability of addressing all possible overlap cases, increase in the overlapped regions among hyperboxes of different classes, and issues related to the contraction process. To eliminate these shortcomings, many variants have been proposed. A total of 21 FMM-based variants have been identified in this review. The identified variants have been carefully investigated concerning the aims, improvements that have been introduced, and the results achieved. These variants have also been critically analysed with respect to their advantages and disadvantages. Many improvements, which include the learning process (expansion, overlap rules, contraction rules), network architecture, and membership function, have been introduced. Several variants have proposed different expansion equations to reduce the overlapped regions of hyperboxes among different classes, e.g. GFMM, EFMM, and EFMM2. EFMM and EFMM2 aim to improve the FMM learning processes (expansion, overlap test, and contraction process). Others, such as EFC, FMCN, GRFMM, DCFMN, and MFMCN, examine the FMM network structure by replacing the contraction process with compensatory neurons in hidden layers to represent the overlap regions of hyperboxes from different classes, as well as introducing new membership functions for the compensatory neurons.

Although FMM variants possess many robust features for tackling pattern classification problems, none of these variants has completely addressed all the FMM-related issues. Our analysis of the existing FMM variants indicates that issues related to hyperbox expansion, overlap test, contraction process, compensatory neurons, and expansion coefficient constitutes the main challenges. Most of the overlap tests processes used in existing FMM variants affect the classification rates, due to inability to detect all possible overlap cases. The compensatory neurons used in some variants increase the number of nodes in the hidden layer, leading to a high degree of network complexity. Furthermore, these models still inherit the limitations related to the original expansion process and overlap test.

Meanwhile, FMM variants that perform a contraction process in their learning stage suffer from membership ambiguity among the overlapped regions and data distortion problems. These affect the network structure and classification performance. The ambiguity problem occurs when a pattern falls in the overlapped boundary region of hyperboxes, whereby it obtains a full membership grade. The analysis reveals that the hyperbox expansion process used by the majority of FMM variants entails the possibility of overlapped regions occurring among hyperboxes from different classes, which result in inaccurate predictions. Importantly, the learning algorithms of all FMM variants require user intervention to specify the expansion coefficient for the learning phase. This user-defined expansion coefficient affects the hyperbox formation process, which in turn causes the learning algorithm susceptible to misclassification.

VI. LIMITATIONS OF THE STUDY

A rigorous search strategy has been conducted to ensure the retrieval of the most relevant papers. Four digital libraries, along with a defined list of search terms, have been used to extract the papers. However, it is not guaranteed that all related papers are included as non-English papers, website articles and other types of articles that are considered as grey studies are ignored. This exclusion is considered as a limitation of this review. In addition, this study has not considered other types of machine learning models, e.g. reinforcement and unsupervised learning, as they are deemed to be unrelated to the supervised learning paradigm of FMM models.

TABLE 4. FMM Variants and their Acronyms.

Variant Name	Acronym
Fuzzy Min–Max Neural Network	FMM
General Fuzzy Min-Max Neural Network	GFMM
Reinforcement Learning using Stochastic FMM Neural Network	Stochastic FMM
An Inclusion/Exclusion Fuzzy Hyperbox Classifier	EFC
Adaptive Inclusion /Exclusion Fuzzy Hyperbox Classifier	Adaptive Inclusion /Exclusion
A Weighted Fuzzy Min-Max Neural Network and Its Application to Feature Analysis	WFMM
A Fuzzy Min-Max Neural Network Classifier with Compensatory Neuron Architecture	FMCN
A General Reflex Fuzzy Min-Max Neural Network	GRFMM
A Modified Fuzzy Min-Max Neural Network with Rule Extraction and its Application to Fault Detection and Classification	MFMM
A Modified Fuzzy Min-Max Neural Network with Genetic-Algorithm-Based Rule Extractor	MFMM–GA
Data-Core-Based Fuzzy Min-Max Neural Network Classification	DCFMN
Modified Fuzzy Min-Max Classifier Using Compensatory Neurons	MFMCN
Multi-Level Fuzzy Min-Max Neural Network Classification	MLF
Fuzzy Min–Max Neural Network for Learning a Classifier with Symmetric Margin	FMNWSM
An Enhanced Fuzzy Min–Max Neural Network	EFMM
Extracting Classification Rules from Modified Fuzzy Min-Max Neural Network for Data with Mixed Attributes	MFMMN
A New Hyperbox Selection Rule and a Pruning Strategy for Enhanced FMM Neural Network	EFMM2
Improving the Fuzzy Min-Max Neural Network with a K-nearest Hyperbox Expansion Rule	KnFMM
Improved Data Classification using Fuzzy Euclidean Hyperbox Classifier	FEHC
Optimized Fuzzy Min-Max Neural Network: An Efficient Approach for Supervised Outlier Detection	FMN-KC
A hybrid Model of Fuzzy Min–Max and Brain Storm Optimization for Feature Selection and Data Classification	FMN-BSO

An extensive filtering process has been adopted (as described in Section III) to identify the most relevant papers that can provide answers to the research questions. However, it is not guaranteed that the filtering process can adequately accomplish this objective. Another limitation of this review is related to the limited digital libraries covered in the search process. Although the selected databases are reliable, the possibilities of missing relevant papers in other sources exist, therefore another limitation of this review.

VII. CONCLUSION

The main contribution of this review is the provision of valuable insights into FMM and its key variants to support researchers and practitioners by understanding their respective advantages and disadvantages. In particular, this review covers FMM and its impact on pattern classification, the key FMM variants with analysis on their objectives, presented improvements, and achieved results. The future trends to address the identified challenges of FMM and its variants are also highlighted. A review methodology has been formulated to achieve the defined objectives of this study. The key related papers have been collected by performing online searching in four digital libraries using specified search terms. The extracted papers have been screened, in which the most relevant ones have been included based on an extensive filtering process. These selected papers have been critically analysed and studied to yield the answers to the defined research questions in this review.

The results of this review indicate that, given the abilities to present nonlinear separability, non-overlapping classes, soft and hard classification decisions and fast training time, FMM exerts a significant impact on pattern classification. While many FMM variants have been proposed to improve the classification performance, a number of limitations, which include hyperbox expansion, overlap test, and contraction, have been identified in this study. Further research opportunities to further enhance FMM have been highlighted, with the ultimate aim to design and develop a new model that can present solutions for the identified limitations.

APPENDIX

See Table 4.

REFERENCES

[1] L. N. Castro, L. N. De Castro, and J. Timmis, *Artificial Immune Systems: A New Computational Intelligence Approach*. London, U.K.: Springer, 2002.

[2] J. J. Wang *et al.*, "Predicting house price with a memristor-based artificial neural network," *IEEE Access*, vol. 6, pp. 16523–16528, 2018.

[3] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral Brain Sciences*, vol. 40, p. e253, Nov. 2017.

[4] J. J. Hopfield, "Artificial neural networks," *IEEE Circuits Devices Mag.*, vol. 4, no. 5, pp. 3–10, Sep. 1988.

[5] S. Walczak, "Artificial neural networks," in *Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction*. Hershey, PA, USA: IGI Global, 2019, pp. 40–53.

[6] B. Coppin, *Artificial Intelligence Illuminated*. Boston, MA, USA: Jones & Bartlett, 2004.

[7] J. Hawkins and S. Blakeslee, *On Intelligence: How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines*. New York, NY, USA: Macmillan, 2007.

[8] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence/Book and Disk*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1992.

[9] L. A. Zadeh, "Fuzzy logic, neural networks, and soft computing," *Commun. ACM*, vol. 37, pp. 77–84, 1994.

[10] P. K. Simpson, "Fuzzy min-max classification with neural networks," in *Proc. IEEE Conf. Neural Netw. Ocean Eng.*, Aug. 1991, pp. 291–300.

[11] D. Dubois and H. Prade, *Fundamentals of Fuzzy Sets*. Dordrecht, The Netherlands: Kluwer, 2012.

[12] O. Castillo, P. Melin, and J. Castro, "Computational intelligence software for interval type-2 fuzzy logic," *Comput. Appl. Eng. Educ.*, vol. 21, no. 4, pp. 737–747, 2013.

[13] Z. Yu and M. Lee, "Real-time human action classification using a dynamic neural model," *Neural Netw.*, vol. 69, pp. 29–43, Sep. 2015.

[14] D. Graupe, *Principles of Artificial Neural Networks*. Singapore: World Sci., 2013.

[15] B. Yegnanarayana, *Artificial Neural Networks*. New Delhi, India: PHI Learning Pvt. Ltd., 2009.

[16] S.-J. Lee, T. Chen, L. Yu, and C.-H. Lai, "Image classification based on the boost convolutional neural network," *IEEE Access*, vol. 6, pp. 12755–12768, 2018.

[17] L. V. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1994.

[18] E. Kayacan, E. Kayacan, and M. A. Khanesar, "Identification of nonlinear dynamic systems using type-2 fuzzy neural networks—A novel learning algorithm and a comparative study," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1716–1724, Mar. 2015.

[19] P. K. Simpson, *Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations*. New York, NY, USA: Pergamon, 1990.

[20] A. Baraldi, P. Blonda, and A. Petrosino, "Fuzzy neural networks for pattern recognition," in *Neural Nets WIRN VIETRI-97*, London, U.K.: Springer, 1998, pp. 35–83.

[21] R. Davtalab, M. H. Dezfoulian, and M. Mansoorizadeh, "Multi-level fuzzy min-max neural network classifier," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 470–482, Mar. 2014.

[22] R. Jafari, W. Yu, and X. Li, "Fuzzy differential equations for nonlinear system modeling with Bernstein neural networks," *IEEE Access*, vol. 4, pp. 9428–9436, 2016.

[23] P. K. Simpson, "Fuzzy min-max neural networks. I. Classification," *IEEE Trans. Neural Netw.*, vol. 3, no. 5, pp. 776–786, Sep. 1992.

[24] P. K. Simpson, "Fuzzy min-max neural networks—Part 2: Clustering," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 1, p. 32, Feb. 1993.

[25] E. Alhroob and N. A. Ghani, "Fuzzy min-max classifier based on new membership function for pattern classification: A conceptual solution," in *Proc. 8th IEEE Int. Conf. Control Syst., Comput. Eng. (ICCSCE)*, Nov. 2018, pp. 131–135.

[26] M. Seera and C. P. Lim, "Online motor fault detection and diagnosis using a hybrid FMM-CART model," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 4, pp. 806–812, Apr. 2014.

[27] B. Gabrys and A. Bargiela, "General fuzzy min-max neural network for clustering and classification," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 769–783, May 2000.

[28] A. Bargiela, W. Pedrycz, and M. Tanaka, "An inclusion/exclusion fuzzy hyperbox classifier," *Int. J. Knowl.-Based Intell. Eng. Syst.*, vol. 8, no. 2, pp. 91–98, Apr. 2004.

[29] A. Bargiela, M. Tanaka, G. Castellano, and A. Fanelli, "Adaptation of exclusion / inclusion hyperboxes for classification of complex data," in *Proc. 5th Asia–Pacific Ind. Eng. Manage. Syst. Conf.*, Gold Coast, QLD, Australia, 2004, pp. 1–12.

[30] A. V. Nandedkar and P. K. Biswas, "A fuzzy min-max neural network classifier with compensatory neuron architecture," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 42–54, Jan. 2007.

[31] A. V. Nandedkar and P. K. Biswas, "A general reflex fuzzy min-max neural network," *Eng. Lett.*, vol. 14, no. 1, pp. 195–205, 2007.

[32] H. Zhang, J. Liu, D. Ma, and Z. Wang, "Data-core-based fuzzy min–max neural network for pattern classification," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2339–2352, Dec. 2011.

[33] Y. Forghani and H. S. Yazdi, "Fuzzy min–max neural network for learning a classifier with symmetric margin," *Neural Process. Lett.*, vol. 42, no. 2, pp. 317–353, Oct. 2015.

[34] M. F. Mohammed and C. P. Lim, "An enhanced fuzzy min–max neural network for pattern classificationc," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 3, pp. 417–429, Mar. 2015.

[35] S. Shinde and U. Kulkarni, "Extracting classification rules from modified fuzzy min–max neural network for data with mixed attributes," *Appl. Soft Comput.*, vol. 40, pp. 364–378, Mar. 2016.

[36] M. F. Mohammed and C. P. Lim, "Improving the fuzzy Min-Max neural network with a K-nearest hyperbox expansion rule for pattern classification," *Appl. Soft Comput.*, vol. 52, pp. 135–145, Mar. 2017.

[37] P. J. García-Laencina, J.-L. Sancho-Gómez, and A. R. Figueiras-Vidal, "Pattern classification with missing data: A review," *Neural Comput. Appl.*, vol. 19, no. 2, pp. 263–282, 2010.

[38] D. Guan and W. Yuan, "A survey of mislabeled training data detection techniques for pattern classification," *IETE Tech. Rev.*, vol. 30, no. 6, pp. 524–530, Nov. 2013.

[39] J. Sotoca, J. S. Sánchez, and R. A. Mollineda, "A review of data complexity measures and their applicability to pattern classification problems," *Actas del III Taller Nacional de Mineria de Datos y Aprendizaje*, pp. 77–83, Jan. 2005. [Online]. Available: https://scholar.google.com.my/scholar?hl=en&as_sdt=0%2C5&q=A+review+of+data+complexity+measures+and+their+applicability+to+pattern+classification+problems&btnG=

[40] R. K. Jambhulkar, "A review on pattern classification using multilevel and other fuzzy min max neural network classifier," *Int. J. Sci. Res.*, vol. 3, no. 12, pp. 898–900, Dec. 2014.
[41] B. Jain and V. Kolhe, "Survey on fuzzy min-max neural networkclassification," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 12, pp. 30–34, Dec. 2015.
[42] O. N. Al Sayaydeh, M. F. Mohammed, and C. P. Lim, "Survey of fuzzy min–max neural network for pattern classification variants and applications," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 4, pp. 635–645, Apr. 2019.
[43] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quart.*, vol. 26, no. 2, pp. 13–23, Jun. 2002.
[44] M. M. Ahmed and N. A. M. Isa, "Knowledge base to fuzzy information granule: A review from the interpretability-accuracy perspective," *Appl. Soft Comput.*, vol. 54, pp. 121–140, May 2017.
[45] F. Hujainah, R. B. A. Bakar, M. A. Abdulgabber, and K. Z. Zamli, "Software requirements prioritisation: A systematic literature review on significance, stakeholders, techniques and challenges," *IEEE Access*, vol. 6, pp. 71497–71523, 2018.
[46] F. Hujainah, R. B. A. Bakar, B. Al-Haimi, and M. A. Abdulgabber, "Stakeholder quantification and prioritisation research: A systematic literature review," *Inf. Softw. Technol.*, vol. 102, pp. 85–99, Oct. 2018.
[47] M. McCloskey and N. J. Cohen, "Catastrophic interference in connectionist networks: The sequential learning problem," *Psychol. Learn. Motivat.*, vol. 24, pp. 109–165, Dec. 1989.
[48] R. Ratcliff, "Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions," *Psychol. Rev.*, vol. 97, no. 2, pp. 285–308, Apr. 1990.
[49] S. Grossberg, "Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors," *Biol. Cybern.*, vol. 23, no. 3, pp. 121–134, Sep. 1976.
[50] G. A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Comput. Vis. Graph. Image Process.*, vol. 37, no. 1, pp. 54–115, Jan. 1987.
[51] R. Polikar, L. Upda, S. S. Upda, and V. Honavar, "Learn++: An incremental learning algorithm for supervised neural networks," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 31, no. 4, pp. 497–508, Nov. 2001.
[52] S. Grossberg, "Adaptive resonance theory: How a brain learns to consciously attend, learn, and recognize a changing world," *Neural Netw.*, vol. 37, pp. 1–47, Jan. 2013.
[53] D. Leite, P. Costa, and F. Gomide, "Evolving granular neural networks from fuzzy data streams," *Neural Netw.*, vol. 38, pp. 1–16, Feb. 2013.
[54] S. Grossberg, "Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, illusions," *Biol. Cybern.*, vol. 23, no. 4, pp. 187–202, Dec. 1976.
[55] E. Parrado-Hernández, E. Gómez-Sánchez, and Y. A. Dimitriadis, "Study of distributed learning as a solution to category proliferation in fuzzy ARTMAP based neural systems," *Neural Netw.*, vol. 16, no. 7, pp. 1039–1057, Sep. 2003.
[56] A. Likas, "Reinforcement learning using the stochastic fuzzy min–max neural network," *Neural Process. Lett.*, vol. 13, no. 3, pp. 213–220, Jun. 2001.
[57] A. Likas and K. Blekas, "A reinforcement learning approach based on the fuzzy min-max neural network," *Neural Process. Lett.*, vol. 4, no. 3, pp. 167–172, Dec. 1996.
[58] H.-J. Kim and H.-S. Yang, "A weighted fuzzy min-max neural network and its application to feature analysis," in *Advances in Natural Computation*, vol. 3612, L. Wang, K. Chen, and Y. S. Ong, Eds. Berlin, Germany: Springer-Verlag, 2005, pp. 1178–1181.
[59] A. Quteishat and C. P. Lim, "A modified fuzzy min–max neural network with rule extraction and its application to fault detection and classification," *Appl. Soft Comput.*, vol. 8, no. 2, pp. 985–995, Mar. 2008.
[60] A. Quteishat, C. P. Lim, and K. S. Tan, "A Modified fuzzy min–max neural network with a genetic-algorithm-based rule extractor for pattern classification," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 40, no. 3, pp. 641–650, May 2010.
[61] R. Davtalab, M. Parchami, M. H. Dezfoulian, M. Mansourizade, and B. Akhtar, "M-FMCN: Modified fuzzy min-max classifier using compensatory neurons," in *Proc. 11th WSEAS Int. Conf. Artif. Intell.*, Feb. 2012, pp. 77–82.
[62] M. F. Mohammed and C. P. Lim, "A new hyperbox selection rule and a pruning strategy for the enhanced fuzzy min–max neural network," *Neural Netw.*, vol. 86, pp. 69–79, Feb. 2017.
[63] C. Azad, A. K. Mehta, and V. K. Jha, "Improved data classification using fuzzy euclidean hyperbox classifier," in *Proc. Int. Conf. Smart Comput. Electron. Enterprise*, Jul. 2018, pp. 1–6.
[64] N. Upasani and H. Om, "Optimized Fuzzy min-max neural network: An efficient approach for supervised outlier detection," *Neural Netw. World*, vol. 28, no. 4, pp. 285–303, 2018.
[65] F. Pourpanah, C. P. Lim, X. Wang, C. J. Tan, M. Seera, and Y. Shi, "A hybrid model of fuzzy min–max and brain storm optimization for feature selection and data classification," *Neurocomputing*, vol. 333, pp. 440–451, Mar. 2019.

ESSAM ALHROOB received the B.Sc. degree in computer information systems from the Al-Zaytoonah University of Jordan, Amman, Jordan, in 2010, and the M.Sc. degree in software engineering from Limkokwing University Malaysia, Cyberjaya, Malaysia, in 2013. He is currently pursuing the Ph.D. degree with the Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang (UMP), Kuantan, Malaysia. He is also a member of a research group that has a collaboration between UMP and the Baoji University of Arts and Sciences. His research interests include pattern classification and fuzzy neural networks.

MOHAMMED FALAH MOHAMMED received the B.Eng. degree in computer engineering from the Technical College of Mosul, Mosul, Iraq, in 2006, and the M.Sc. degree in wireless and mobile system and the Ph.D. degree in computational intelligence from the School of Electrical and Electronic Engineering, University of Science Malaysia, Nibong Tebal, Malaysia, in 2010 and 2014, respectively. His current research interests include fuzzy neural networks, multiagent systems, and pattern classification.

CHEE PENG LIM received the Ph.D. degree in intelligent systems from The University of Sheffield, Sheffield, U.K., in 1997. He is currently an Associate Professor with the Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, VIC, Australia. He has authored more than 350 technical papers in journals, conference proceedings, and books, and edited three books and 12 special issues in journals. His research interests include computational intelligence-based systems for data analytics, condition monitoring, optimization, and decision support.

Dr. Lim received seven best paper awards. He served on the Editorial Board for five international journals.

HAI TAO received the B.Sc. degree from the Department of Computer and Information Science, Northwest University of Nationalities, in 2004, the M.S. degree from the School of Mathematics and Statistics, Lanzhou University, in 2009, and the Ph.D. degree from the Faculty of Computer System and Software Engineering, University Malaysia Pahang, in 2012. He is currently an Associate Professor with the Baoji University of Arts and Sciences. His current research interests include machine learning, the Internet of Things, and optimization computation.