

A Fuzzy Brain Emotional Learning Classifier Design and Application in Medical Diagnosis

Yuan Sun, Chih-Min Lin*

Department of Electrical Engineering, Yuan Ze University, Tao-Yuan 320, Taiwan,
sungirl609@126.com

*Corresponding author, cml@saturn.yzu.edu.tw

Abstract: This paper aims to propose a fuzzy brain emotional learning classifier and applies it to medical diagnosis. To improve the generalization and learning ability, this classifier is combined with a fuzzy inference system and a brain emotional learning model. Meanwhile, different from a brain emotional learning controller, a novel definition of the reward signal is developed, which is more suitable for classification. In addition, a stable convergence is guaranteed by utilizing the Lyapunov stability theorem. Finally, the proposed method is applied for the leukemia classification and the diagnosis of heart disease. A comparison between the proposed method with other algorithms shows that this proposed classifier can be viewed as an efficient way to implement medical decision and diagnosis.

Keywords: brain emotional learning classifier; neural network; medical diagnosis

1 Introduction

Computational intelligent models have been widely used over the past decade, some popular approaches are fuzzy inference systems and fuzzy control systems [1-7], evolutionary computing techniques [8, 9] and cerebellar model articulation controllers (CMACs) [10-15]. Although, emotion was traditionally considered playing an insignificant role in intelligence for a long time, there has been a discovery of important value of emotion in human mind and behavior announced by J. E. LeDoux in 1990s [16, 17]. On the basis of this theory, a brain emotional learning (BEL) model based on neurophysiology was created by J. Moren and C. Balkenius in the early 21st century [18-20]. They have implemented a computational model of the amygdala and the orbitofrontal cortex, and tested that in simulation. Thus, there has been an increasing interest in constructing the model of emotional learning process in recent years. One of the successful implementations of this model, the Brain Emotional Learning Based Intelligent Controller (BELBIC), was introduced by C. Lucas et al. [21, 22], which became an effective method for control systems. This model is composed of two parts, mimics to the Amygdala and the Orbitofrontal cortex, respectively. The sensory

signals and an emotional cue signal are combined to generate the proper action regarding the emotional situation of the system. As an adaptive controller with fast self-learning, simple implementation and good robustness, the BELBIC has been developed to numerous applications, such as power systems [23-25], nonlinear systems [26, 27], and motor drives [28, 29].

In addition to engineering techniques, intelligent algorithms applied to medical diagnoses are also experiencing continuous attention in recent years. Early prediction with computer-aided diagnosis (CAD) acts a significant role in the clinical medicine, which can greatly increase the cure rate. Several approaches have been utilized for disease prediction or classification. For example, the logistic regression could be applied to medical diagnosis [30]. Moreover, a support vector machine [31, 32] and some other neural networks [33, 34] are also designed in this field. Most of the medical diagnosis problems are nonlinear and complex, which may be difficult to obtain definitely results even for a medical expert. Thus, it has motivated the design of more accurate and robust medical diagnosis CAD algorithms.

In this paper, we propose a classifier by incorporating fuzzy inference system with a BEL model, called FBELC model, which could not only offer a unique and flexible framework for knowledge representation, but also processes the quick learning ability of a BEL. Moreover, the online parameter adaptation laws are derived and the stable convergence is analysis for the proposed FBELC classifier. And the suggested model is designed for classifying different diseases and distinguishing presence or absence of heart disease using the posed conditions.

The specific contribution of this work involves the learning model with the definition of the emotional signal in the learning rules for classification problems. The desired goal could be obtained by appropriately choosing the system's emotional condition, which means, with the suitable definition of the reinforcing signal, the generalization quality and accuracy of prediction could be improved. The simulation results and analyses are performed to illustrate the effectiveness of the proposed model.

The remainder of the article is organized as follows: following this introductory section, Section 2 introduces the fuzzy brain emotional learning classifier, including the structure of the network, the learning algorithm and the convergence Analyses. The classifying process of diagnosis of diseases are described and the simulation analysis are shown in Section 3. Conclusion is given at last.

2 Fuzzy Brain Emotional Learning Classifier Design

This section reviews the structure of a fuzzy brain emotional learning classifier and the corresponding learning algorithm. As a whole, the general BEL model could be divided into two parts, the amygdala and the orbitofrontal cortex.

The former part receives inputs from the thalamus and from cortical areas, while the latter part receives inputs from the cortical areas and the amygdala. The output of the whole model is the output from the amygdala subtracting the inhibitory output from the orbitofrontal cortex. Usually, for control systems, the sensory input must be a function of plant output and controller output. Moreover, another reinforcing signal should be considered as a function of other signals, which is supposed as a cost function validation i.e. award and punishment are applied based on pervious defined cost function [35]. However, as a classifier, the sensory input and the reinforcing signal need to be reconsidered due to the characteristics of the input features and the cost function of the model. Detailed descriptions are presented in the following.

2.1 Structure of a FBELC

Fig. 1 shows a FBELC model with six spaces: the sensory input, sensory cortex, thalamus, orbitofrontal cortex, amygdala, and the output space. The fuzzy inference rules are defined as

$$\text{If } I_1 \text{ is } S_{1j} \text{ and } I_2 \text{ is } S_{2j}, \dots, \text{ and } I_{n_i} \text{ is } S_{n_i j}, \text{ Then } A = V_{ij} \quad (1)$$

$$\text{If } I_1 \text{ is } S_{1j} \text{ and } I_2 \text{ is } S_{2j}, \dots, \text{ and } I_{n_i} \text{ is } S_{n_i j}, \text{ Then } O = W_{ij} \quad (2)$$

for $j = 1, 2, \dots, n_j$

where n_i is the number of input dimension, S_{ij} is the fuzzy set for the i -th input and j -th layer, and A is the output of amygdala and O is the output of orbitofrontal cortex; V_{ij} is the output weight of the amygdala and W_{ij} is the output weight of orbitofrontal cortex.

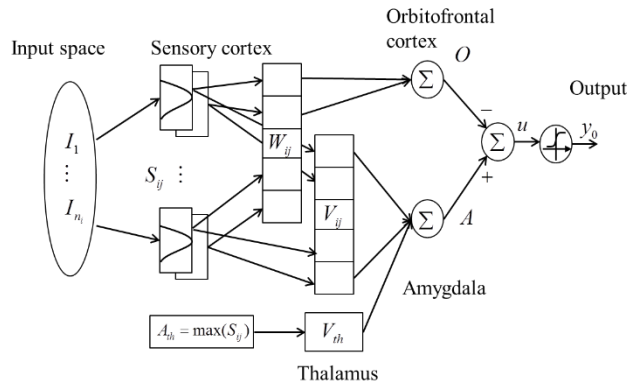


Figure 1
Architecture of a FBELC

The main functions of these six spaces are as follows:

a) *Input space*: For the input space $I = [I_1, \dots, I_i, \dots, I_{n_i}]^T \in R^{n_i}$, each input state variable, I_i can be quantized into discrete regions (called elements or

neurons). n_i is the number of input state variables. According to a given classification problem, it can also be considered as the feature dimension.

b) Sensory cortex: In this space, the sensory input should be defined and transmitted then to the orbitofrontal cortex and amygdala space. Different from the definition of the sensory input of a normal BEL model, this fuzzy brain emotional learning classifier only confirms a few activated neurons entering into the subsequent space. This is due to the consideration of improving the generalization ability and operating speed. Each block performs a fuzzy set excitation of the sensory input. Gaussian function is adopted here as the membership function, which can be represented as

$$S_{ij} = \exp\left[\frac{-(I_i - m_{ij})^2}{\sigma_{ij}^2}\right],$$

for $i = 1, 2, \dots, n_i, j = 1, 2, \dots, n_j$ (3)

where S_{ij} represents the j -th block and the i -th sensory with the mean m_{ij} and variance σ_{ij} .

c) Amygdala space: Each sensory fired value S_{ij} is multiplied by a weight V_{ij} , then the *Amygdala space* output will be

$$A = \sum_{i=1}^M \sum_{j=1}^N S_{ij} V_{ij} \quad (4)$$

d) Orbitofrontal cortex: Similarly, the orbitofrontal cortex will be

$$O = \sum_{i=1}^M \sum_{j=1}^N S_{ij} W_{ij} \quad (5)$$

Equations (4) and (5) correspond to the weighted sum outputs of the fuzzy rules in (1) and (2), respectively.

Both the weights V_{ij} and W_{ij} could be adjusted by certain learning rules. This will be addressed in Section 2.2.

e) Output space: A single-output fuzzy brain emotional learning classifier is designed as

$$y_0 = 1/[1 + \exp(-u)] \quad (6)$$

where u sums all the output from amygdala (including $A_{th}V_{th}$ term), and then subtracts the inhibitory outputs from the orbitofrontal cortex, as

$$u = A - O \quad (7)$$

2.2 The Learning Algorithm for FBELC

The learning rule of amygdala is given as follows [35]:

$$\Delta V_{ij} = \lambda_V (S_{ij} \max(0, REW - A)) \quad (8)$$

where λ_V is the learning rate in amygdala. It needs to be emphasized that, the reward signal REW , as the emotional signal, is flexible and should be determined by the system according to the biological knowledge. Many terms contribute to the definition of the reward signal for control system. Exploring the prior researches, there are different formulas used for developing the reward signals. If the network is applied to control system, the reward signal is always connected with control error, frequency deviation [36], or some other control signals, such as sliding-mode control signal [37]. Likewise, for the classification problem, it is important to define this reward signal properly. And, the formula should be determined based on the knowledge of the problem. Meanwhile, the definition of the emotional signal and tuning of the gains are not complicated. In this work, the reward signal REW can be arbitrary function of error signal and output of the model, which is selected as:

$$REW = k_1 e + k_2 u \quad (9)$$

where e is the error of the model. k_1 and k_2 are both weighting factors, which are tuned for the expectation of error reduction and output separately. Usually, the value of k_1 is bigger than that of k_2 since the error of the model is always increasingly smaller in the process of learning.

$$\Delta W_{ij} = \lambda_W (S_{ij}(A - A_{th} V_{th} - O - REW)) \quad (10)$$

From (8), obviously, ΔV_{ij} has the same plus or minus with S_{ij} , which means, in the amygdala space, once an appropriate emotional reaction is learned, it should be permanent. However, in the orbitofrontal cortex, in order to inhibit or correct signals in the amygdala space and speed up the learning process toward to the expected value, ΔW_{ij} can be increased or decreased, shown from (10).

Define the output error as

$$e = T_0 - y_0 \quad (11)$$

where T_0 is the known target of samples and y_0 is the actual output.

In most situations of a BEL controller, the sensory input is computed in the sensory cortex and is directly sent to the orbitofrontal cortex and the amygdala space. There is not any learning process in the sensory cortex. In this architecture, from (2), updating of the mean and variance of the Gaussian function should be considered, which means the sensory cortex has the learning rules. Here, the gradient descent is applied to adjust the parameters and the adaptive law of mean and variance of Gaussian function are represented as

$$\begin{aligned} \Delta m_{ij} &= -\lambda_m \frac{\partial E_0}{\partial m_{ij}} = -\lambda_m \frac{\partial E_0}{\partial e} \frac{\partial e}{\partial y_0} \frac{\partial y_0}{\partial u} \frac{\partial u}{\partial S_{ij}} \frac{\partial S_{ij}}{\partial m_{ij}} \\ &= \lambda_m e \cdot y_0 (1 - y_0) \cdot (v_{ij} - w_{ij}) \cdot S_{ij} \cdot \frac{2(l_i - m_{ij})}{\sigma_{ij}^2} \equiv \lambda_m e y_m \end{aligned} \quad (12)$$

$$\begin{aligned}
\Delta\sigma_{ij} &= -\lambda_\sigma \frac{\partial E_0}{\partial \sigma_{ij}} = -\lambda_\sigma \frac{\partial E_0}{\partial e} \frac{\partial e}{\partial y_o} \frac{\partial y_o}{\partial u} \frac{\partial u}{\partial S_{ij}} \frac{\partial S_{ij}}{\partial \sigma_{ij}} \\
&= \lambda_\sigma e \cdot y_o(1 - y_o) \cdot (v_{ij} - w_{ij}) \cdot S_{ij} \cdot \frac{2(I_i - m_{ij})^2}{\sigma_{ij}^3} \equiv \lambda_\sigma e y_\sigma
\end{aligned} \tag{13}$$

where $E_0 = \frac{1}{2}e^2$, $y_m = y_o(1 - y_o) \cdot (v_{ij} - w_{ij}) \cdot S_{ij} \cdot \frac{2(I_i - m_{ij})}{\sigma_{ij}^2}$ and $y_\sigma = y_o(1 - y_o) \cdot (v_{ij} - w_{ij}) \cdot S_{ij} \cdot \frac{2(I_i - m_{ij})^2}{\sigma_{ij}^3}$.

The learning objective could be divided into two parts. First, the parameters of the fuzzy part, m_{ij} and σ_{ij} , are adjusted by the gradient descent algorithm, shown in equations (12) and (13), respectively, which could minimize the training error, theoretically. And the other weights ΔV_{ij} and ΔW_{ij} are updated by equations (8) and (10), respectively, which are adjusted according to the structure of the brain emotional learning model. These special weights adaptation laws are the major feature that the brain emotional learning model distinguishes from the other intelligent algorithms.

A summary of this proposed FBELC model is given as below: first, original input signals are received from the features of samples. After introducing sensory input and reward signal by (2) and (9) respectively, the sensory input S_{ij} is used to form the thalamus input. From (3), a maximum term of the sensory input is selected as A_{th} . Then, signals are entered to the orbitofrontal cortex and the amygdala space by (4) and (5), respectively. By this process, the S_{ij} signal is used for both the orbitofrontal cortex and the amygdala space; however, A_{th} signal is only used for amygdala space. Before giving the total output, a comparison between the orbitofrontal cortex and the amygdala space is generated from (7). Moreover, in order to obtain a probability value between 0 to 1 for the binary classification problem, the sigmoid function is selected, as shown from (6). At last, learning processes exist in both the orbitofrontal cortex and the amygdala space by (8) and (10). They are related to the sensory inputs, reward signal, and the outputs. Besides, learning process is also done in the sensory cortex, and the adaptive law of mean and variance of Gaussian function are generated by (12) and (13).

2.3 Convergence Analyses

In the previous discussion, the learning laws in (12) and (13) call for a proper choice of the learning rates. Large values of learning rates could speed up the convergence; however, it may lead to more unstable. Therefore, we introduce the convergence theorem for selecting proper learning rates for FBELC to guarantee the stable convergence of the system.

Theorem 1: Let λ_m and λ_σ be the learning rates for the parameter of FBELC m_{ij} and σ_{ij} , respectively. Then, the stable convergence is guaranteed if λ_m and λ_σ are chosen as

$$0 < \lambda_m < \frac{2}{y_m^2} \quad (14)$$

and

$$0 < \lambda_\sigma < \frac{2}{y_\sigma^2} \quad (15)$$

Proof: A Lyapunov function is selected as

$$L(k) = \frac{1}{2} e^2(k) \quad (16)$$

The change of the Lyapunov function is

$$\Delta L(k) = L(k+1) - L(k) = \frac{1}{2} [e^2(k+1) - e^2(k)] \quad (17)$$

The predicted error can be represented by

$$e(k+1) = e(k) + \Delta e(k) \cong e(k) + \left[\frac{\partial e(k)}{\partial m_{ij}} \right] \Delta m_{ij} \quad (18)$$

where Δm_{ij} denotes the change of m_{ij} .

Using (12), it is obtained that

$$\frac{\partial e(k)}{\partial m_{ij}} = -y_o(1-y_o) \cdot (A-O) \cdot S_{ij} \cdot \frac{2(i-m_{ij})}{\sigma_{ij}^2} \quad (19)$$

Substituting (12) and (19) into (18), yields

$$e(k+1) = e(k) - \lambda_m e(k) [y_o(1-y_o) \cdot (A-O) \cdot S_{ij} \cdot \frac{2(i-m_{ij})}{\sigma_{ij}^2}]^2 = e(k) - \lambda_m e(k) y_m^2 \quad (20)$$

Thus,

$$\Delta L(k) = \frac{1}{2} [e^2(k+1) - e^2(k)] = \frac{1}{2} e^2(k) [(1 - \lambda_m y_m^2)^2 - 1] \quad (21)$$

If λ_m is chosen as (14), $\Delta L(k)$ in (21) is less than 0. Therefore, the Lyapunov stability of $L(k) > 0$ and $\Delta L(k) < 0$ is guaranteed. The proof for λ_σ can be derived in similar method, which should be chosen as (15). This completes the proof.

3 Simulation Results

The proposed FBELC model is evaluated for two illustrative examples, including classification of cancer using gene microarray data and prediction of heart disease.

These two research data sets used for the medical classification in this analysis respectively are called the Leukemia ALL/AML dataset and Statlog Heart Dataset.

The design method of the classification system consists of the following steps:

- Step 1. Obtain the information of the dataset, including the the attributes of samples and their labels. Partial attributes or entire attributes are selected as the features of samples, which are also the inputs of the model.
- Step 2. Divide the data into training set and testing set. A training set could be selected from the dataset in a certain proportion, and the rest as the test set. Sometimes, k-fold cross-validation or other cross validation methods are also used for the partition.
- Step 3. Set the initial conditions and inputs. In general, the initial values of parameters of the learning algorithm, such as m_{ij} , σ_{ij} , V_{ij} , W_{ij} , V_{th} are chosen as random. Stop criterion is set when the mean square error equals to a certain defined small value, or the iteration value reaches an upper set limit value. Besides, the number of blocks and the values of learning-rates could be determined by trial-and-error.
- Step 4. The selected features are put as the input to the model, and the output are obtained by formula (2)-(7). Then, the differences between the actual output and the given label of the training samples are adopted to modify the parameters in the network, by formula (8)-(13).
- Step 5. After training for a certain number of times, the test set are applied to the trained model and the performance are available.

3.1 Leukemia ALL/AML Dataset

a) Data description

In this experiment, we target the prediction of leukemia disease using the standard Leukemia ALL/AML data [38], which is available at website <http://www.molbio.princeton.edu/colondata>. This dataset contains 72 samples taken from 47 patients with acute lymphoblastic leukemia (ALL) and 25 patients with acute myeloid leukemia (AML). Each sample has its class label, 1 and 0, which is either ALL or AML. That means, this prediction problem could be modelled as a binary classification problem. Also, 7129 gene expression values corresponding to each patient are provided.

b) Experiment Methods

Previously, various types of gene selection methods are applied for classification on the Leukemia Datasets. These studies mostly consider how to automatically select appropriate genes, which could be associated with medical knowledge, and then obtain good results. Indeed, it is confirmed that no more than 10% of these 7129 genes are relevant to leukemia classification [38]. And without gene

selection, it is unnecessary or even harmful to classify such a few samples in such a high dimensional space. That means, it is impossible to use all the 7129 genes as features to classify this problem. Different from previous research, this work focuses on the comparison of the performance with other approaches using the same genes, and also the analysis of stability of the performance using different genes, simply because we mainly intend to illustrate the classification performance of the FBELC model. Our experiments are carried out using different selected genes as the inputs of the classifier, and the results are compared with other models.

For the classification system, the number of inputs is exactly the number of genes which are selected. If five genes are selected as features, that means, the input consists of five dimensions. The number of blocks could be adjusted from 10 to 50, determined by trial-and-error. It mainly influences the training time in the learning process. The training data are used to train the proposed FBELC offline. Other parameters, including m_{ij} , σ_{ij} , V_{ij} , W_{ij} , V_{th} are randomly initialized and the weighting factors k_1 and k_2 are tuned to be 100 and 1, respectively. Stop criterion is set to the limit of 200 training epochs with the learning rate $\lambda_V = \lambda_W = \lambda_m = \lambda_\sigma = 0.0001$.

c) Result Analysis

In our experiment, two evaluation methods are considered to divide the available data into a training dataset and a test dataset. Firstly, as provided by the standard Leukemia ALL/AML data, the training dataset consists of 38 samples (27 ALL and 11 AML) from bone marrow specimens, while the testing dataset has 34 samples (20 ALL and 14 AML), which are prepared under different biological experimental conditions.

Table 1 shows the comparison of performance of different methods on Leukemia dataset by training on 38 samples and testing on 34 samples. We adopt the accuracy to measure the performance of all the approaches. Shujaat Khan [39] proposes an RBF network with a novel kernel and selects the top five genes [40] for the experiment and provides 97.07% of accuracy. Meanwhile, using the same five genes as the features of our model, the same result is obtained. It is evident that the 97.07% of accuracy means there are only one of the 34 samples has been misclassified. Considering the small sample size for testing, this result is acceptable. The other two approaches, those of Tang [41], and Krishna Kanth [42], respectively select 15 genes and 2 genes as inputs of each classifier by achieving both 100% accuracy, that means the 34 testing samples are all classified correctly. It is shown in Table 1 that, using the FBELC model, whether choosing 15 genes or 2 genes, the accuracy is both 97.07%. This result illustrates that the benefit of the proposed model seems to be its stability and high-accuracy performance.

Moreover, another evaluation method, Leave-One-Out Cross Validation (LOOCV) is also used to verify the proposed classifier. The LOOCV method is usually used to select a model with good generalization and to evaluate predictive performance.

At each LOOCV step, this method holds out one sample for testing while the remaining samples are used for training. The overall test accuracy is calculated based on each testing samples. The LOOCV method has the advantages of producing model estimates with less bias and more ease in smaller samples.

Table 1

Comparison of performance of different methods on Leukemia dataset by training on 38 samples and testing on 34 samples

| Method | No.of genes/features | Gene accession number/Gene index | Accuracy(%) | |
|------------------------------|----------------------|--|-------------|----------|
| | | | References | Our work |
| RBF with a novel kernel [39] | 5 | Top-5-ranked | 97.06 | 97.06 |
| FCM-SVM-RFE [41] | 15 | M83652, X85116, D49950, U50136, M24400, Y12670, L20321, M23197, M20902, X95735, M19507, L08246, M96326, X05409, M29610 | 100 | 97.06 |
| MFHSNN [42] | 2 | X95735, M27891 | 100 | 97.06 |

Hence, for this Leukemia ALL/AML dataset, LOOCV method consists of splitting the dataset randomly into 72 samples. At each of the 72-th iteration, 71 samples will be used as training sample and the left-out sample will be used as the test sample. For each step we obtain a test accuracy and the final accuracy equals to the average value of 72 iterations. If we get the 98.61% accuracy, it means in each test of LOOCV, there is only one error of 72 times test.

Table 2

Comparison of performance of different methods on Leukemia dataset by LOOCV

| Method | No.of genes/features | Gene accession number/Gene index | Accuracy(%) | |
|--------------------------|----------------------|----------------------------------|-------------|----------|
| | | | References | Our work |
| Wrapper method +SVM [43] | 5 | Top-5-ranked | 98.61 | 98.61 |
| NB [43] | 4 | Top-4-ranked | 94.44 | 98.61 |
| Wrapper method + NB [43] | 3 | Top-3-ranked | 98.61 | 98.61 |
| SVM [44] | 3 | X95735, M31523, M23197 | 97.22 | 98.61 |

Table 2 compares the performance of different methods on Leukemia dataset by LOOCV. The first three approaches, those of Peng [43], report 98.61% of accuracy with the top-5-ranked genes by wrapper method and SVM, 94.44% of accuracy with the top-4-ranked genes by NB and RBF and 98.61% of accuracy

with the top-3-ranked genes by NB method. The last approach is that of Wang [44], involves the model of MFHSNN and get the 97.22% of accuracy. However, according to the same selected genes in previous work, respectively, from the results in Table 2, the proposed FBELC model, gives 98.61% accuracy consistently. In other word, we can confirm in spite of that the attained performance being similar to other methods, the proposed FBELC is superior to other approaches for its good generalization.

3.2 Statlog Heart Dataset

a) Sample Datasets

The Statlog heart disease dataset used in our work is published and shared in the UCI machine learning database [45]. It contains 270 observations, which belongs two classes: the presence and absence of heart disease. Every sample includes 13 different features, including some conditions and symptoms of the patients. Thus, some of the attributes are real value and some are binary or nominal type.

b) Performance Evaluation

This Statlog dataset is commonly used among researchers for classification. Some studies used all the 13 features as inputs of the classifiers. Others proposed some of features are irrelevant to the learning process and feature selection was used to improve learning accuracy and decrease training and testing time. To illustrate the effectiveness of the proposed algorithm, both of these scenarios are considered and the results are compared with other studies.

Table 3

Comparison of performance of different methods on Statlog dataset with all features

| Author(Year) | Method | Training-test partitions | Accuracy(%) |
|--------------------------|----------|--------------------------|-------------|
| Subbulakshmi [46] (2012) | ELM | 70%-30% | 87.5 |
| Lee [47] (2015) | NEWFM | 5-fold CV | 81.12 |
| Hu [48] (2013) | RSRC | 5-fold CV | 84.0 |
| | Our work | 70%-30% | 89.41 |
| | Our work | 5-fold CV | 85.93 |
| | Our work | 75%-25% | 92.54 |

Table 3 compares the accuracies of our algorithm with other approaches using all 13 features. The first method, that of Subbulakshmi [46], involves an extreme learning machine (ELM) for two category data classification problems and evaluated on the Statlog datasets. And the accuracy of 87.5% is the mean value for 50 runs. Those of Lee [47] and Hu [48], use 5-fold cross validation (CV) method to make results more credible and the accuracies of 81.12% and 84.0% are obtained, respectively. The results in our study, separately using the same training-test partitions as other approaches, shows better performance evidently.

Besides, using 75%-25% training-test partitions, the highest classification performance is also provided in Table 3.

On the other hand, there are some other feature selection and classification approaches based on this Statlog dataset. They use not all the 13 features as the inputs of models. The comparison of the classification accuracies of our study and previous methods, not using all features, are summarized in Table 4. The first approach [49], based on the LSTSVM model with 11 features, has achieved the accuracies of 85.19%, 87.65% and 83.93% using 50-50%, 70-30% and 80-20% partitions, respectively. The other method, that of Lee [47], selects 10 features and obtains the accuracy of 82.22%, using 5-fold CV. Using the same selected features and training-test partitions, the results shows our proposed method obtains superior and consistent performance. Indeed, the last two approaches select fewer features and obtain higher accuracies than other researches. For example, Liu [50] uses four classifiers with the same 7 features, and achieves the accuracy of 83.33%, 85.19%, 87.03% and 92.59%, respectively. Ertugrul [51] selects only 3 features, which are all nominal type and obtains the accuracy of 85.93% by extreme learning machine. All these feature selection methods are combined to certain algorithms. The accuracy displayed in Table 4 remains relatively high value, which demonstrates good robustness of the proposed model.

The results of our proposed algorithm are based on the fixed structure of the FBELC model and the parameters setting for the simulations using different selected features are presented in Table 5. It exhibits that the weighting factors, number of blocks and learning rate does not change much when using different selected features.

From the simulation results of two examples, with different feature selection or different training-testing partitions, the fuzzy emotional learning classifier can always perform well. From theoretical analysis, it is proved that satisfactory performance could be obtained by choosing appropriate emotional signals, according to the characteristics of the classification problems.

Table 4

Comparison of performance of different methods on Statlog dataset with certain features

| Author(Year) | Method | No. of genes/features | Training-test partitions | Accuracy(%) | |
|----------------------|--------|-----------------------|--------------------------|-------------|-------|
| Tomar [49] (2014) | LSTSVM | 11 | 50%-50% | 85.19 | 87.41 |
| | | | 70%-30% | 87.65 | 88.89 |
| | | | 80%-20% | 83.93 | 90.74 |
| Lee [47] (2015) | NEWFM | 10 | 5-fold CV | 82.22 | 85.19 |
| Liu [50] (2017) | RFRS | 7 | 70%-30% | 92.59 | 85.19 |
| Ertugrul [51] (2016) | ELM | 3 | 9-fold CV | 85.93 | 83.33 |

Table 5
List of classification parameters using different features

| Parameters | Value | | | |
|---------------|-------------|-------------|------------|------------|
| | 11 features | 10 features | 7 features | 3 features |
| k_1 | 50 | 50 | 30 | 30 |
| k_2 | 1 | 1 | 1 | 1 |
| No. of blocks | 50 | 50 | 30 | 30 |
| Learning rate | 0.00001 | 0.00001 | 0.00001 | 0.00001 |

Conclusions

This study has successfully proposed a FBELC for classification. The novelty of this paper lies in the proposed approach: the incorporation of the fuzzy inference system and a BEL model, and the application to medical diagnosis. The classification efficiency can be improved specifically because of the fuzzy set and the novel setting of reward signal in the model, which can cause better generalization and faster learning. Meanwhile, two medical datasets are applied to test the developed FBELC model. From the simulation results, it is shown that the proposed algorithm can perform high generalization and good accuracy, while being simple and easily implementable. Therefore, the results indicate that the proposed classifier can be used as a promising alternative tool in medical decision and diagnosis. The data used for simulations all come from public medical experimental datasets; in the future, we could cooperate with some medical institutions and apply our algorithm in practical experiments.

References

- [1] T. Islam, P. K. Srivastava, M. A. Rico-Ramirez, Q. Dai, D. Han and M. Gupta, "An exploratory investigation of an adaptive neuro fuzzy inference system (ANFIS) for estimating hydrometeors from TRMM/TMI in synergy with TRMM/PR," *Atmospheric Research*, Vol. 145-146, pp. 57-68, 2014
- [2] K. Subramanian, S. Suresh and N. Sundararajan, "A metacognitive neuro-fuzzy inference system (McfIS) for sequential classification problems," *IEEE Transactions on Fuzzy Systems*, Vol. 21, No. 6, pp. 1080-1095, 2013
- [3] M. Shanbedi, A. Amiri, S. Rashidi, SZ. Heris and M. Baniadam, "Thermal performance prediction of two-phase closed thermosyphon using adaptive neuro-fuzzy inference system," *Heat Transfer Engineering*, Vol. 36, No. 3, pp. 315-324, 2015
- [4] S. Babu Devasenapati and K. I. Ramachandran, "Hybrid fuzzy model based expert system for misfire detection in automobile engines," *International Journal of Artificial Intelligence*, Vol. 7, No. A11, pp. 47-62, 2011
- [5] T. Haidegger, L. Kovács, R. E. Precup, B. Benyo, Z. Benyo, S. Preitl, "Simulation and control for telerobots in space medicine," *Acta Astronautica*, Vol. 81, No. 1, pp. 390-402, 2012

- [6] M. Jovic, E. Pap, A. Szakál, D. Obradovic, Z. Konjovic, "Managing big data by directed graph node similarity," *Acta Polytechnica Hungarica*, Vol. 14, No. 2, pp. 183-200, 2017
- [7] S. Vrkalic, E. C. Lunca and I. D. Borlea, "Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants," *International Journal of Artificial Intelligence*, Vol. 16, No. 2, pp. 208-222, 2018
- [8] F. Jiménez, G. Sánchez and JM. Juárez, "Multi-objective evolutionary algorithms for fuzzy classification in survival prediction," *Artificial Intelligence in Medicine*, Vol. 60, No. 3, pp. 197-219, 2014
- [9] A. Gotmare, SS. Bhattacharjee, R. Patidar and NV. George, "Swarm and evolutionary computing algorithms for system identification and filter design: A comprehensive review," *Swarm and Evolutionary Computation*, Vol. 32, pp. 68-84, 2017
- [10] J. S. Guan, G. L. Ji, L. Y. Lin, C. M. Lin, T. L. Le and I. J. Rudas, "Breast tumor computer-aided diagnosis using self-validation cerebellar model neural networks," *Acta Polytechnica Hungarica*, Vol. 13, No. 4, pp. 39-52, 2016
- [11] C. M. Lin and Y. F. Peng, "Adaptive CMAC-based supervisory control for uncertain nonlinear systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 34, No. 2, pp. 1248-1260, 2004
- [12] C. M. Lin, and H. Y. Li, "Adaptive dynamic sliding-mode fuzzy CMAC for voice coil motor using asymmetric Gaussian membership function," *IEEE Transactions on Industrial Electronics*, Vol. 61, No. 10, pp. 5662-5671, 2014
- [13] C. H. Lee, F. Y. Chang and C. M. Lin, "An efficient interval type-2 fuzzy CMAC for chaos time-series prediction and synchronization," *IEEE Transactions on Cybernetics*, Vol. 44, No. 3, pp. 329-341, 2014
- [14] C. M. Lin and C. H. Chen, "Robust fault-tolerant control for a biped robot using a recurrent cerebellar model articulation controller," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 37, No. 1, pp. 110-123, 2007
- [15] C. M. Lin and H. Y. Li, "Intelligent control using the wavelet fuzzy CMAC backstepping control system for two-axis linear piezoelectric ceramic motor drive systems," *IEEE Transactions on Fuzzy Systems*, Vol. 22, No. 4, pp. 791-802, 2014
- [16] J. E. LeDoux, *The Amygdala: Neurobiological Aspects of Emotion*, Wiley-Liss, New York, 1992, pp. 339-351
- [17] J. E. LeDoux, "Emotion: clues from the brain," *Annual Review of Psychology*, Vol. 46, pp. 209-235, 1995

-
- [18] J. Morén and C. Balkenius, "A computational model of emotional learning in the amygdala," *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, Cambridge, Mass. The MIT Press, 2000
- [19] C. Balkenius and J. Moren, "Emotional learning: A computational model of the amygdala," *Cybernetics and Systems*, Vol. 32, No. 6, pp. 611-636, 2001
- [20] J. Moren, *Emotion and Learning-A Computational Model of the Amygdala*, PhD Dissertation, Lund University, 2002
- [21] C. Lucas, D. Shahmirzadi and N. Sheikholeslami, "Introducing BELBIC: Brain emotional learning based intelligent controller," *International Journal of Intelligent Automation and Soft Computing*, Vol. 10, No. 1, pp. 11-21, 2004
- [22] M. Roshanaei, E. Vahedi and C. Lucas, "Adaptive antenna applications by brain emotional learning based on intelligent controller," *IET Microwaves, Antennas & Propagation*, Vol. 4, No. 12, pp. 2247-2255, 2010
- [23] M. Hosseinzadeh Soreshjani, G. Arab Markadeh, E. Daryabeigi, N. R. Abjadi and A. Kargar, "Application of brain emotional learning-based intelligent controller to power flow control with thyristor-controlled series capacitance," *IET Generation, Transmission & Distribution*, Vol. 9, No. 14, pp. 1964-1976, 2015
- [24] S. A. Aghaei, C. Lucas, and K. Amiri Zadeh, "Applying brain emotional learning based intelligent controller (BELBIC) to multiple-area power systems," *Asian Journal of Control*, Vol. 14, No. 6, pp. 1580-1588, 2012
- [25] S. A. N. Niaki, R. Irvani and M. Noroozian, "Power-flow model and steady-state analysis of the hybrid flow controller," *IEEE Transactions on Power Delivery*, Vol. 23, No. 4, pp. 2330-2338, 2008
- [26] G. Huang, Z. Zhen and D. Wang, "Brain emotional learning based intelligent controller for nonlinear system," *2008 Second International Symposium on Intelligent Information Technology Application*, Shanghai, 2008, pp. 660-663
- [27] M. M. Polycarpou, "Fault accommodation of a class of multivariable nonlinear dynamical systems using a learning approach," *IEEE Transactions Automatic Control*, Vol. 46, pp. 736-742, 2001
- [28] H. A. Zarchi, E. Daryabeigi, G. R. A. Markadeh and J. Soltani, "Emotional controller (BELBIC) based DTC for encoderless synchronous reluctance motor drives," *2011 2nd Power Electronics, Drive Systems and Technologies Conference*, Tehran, 2011, pp. 478-483
- [29] M. A. Rahman, R. M. Milasi, C. Lucas, B. N. Araabi and T. S. Radwan, "Implementation of emotional controller for interior permanent-magnet synchronous motor drive," *IEEE Transactions on Industry Applications*, Vol. 44, No. 5, pp. 1466-1476, 2008

- [30] S. K. Agarwal and R. Kumar, "Explication of a logistic regression driven hypothesis to strengthen derivative approach driven classification for medical diagnosis," 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, Bhopal, 2014, pp. 1-6
- [31] H. Azzawi, J. Hou, Y. Xiang and R. Alanni, "Lung cancer prediction from microarray data by gene expression programming," *IET Systems Biology*, Vol. 10, No. 5, pp. 168-178, 2016
- [32] M. Tan, B. Zheng, J. K. Leader and D. Gur, "Association between changes in mammographic image features and risk for near-term breast cancer development," *IEEE Transactions on Medical Imaging*, Vol. 35, No. 7, pp. 1719-1728, 2016
- [33] J. Zhao, L. Y Lin, and C. M. Lin, "A general fuzzy cerebellar model neural network multidimensional classifier using intuitionistic fuzzy sets for medical identification," *Computational Intelligence and Neuroscience*, 2016
- [34] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Transactions on Medical Imaging*, Vol. 35, No. 5, pp. 1207-1216, 2016
- [35] AR Mehrabian, C. Lucas, "Emotional learning based intelligent robust adaptive controller for stable uncertain nonlinear systems," *World Academy of Science, Engineering and Technology* Vol. 19, pp. 1027-1033, 2008
- [36] E. Bijami, R. Abshari, S. M. Saghaiannejad and J. Askari, "Load frequency control of interconnected power system using brain emotional learning based intelligent controller," 2011 19th Iranian Conference on Electrical Engineering, Tehran, 2011, pp. 1-6
- [37] C. F. Hsu, C. T. Su and T. T. Lee, "Chaos synchronization using brain-emotional-learning-based fuzzy control," *IEEE Joint, International Conference on Soft Computing and Intelligent Systems*, 2016, pp. 811-816
- [38] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield and E. S. Lander, "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring," *Science*, Vol. 286, No. 5439, pp. 531-537, 1999
- [39] S. Khan, I. Naseem, R. Togneri and M. Bennamoun, "A novel adaptive kernel for the RBF neural networks," *Circuits Systems and Signal Processing*, Vol. 36, No. 4, pp. 1639-1653, 2017
- [40] H. C. Peng, F. H. Long and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine*

- Intelligence, Vol. 27, No. 8, pp. 1226-1238, 2005
- [41] Y Tang, Y. Q. Zhang and Z Huang, "FCM-SVM-RFE gene feature selection algorithm for leukemia classification from microarray gene expression data," The 14th IEEE International Conference on Fuzzy Systems, Reno, 2005, pp. 97-101
- [42] B. B. M. Krishna Kanth, U. V. Kulkarni and B. G. V. Giridhar, "Gene expression based acute leukemia cancer classification: a neuro-fuzzy approach," International Journal of Biometric & Bioinformatics, Vol. 4, No. 4, pp. 136-146, 2010
- [43] Y Peng, W. Li, and Y. Liu, "A hybrid approach for biomarker discovery from microarray gene expression data for cancer classification," Cancer Information, Vol. 2, No. 1, pp. 301-311, 2006
- [44] S Wang, H Chen, R Li and D. Zhang, "Gene selection with rough sets for the molecular diagnosing of tumor based on support vector machines," International Computer Symposium, Taiwan, 2006, pp. 1368-1373
- [45] UCI Repository of Machine Learning Databases, <http://archive.ics.uci.edu/ml/datasets/Statlog+%28Heart%29>
- [46] C. V. Subbulakshmi, S. N. Deepa and N. Malathi, "Extreme learning machine for two category data classification," 2012 IEEE International Conference on Advanced Communication Control and Computing Technologies, Ramanathapuram, 2012, pp. 458-461
- [47] S. H. Lee, "Feature selection based on the center of gravity of BSWFMs using NEWFM," Engineering Applications of Artificial Intelligence, Vol. 45, pp. 482-487, 2015
- [48] Y. C. Hu, "Rough sets for pattern classification using pairwise-comparison-based tables," Applied Mathematical Modelling, Vol. 37, No. 12-13, pp. 7330-7337, 2013
- [49] D. Tomar and S. Agarwal, "Feature selection based least square twin support vector machine for diagnosis of heart disease," International Journal of Bio-Science and Bio-Technology, Vol. 6, No. 2, pp. 69-82, 2014
- [50] X. Liu, X. Wang, Q. Su, M. Zhang, Y. Zhu, Q. Wang and Q. Wang, "A hybrid classification system for heart disease diagnosis based on the rfrs method," Computational and Mathematical Methods in Medicine, Vol. 2017, No. 3, pp. 1-10, 2017
- [51] F. Ertugrul, "Determining the order of risk factors in diagnosing heart disease by extreme learning machine," International Conference on Natural Science and Engineering, Kilis, 2016, pp. 10-19