Breast Cancer Diagnosis Using K-Means Type-2 Fuzzy Neural Network

Tien-Loc Le
[a] Department of Electrical Engineering
Yuan Ze University
Taoyuan, Taiwan
[b] Department of Electrical Electronic and Mechanical Engineering
Lac Hong University
Bien Hoa, Vietnam
tienloc@saturn.yzu.edu.tw

Tuan-Tu Huynh
[a] Department of Electrical Engineering
Yuan Ze University
Taoyuan, Taiwan
[b] Department of Electrical Electronic and Mechanical Engineering
Lac Hong University
Bien Hoa, Vietnam
huynhtuantu@gmail.com

Chih-Min Lin*, *Fellow, IEEE*
Department of Electrical Engineering
Yuan Ze University
Taoyuan, Taiwan
cml@saturn.yzu.edu.tw

Fei-Chao
Department of Cognitive Science
Xiamen University
Fujian, China.
fchao@xmu.edu.cn

***Abstract*— This paper aims to design a classifier using the K-means clustering algorithm and the interval type-2 fuzzy neural network (IT2FNN). Firstly, the K-means clustering algorithm will classify the training data into *k* groups, according to its characteristics. After that, the IT2FNN will train the *k* classifiers' structure with these data. The testing data will be also determined that they will belong to which classifier. With this parallel structure, the performance of the proposed classifier is competitive with some state-of-the-art techniques. The parameter adaptive laws of the network are derived by using the steepest descent gradient approach. The convergence and stability of the proposed algorithm are guaranteed using the Lyapunov function. The system performance is evaluated by the breast cancer datasets of the University of California at Irvine (UCI). Comparison with other classifiers is also conducted. The experimental results have shown the effectiveness of the proposed method.**

Keywords— classification problems, interval type-2 fuzzy logic system, K-means clustering, breast cancer diagnosis

I. Introduction

The K-means classification algorithm is one of the unsupervised learning algorithms that is used to solve the clustering problems. It was introduced by Steinhaus in 1956 [1] and had been modified by MacQueen in 1967 [2]. The goal of data clustering is to divide data into groups or clusters such that objects in the same cluster have the similar characteristics or features. In the past years, K-means clustering has been used widely by its advantages such as: simple in computation, easy to understand and implements [3]. In 2016 Kumari et al. proposed K-mean clustering for anomaly detection in network traffic [4]. In 2017, Shanker et al. introduced K-mean clustering and hierarchical centroid shape descriptor for segmentation of tumor and edema [5]. However, the K-means algorithm also has several disadvantages such as: sensitive to noisy data and outliers, request to specify the number of the cluster in advanced [6, 7].

The type-2 fuzzy logic system (T2FLS) was introduced by Zadeh in 1975 [8], which can be known as an extension of the type-1 fuzzy logic system (T1FLS). In the past years, there have many studies indicated that the T2FLS is better to deal with uncertainties than T1FLS [9-11]. The interval type-2 fuzzy logic system (IT2FLS) with type-2 membership functions (T2MFs) was proposed by Liang and Mendel in 2000 [12], which was supposed that it can reduce the computation load in T2FLS. In recent year, the IT2FNN has been widely applied to solve the nonlinear problems [13-17]. In 2015, Kim and Chwa proposed the wheeled mobile robots obstacle avoidance method using an IT2FNN [13]. In 2017, Eyoh et al. introduced the interval type-2 intuitionistic fuzzy logic for regression problem [17].

Nowadays, along with the rapid development of science and technology, the computer-aided diagnosis (CAD) systems have been widely applied to assist doctors in avoiding misdiagnosis. In 2014, Lin et al. provided the breast nodules computer-aided diagnostic system design using fuzzy cerebellar model neural networks [18]. In 2015, Bhardwaj et al. introduced the breast cancer diagnosis using genetically optimized neural network model [19]. After that, in 2017, Zhang et al. developed the longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks [20].

In this study, we use the K-means algorithm to classify the training data into *k* sets, and after that, the interval type-2 fuzzy neural network is applied to train *k* classifiers structure network with the training data. After training process, we have *k* classifiers with different parameters. Then, each testing data sample will be determined which classifier structure it belongs to. By using the K-means algorithm as the pre-classifier and the parallel structure of the IT2FNN as the main-classifier, the performance of the proposed method is superior to other methods. Then, the proposed classifier is applied to build the medical diagnosis system. The numerical experiments in breast cancer datasets of the University of California at Irvine (UCI)

2577-1655/18/$31.00 ©2018 IEEE
DOI 10.1109/SMC.2018.00703

are provided to illustrate the effectiveness of the proposed approach.

II. K-Means Interval Type-2 Fuzzy Neural Network

The general operating mechanism of the K-mean IT2FNN classifier is shown in Fig. 1, where the K-mean algorithm is the pre-classifier and the IT2FNN is the main classifier. The training process and testing process are depicted in Fig. 1a and Fig. 1b, respectively.

In the training process, firstly, the training data is divided into k groups based on its characteristics. By using the K-means algorithm, the samples in the same group will have more similar characteristics than the other groups. Consequently, each group will be used to training the IT2FNN classifier. In the testing process, in the same way, the testing data will be classified into which classifier structure it belongs to. Finally, the system performance measurements are applied to calculate the accuracy, the sensitivity and the specificity of the breast cancer datasets.

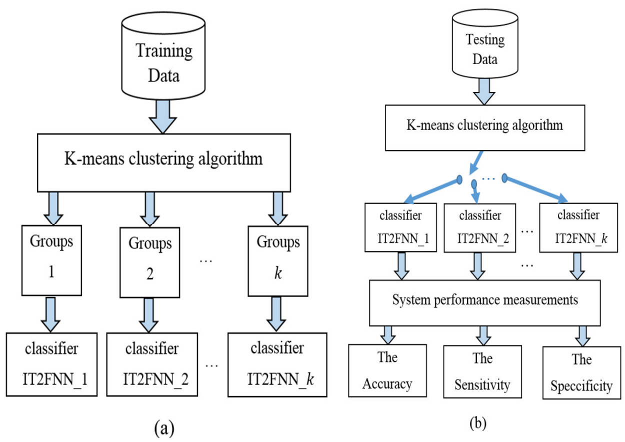

Fig. 1. The operating mechanism of K-means type-2 fuzzy neural network for training and testing process.

A. The K-Means Clustering Algorithm

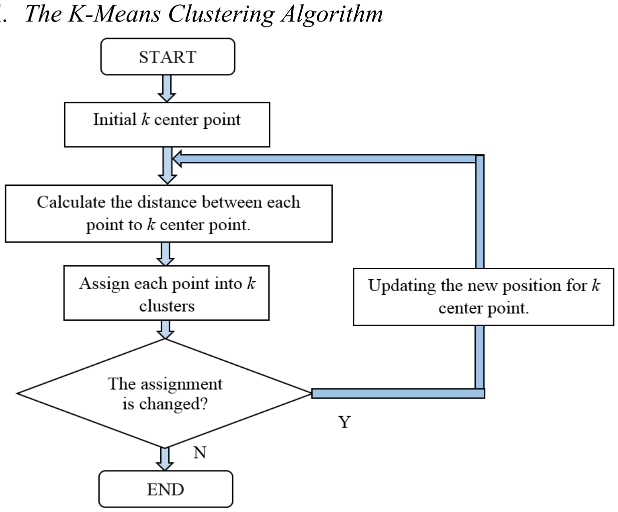

Fig. 2. The flowchart for the K-means clustering algorithm.

The financial support in part from The Ministry of Science and Technology of Republic of China under grant MOST 106-2221-E-155-MY3.
*Corresponding author

The K-means clustering algorithm is the basic algorithm for solve clustering problem. In which, data will be assigned to exactly one of k clusters defined by centroids. The operation mechanism of K-means clustering is the optimization of data distribution by minimizing the sum of the squared distances between the centroids and the data points. Thus, the K-means objective function can be given by

$$J_n = \sum_{j=1}^{k} \sum_{i}^{n} \left\| x_i^j - c_j \right\|^2 \tag{1}$$

where x_i^j is the data point position, c_j is the cluster center, j=1,2,…k is the number of clusters, and i =1,2,…n is the number of data points. The flowchart of the K-means clustering algorithm is shown in Fig. 2.

B. The Interval Type-2 Fuzzy Neural Network

The block diagram of the proposed K-mean IT2FNN classifier system is shown in Fig. 3. The structure of an interval type-2 fuzzy neural network as shown in Fig. 4 is described as below.

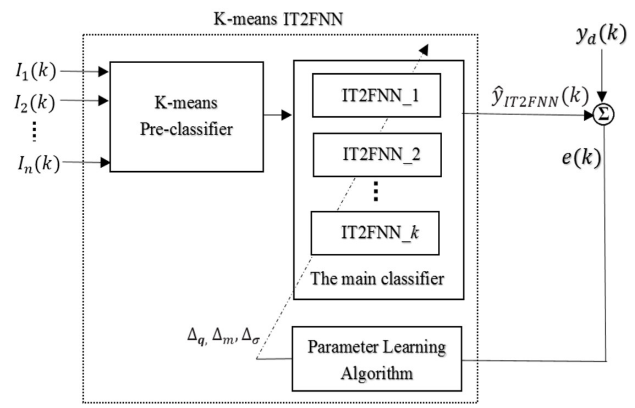

Fig. 3. Classification scheme using K-means IT2FNN.

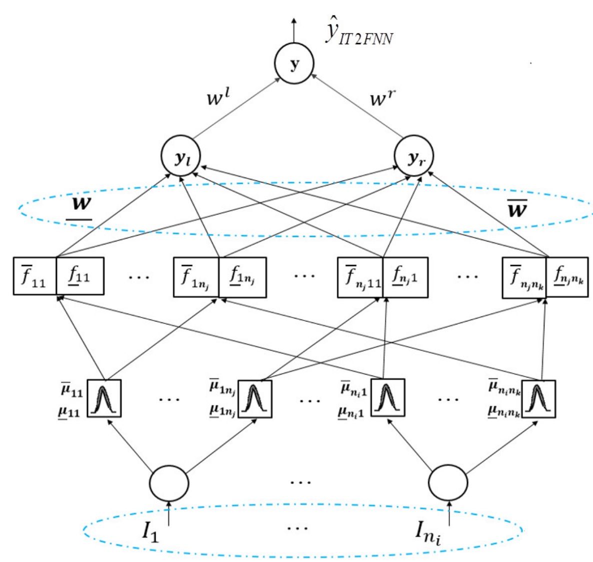

Fig. 4. Structure of the IT2FNN.

1) The Input Layer: It directly transfers the input variables into the next layer.
2) The Membership Function Layer: It can be determined based on type-2 Gaussian membership function (T2GMF)

$$\bar{\mu}_{ij} = \exp\left\{-\frac{1}{2}\left(\frac{x_i - m_{ij}}{\bar{\sigma}_{ij}}\right)^2\right\} \quad and \quad \underline{\mu}_{ij} = \exp\left\{-\frac{1}{2}\left(\frac{x_i - m_{ij}}{\underline{\sigma}_{ij}}\right)^2\right\} \tag{2}$$

3) The Firing Layer: It uses the t-norm operator. The firing strength of the ith rule is an interval value $F^i = \left[\underline{f}^i, \bar{f}^i\right]$ where

$$\bar{f}^i = \prod_{j=1}^{n} \bar{\mu}_{ij} \quad and \quad \underline{f}^i = \prod_{j=1}^{n} \underline{\mu}_{ij} \tag{3}$$

4) The Weight Layer: The initial values of weights should be given and then they will be updated by the parameter learning algorithm. The interval vector of the weight layer is symbolized by $[\bar{\boldsymbol{w}}, \underline{\boldsymbol{w}}]$.

5) The Pre-output Layer: This layer calculates the left and right outputs $\left[y_l, y_r\right]$.

$$y_l = \frac{\sum_{i=1}^{M} f_l^i \underline{w}^i}{\sum_{i=1}^{M} f_l^i} \quad and \quad y_r = \frac{\sum_{i=1}^{M} f_r^i \bar{w}^i}{\sum_{i=1}^{M} f_r^i} \tag{4}$$

where M is the number of fuzzy rules and the firing strength f_l^i and f_r^i are chosen as

$$f_l^i = \begin{cases} \bar{f}^i, & i \le L \\ \underline{f}^i, & i > L \end{cases} \quad and \quad f_r^i = \begin{cases} \underline{f}^i, & i \le R \\ \bar{f}^i, & i > R \end{cases} \tag{5}$$

where L and R represent the left and right switch points, respectively, which can be obtained using the Karnik-Mendel algorithm [21].

6) Output Layer: the final output is:

$$y = \frac{y_l + y_r}{2} \tag{6}$$

C. Parameter Learning Algorithm

The Lyapunov cost function is chosen as

$$E(k) = \frac{1}{2} e^2(k) \tag{7}$$

Then

$$\dot{E}(k) = e(k)\dot{e}(k) \tag{8}$$

where $e(k)$ is the feedback error, which can be obtained by

$$e(k) = y_d(k) - \hat{y}_{IT2FNN}(k) \in \Re \tag{9}$$

where $y_d(k)$ is desired output and $\hat{y}_{IT2FNN}(k)$ is the estimated output of IT2FNN.

The online learning gradient descent algorithm is applied to find the updated laws:

$$\hat{\underline{\boldsymbol{w}}}^i(k+1) = \hat{\underline{\boldsymbol{w}}}^i(k) + \frac{1}{2}\hat{\eta}_w e(k) \frac{f_l^i}{\sum_{i=1}^{M} f_l^i} \tag{10}$$

$$\hat{\bar{\boldsymbol{w}}}^i(k+1) = \hat{\bar{\boldsymbol{w}}}^i(k) + \frac{1}{2}\hat{\eta}_w e(k) \frac{f_r^i}{\sum_{i=1}^{M} f_r^i} \tag{11}$$

$$\hat{m}_j^i(k+1) = \hat{m}_j^i(k) + \frac{1}{2}\hat{\eta}_m s(k)\left(\frac{\left(\underline{w}^i - y_l\right)}{\sum_{i=1}^{M} f_l^i}\frac{\partial f_l^i}{\partial \hat{m}_j^i} + \frac{\left(\bar{w}^i - y_r\right)}{\sum_{i=1}^{M} f_r^i}\frac{\partial f_r^i}{\partial \hat{m}_j^i}\right) \tag{12}$$

$$\hat{\sigma}_{j1}^i(k+1) = \hat{\sigma}_{j1}^i(k) + \frac{1}{2}\hat{\eta}_\sigma s(k)\left(\frac{\left(\underline{w}^i - y_l\right)}{\sum_{i=1}^{M} f_l^i}\frac{\partial f_l^i}{\partial \hat{\sigma}_{j1}^i} + \frac{\left(\bar{w}^i - y_r\right)}{\sum_{i=1}^{M} f_r^i}\frac{\partial f_r^i}{\partial \hat{\sigma}_{j1}^i}\right) \tag{13}$$

$$\hat{\sigma}_{j2}^i(k+1) = \hat{\sigma}_{j2}^i(k) + \frac{1}{2}\hat{\eta}_\sigma s(k)\left(\frac{\left(\underline{w}^i - y_l\right)}{\sum_{i=1}^{M} f_l^i}\frac{\partial f_l^i}{\partial \hat{\sigma}_{j2}^i} + \frac{\left(\bar{w}^i - y_r\right)}{\sum_{i=1}^{M} f_r^i}\frac{\partial f_r^i}{\partial \hat{\sigma}_{j2}^i}\right) \tag{14}$$

Based on (5), f_l^i and f_r^i in (10) - (14) can be $\underline{f}^i$ or $\bar{f}^i$.

Appling the online parameter tuning laws in (10)-(14), the K-mean IT2FNN classifier is obtained and the system can achieve the desired classification performance.

Proof of the converge:

$$V(k) = E(k) = \frac{1}{2} e^2(k) \tag{15}$$

Therefore,

$$\Delta V(k) = V(k+1) - V(k) = \frac{1}{2}\left[e^2(k+1) - e^2(k)\right] \tag{16}$$

Apply the Taylor expansion, obtained

$$e(k+1) = e(k) + \Delta e(k) \cong e(k) + \left[\frac{\partial e(k)}{\partial \hat{\underline{w}}_j}\right]\Delta \hat{\underline{w}}_j \tag{17}$$

From (10), yield

$$\frac{\partial e(k)}{\partial \hat{\underline{w}}_j} = -\frac{1}{2}\frac{f_j^l}{\sum_{j=1}^{n_j} f_j^l} = \boldsymbol{\psi} \tag{18}$$

Rewrite (17), using (18) and (10)

$$e(k+1) = e(k) - \boldsymbol{\psi}\left(\hat{\eta}_w e(k)\boldsymbol{\psi}\right) = e(k)\left[1 - \hat{\eta}_w \boldsymbol{\psi}^2\right] \tag{19}$$

From (19), rewrite (16)

$$\begin{aligned} \Delta V(k) &= \frac{1}{2} e^2(k)\left[\left(1 - \eta_w \boldsymbol{\psi}^2\right)^2 - 1\right] \\ &= \frac{1}{2} e^2(k)\left[\left(\hat{\eta}_w \boldsymbol{\psi}^2\right)^2 - 2\hat{\eta}_w \boldsymbol{\psi}^2\right] \\ &= \frac{1}{2}\hat{\eta}_w e^2(k)\boldsymbol{\psi}^2\left(\hat{\eta}_w \boldsymbol{\psi}^2 - 2\right) \end{aligned} \tag{20}$$

This result has shown that, if the learning rates are chosen as a positive value and satisfy with $\hat{\eta}_w < \frac{2}{\psi^2}$, then $\Delta V(k) < 0$, therefore can obtain $V(k) > 0$, and then the stability of the system is guaranteed. By the similar way, the values for $\hat{\eta}_m, \hat{\eta}_\sigma$ can be proved.

III. Experimental Results and Discussion

The effectiveness of the designed medical diagnosis system is evaluated by three factors:

$$Accuracy(ACC) = \frac{TP + TN}{TN + TP + FP + FN} \times 100\% \tag{15}$$

$$Sensitivity(SEN) = \frac{TP}{TP + FN} \times 100\% \tag{16}$$

$$Specificity(SPE) = \frac{TN}{TN + FP} \times 100\% \tag{17}$$

where TP and TN are the number of true positives and the number of true negatives, respectively. FP and FN are the number of false positives and the number of false negatives, respectively. True or false indicates that the prediction is correct or not. Negative or positive indicates that the sample is healthy or not.

In order to show the effectiveness of the K-means FIT2FNN classifier, the Wisconsin breast cancer dataset (WBCD) is used in this study; the data of WBCD was collected at Wisconsin-Madison hospital and published in the University of California at Irvine (UCI) Machine Learning website. This dataset contains 683 sets of sample, each sample includes 9 characteristics and 1 output target. The output has 2 status, which can be represented by the binary value with 1 corresponding to the benign cases and 0 corresponding to the malignant cases. The detail of all characteristics attribute and output target are shown in Table I.

Firstly, the dataset will be separated into a training set and testing set by the ratio 70% and 30%, respectively. And then, the training data is divided into k groups based on its characteristics. For choosing the suitable value of k, the experiment was conducted 5 times with k = 1, 2, …, 5. Table II shows the experimental results of the K-mean IT2FNN classifier applied for the breast cancer diagnosis in various value of k. Choosing the k value for K-mean algorithm significantly affect system performance. From Table II, it can be seen that the proposed method can achieve the highest performance with k = 2. Figure. 5 shows the accuracy during the training process of K-mean IT2FNN with k = 2. Finally, the comparison results of the proposed method with other methods is shown in Table 3. It shows the proposed algorithm can achieve the best classification performance. For fair comparison, the experiment results in Table II is the average of 10 times with random pick up data for training and testing.

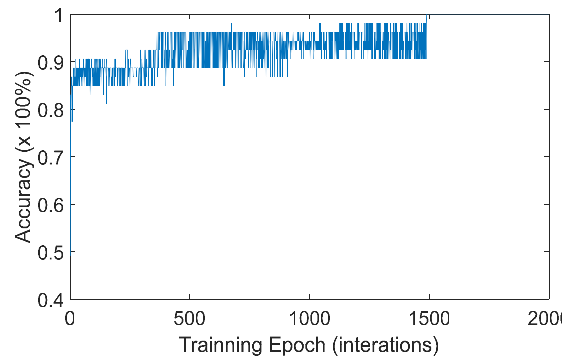

Fig. 5. The accuracy during training process of IT2FNN with K-mean algorithm (for k = 2).

TABLE I. Wisconsin Breast Cancer Dataset

No	Attribute	Description
1	Clump thickness	Measurement of thickness of clustered mass tissues
2	Uniformity of cell size	Degree of consistent cell size
3	Uniformity of cell shape	Having one form of shape
4	Marginal adhesion	The stable joining of parts to one another, which may occur abnormally
5	Single epithelial cell size	Number of layers present in epithelium
6	Bare nuclei	Having sufficient nucleus
7	Bland Chromatin	Unperturbed genetic
8	Normal nuclei	Normal round granular body composed of protein and RNA in the nucleus of a cell
9	Mitoses	The entire process of cell division of the nucleus and the cytoplasm

TABLE II. Experimental Results of the K-mean IT2FNN Classifier

k	Accuracy (%)		Sensitivity (%)		Specificity (%)	
	Training	Testing	Training	Testing	Training	Testing
1	100	96.15	100	96.61	100	95.40
2	100	97.66	100	98.02	100	97.08
3	100	96.78	100	97.38	100	95.82
4	100	96.88	100	97.23	100	96.30
5	100	97.02	100	96.95	100	97.15

TABLE III. The Comparison Result of WBC Dataset

Year	Author	Method	Accuracy (%)
2013	Stoean and Stoean [22]	SVM and evolutionary algorithm	97.23
2014	Zheng et al. [23]	K-mean and SVM	97.38
2015	Lim and Chan [24]	BK with IVFS	95.26
2016	Guan et al. [25]	SVCMAC	96.50
2018	Our method	K-mean IT2FNN (k = 2)	97.66

IV. Conclusion

In this paper, the K-mean IT2FNN classifier has been developed and successfully applied to the breast cancer diagnosis. The main contribution of this study is to successful develop a classification system, which combines the K-mean pre-classifier and the main-classifier IT2FNN. The experimental results show that the system performance can be improved by using the K-mean pre-classifier. The parameters of system can be updated using the adaptive tunning laws. Finally, the performance of the proposed system is verified by the comparison with other methods. Besides the application for breast cancer diagnosis, the proposed method also can be applied to the other classification systems.

REFERENCES

[1] H. Steinhaus, "Sur la division des corp materiels en parties," *Bull. Acad. Polon. Sci,* vol. 1, p. 801, 1956.

[2] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281-297.

[3] J. Qin, W. Fu, H. Gao, and W. X. Zheng, "Distributed K-means algorithm and fuzzy C-means algorithm for sensor networks based on multiagent consensus theory," *IEEE Transactions on Cybernetics,* vol. 47, pp. 772-783, 2017.

[4] R. Kumari, M. Singh, R. Jha, and N. Singh, "Anomaly detection in network traffic using K-mean clustering," in *Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on*, 2016, pp. 387-393.

[5] R. Shanker, R. Singh, and M. Bhattacharya, "Segmentation of tumor and edema based on K-mean clustering and hierarchical centroid shape descriptor," in *Bioinformatics and Biomedicine (BIBM), 2017 IEEE International Conference on*, 2017, pp. 1105-1109.

[6] G. Gan and M. K. P. Ng, "K-means clustering with outlier removal," *Pattern Recognition Letters,* vol. 90, pp. 8-14, 2017.

[7] P. Vora and B. Oza, "A survey on k-mean clustering and particle swarm optimization," *International Journal of Science and Modern Engineering,* vol. 1, pp. 1-14, 2013.

[8] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Information Sciences,* vol. 8, pp. 199-249, 1975.

[9] J. M. Mendel, "A quantitative comparison of interval type-2 and type-1 fuzzy logic systems: First results," in *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, 2010, pp. 1-8.

[10] O. Castillo, R. Martínez-Marroquín, P. Melin, F. Valdez, and J. Soria, "Comparative study of bio-inspired algorithms applied to the optimization of type-1 and type-2 fuzzy controllers for an autonomous mobile robot," *Information Sciences,* vol. 192, pp. 19-38, 2012.

[11] C. M. Lin, T. L. Le, and T. T. Huynh, "Self-evolving function-link interval type-2 fuzzy neural network for nonlinear system identification and control," *Neurocomputing,* vol. 275, pp. 2239-2250, 2018.

[12] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Transactions on Fuzzy Systems,* vol. 8, pp. 535-550, 2000.

[13] C. J. Kim and D. Chwa, "Obstacle avoidance method for wheeled mobile robots using interval type-2 fuzzy neural network," *IEEE Transactions on Fuzzy Systems,* vol. 23, pp. 677-687, 2015.

[14] X. Ma, P. Wu, L. Zhou, H. Chen, T. Zheng, and J. Ge, "Approaches based on interval type-2 fuzzy aggregation operators for multiple attribute group decision making," *International Journal of Fuzzy Systems,* vol. 18, pp. 697-715, 2016.

[15] M. M. Zirkohi and T. C. Lin, "Interval type-2 fuzzy-neural network indirect adaptive sliding mode control for an active suspension system," *Nonlinear Dynamics,* vol. 79, pp. 513-526, 2015.

[16] C. M. Lin and T. L. Le, "PSO-self-organizing interval type-2 fuzzy neural network for antilock braking systems," *International Journal of Fuzzy Systems,* pp. 1-13, 2017.

[17] I. Eyoh, R. John, and G. De Maere, "Interval type-2 intuitionistic fuzzy logic for regression problems," *IEEE Transactions on Fuzzy Systems,* 2017.

[18] C. M. Lin, Y. L. Hou, T. Y. Chen, and K. H. Chen, "Breast nodules computer-aided diagnostic system design using fuzzy cerebellar model neural networks," *IEEE Transactions on Fuzzy Systems,* vol. 22, pp. 693-699, 2014.

[19] A. Bhardwaj and A. Tiwari, "Breast cancer diagnosis using genetically optimized neural network model," *Expert Systems with Applications,* vol. 42, pp. 4611-4620, 2015.

[20] S. Zhang, E. Grave, E. Sklar, and N. Elhadad, "Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks," *Journal of Biomedical Informatics,* vol. 69, pp. 1-9, 2017.

[21] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*: Prentice Hall PTR Upper Saddle River, 2001.

[22] R. Stoean and C. Stoean, "Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection," *Expert Systems with Applications,* vol. 40, pp. 2677-2686, 2013.

[23] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications,* vol. 41, pp. 1476-1482, 2014.

[24] C. K. Lim and C. S. Chan, "A weighted inference engine based on interval-valued fuzzy relational theory," *Expert Systems with Applications,* vol. 42, pp. 3410-3419, 2015.

[25] J. S. Guan, L. Y. Lin, G. L. Ji, C. M. Lin, T. L. Le, and I. J. Rudas, "Breast tumor computer-aided diagnosis using self-validating cerebellar model neural networks," *Acta Polytechnica Hungarica,* vol. 13, pp. 39-52, 2016.