

## Research Article

# A Fuzzy Neural Network-Based System for Alleviating Students' Boredom in English Learning

Liuhui Yang and Xiuying Wu 

*Department of Foreign Languages and General Studies, Shenyang Urban Construction University, Shenyang, Liaoning, China 110167*

Correspondence should be addressed to Xiuying Wu; [wuxiuying@syucu.edu.cn](mailto:wuxiuying@syucu.edu.cn)

Received 14 March 2022; Revised 29 March 2022; Accepted 7 April 2022; Published 25 May 2022

Academic Editor: Chia-Huei Wu

Copyright © 2022 Liuhui Yang and Xiuying Wu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to explore students' boredom in English learning, a recognition algorithm based on fuzzy neural network is proposed. The algorithm selects Gaussian membership function and initializes the clustering center obtained by fuzzy c-means algorithm to the center of Gaussian function, and the width of Gaussian function is obtained by the membership and center of fuzzy c-means clustering algorithm. In the construction of base classifier, diversity strategy is adopted to increase its diversity and complementarity. In the selection of base classifiers, the combination of contour coefficient and clustering algorithm is used to determine the number of classifiers to be fused, and the inconsistency measurement method is used to evaluate their differences. In the combination strategy, we learn from the Bayesian thought and dynamically adapt the weight of learning base classifier based on its a priori probability and class conditional probability. The experimental results show that 10 times of cross-validation are carried out, respectively, and the accuracy of each algorithm is given. The algorithm based on tree structure obviously has better performance, followed by the rule-based algorithm, and finally the fuzzy neural network algorithm based on neural network, while the accuracy of SVM and logistic regression algorithm LR is lower. It is proved that the fuzzy neural network can effectively identify students' boredom in English learning.

## 1. Introduction

With the continuous development of the computer industry, the relationship between computers and people is becoming closer and closer. It is of great significance for computers to perceive human emotions for improving human-computer interaction experience. Voice emotion sensing technology bridges this field and makes it possible for machines to understand people's emotions. Voice emotion sensing technology, that is, voice emotion sensing technology, mainly refers to allowing the computer to recognize different human emotional states by receiving voice signals. Human speech in different emotional states has very obvious features, such as sound quality features, prosodic features, spectral features, and other features. These features can help computers better distinguish different emotional states [1]. The deep neural network originates from the artificial neural network and

the multilayer perceptron model. Based on the original two-layer perceptron model, the deep neural network can approach any mathematical continuous function in theory, and the feature expression ability of the model has been better expanded. At the same time, due to the increase of levels, the features can become more abstract, more features can be expressed with the same number of features, and the classification effect becomes better. Convolutional neural network and cyclic neural network are very popular in the field of deep learning. They have made breakthroughs in speech recognition, expression recognition (Figure 1), and natural language processing [2].

Human beings have invented words to record the communication information in time and space. In the process of communication, the voices of both sides of communication are different, and they will have different interpretations of each other. Only the description of words, the dialogue

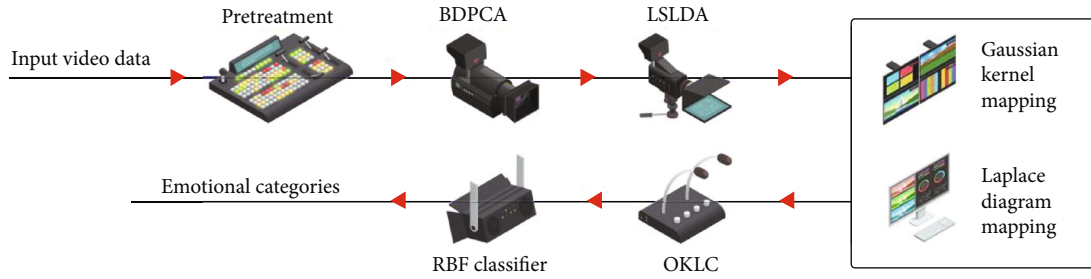


FIGURE 1: Expression recognition system.

will have no emotion and be straightforward. Such information lacks dimension. In modern society, people have also invented many ways to record information, such as pictures, movies, operas, and novels [3]. These inventions add more interest and convenience to life and bring massive data to life at the same time. Processing and analyzing these data is a time-consuming and human task, and facial emotion, as a supplement to language communication, can more three-dimensionally express people's real situation in the case of strong language subjectivity. In interpersonal communication, people can have 65% cognitive judgment of each other in the first seven seconds of acquaintance. People's facial emotions account for 55% of the communication. In addition, the tone of voice accounts for 38% of the communication, and only 7% of the real information about each other's communication content. From the above data, it can be seen that facial emotion has a great influence on communication and is an indispensable part. As a member of the tide of artificial intelligence, deep learning has become the focus of academic and industrial circles with its high-performance performance and theoretical precipitation over the years. At present, the demand for intelligent manufacturing in the industry is becoming stronger and stronger, and face emotion recognition has also got a good development opportunity. Facial emotion can reflect the individual's internal psychological state to the greatest extent in facial information, which has great research value. The following are the related applications of facial emotion recognition [4].

## 2. Literature Review

Emotion is a physiological and psychological state accompanied by the process of cognition and consciousness. It plays a very important role in interpersonal communication. In recent years, the first mock exam has been done by many researchers at home and abroad. Most of these researches are to distinguish six basic facial expressions from single modal emotion recognition to multimodal emotion recognition. Kahaki proposed a multimodal method of facial expression and speech based on the decision-making level fusion mechanism and achieved a certain fusion effect [5]. Fatahi proposed a multimodal method of facial expression and speech based on Multi-Stream Hidden Markov Model (HMM). After fusion, the recognition rate was 72.42% [6]. Kanya used the kernel cross model factor analysis method

to reduce the dimension and fuse the features of speech mode and facial expression mode. However, for the same facial expression, the expression intensity is different, and the emotional state will be different. For example, the expression degree of anger can be divided into different levels from anger to anger. Anger indicates a slight degree of anger; anger means that people are in a very angry state [7]. Calculating the intensity of facial expressions is an extension of the process of recognizing emotions, and there are relatively few researchers in this regard. For Kanya, the sun used a 6-level classification of binary classifiers for hierarchical division. They assessed the intensity of Au from a unit of activity representing several infant facial features [8]. However, it is not possible to avoid class overlap between classes, which makes these methods more robust. In Soria's study, they used a picture-based classification method to calculate intensity at three levels (low, medium, and high). From the above research, it can be observed that the level division of expression is preset. Therefore, this paper proposes an expression level subdivision method based on fuzzy clustering. This method adopts a clustering fusion algorithm, which can automatically obtain the intensity level without setting the intensity level in advance [9]. Jebadass inputs a large number of facial expression pictures of different races, ages and genders into the computer and observes the changes of important facial feature points such as the texture and wrinkles of all faces and the changes of shape through the algorithm, so as to identify people's current emotions. The algorithm can accurately identify seven emotions: happiness, sadness, surprise, anger, contempt, disgust, and fear [10]. Cui experimentally defined human emotions as six basic emotions: happiness, anger, disgust, sadness, fear, and surprise. This emotion model has the universal applicability of problem description and problem-solving for easily identifiable emotions. However, this discrete emotion model does not have wide adaptability in subtle, complex, and abstract emotion description. At the same time, there are strong subjective factors in the standard of emotional judgment. For example, people have individual differences in the distinction between fear and disgust. Secondly, there is no quantifiable dimensional system to make the whole emotion model have the ability to describe the relationship and degree, and it is impossible to use the tools of vector and matrix for quantitative description. Hinton proposed an algorithm for parameter initialization of multi-layer structure deep neural network in 2006. Using this

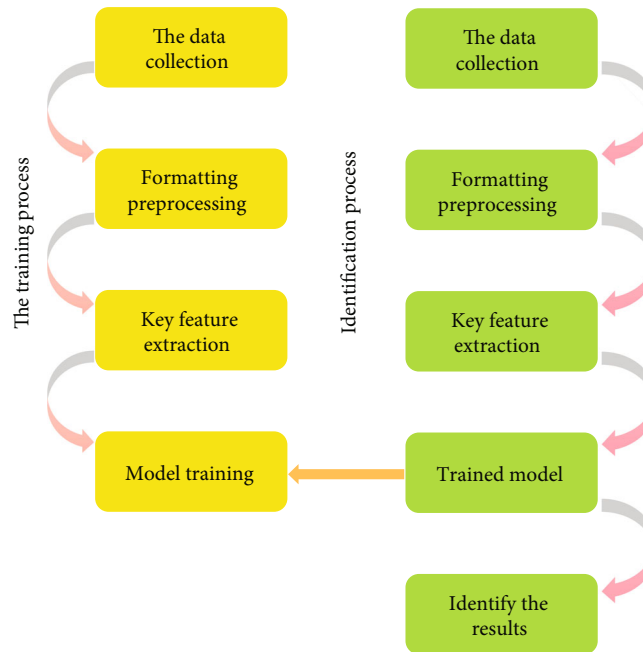


FIGURE 2: Training and recognition process of speech emotion perception.

algorithm, it is easier to make the network converge, and the network can be designed deeper. This contribution is considered to be the cornerstone and beginning of modern deep neural networks. With the improvement of computer computing power, the optimization of computing framework and the accelerated development of image datasets with the advent of the Internet era. The literature proposes to use artificial neural network ANN for nonlinear model mapping, which is divided into three types: multilayer perceptron MLP, cyclic neural network RNN, and restricted Boltzmann neural network RBN. The recognition effect of ANN is poor, with an average of about 65% [11]. Sabri has developed a speech emotion perception platform, which can recognize emotions through sound quality analysis. At present, the emotion states that can be analyzed include five basic emotions, and the recognition accuracy is 70% to 80%, which is 60% higher than the human average [12].

Based on the current research, a recognition algorithm based on fuzzy neural network is proposed. The algorithm selects Gaussian membership function and initializes the clustering center obtained by fuzzy c-means algorithm to the center of Gaussian function, and the width of Gaussian function is obtained by the membership and center of fuzzy c-means clustering algorithm. In the construction of base classifier, diversity strategy is adopted to increase its diversity and complementarity; in the selection of base classifiers, the combination of contour coefficient and clustering algorithm is used to determine the number of classifiers to be fused, and the inconsistency measurement method is used to evaluate their differences; in the combination strategy, we learn from Bayesian thought and dynamically adapt the weight of learning base classifier based on its a priori probability and class conditional probability.

### 3. Traditional and Fuzzy Neural Emotion Recognition System

#### 3.1. Traditional Recognition Model

*3.1.1. Overall Design of System Prototype.* Figure 2 shows the overall process of speech emotion perception, which is in line with the general process of pattern recognition, machine learning, and classification.

Every step in the process of speech emotion perception is very key and will have a very important impact on the results.

(1) *Data Collection.* Whether recognition or training, we need to collect samples and collect the original data of the research object. However, for the collected original data, we should pay attention to whether the distribution of the samples is uniform and whether the samples truthfully reflect the essence of the characteristics. The quality of the data directly determines the quality of the recognition results [13]. At the same time, the model also depends on the experimental samples, so it is necessary to experiment on multiple data samples to determine the independence and universality of the model. The samples used in this paper include CASIA, EMO, and SAVEE. The author will carry out experiments on these three databases in the future.

(2) *Format Preprocessing.* The collected samples cannot always be taken directly for use. It is necessary to sort and label the data, which is called formatting. For voice data, endpoint detection, pre-emphasis, windowing, and framing are also needed. At the same time, the data needs to be transformed to meet the needs of the model.

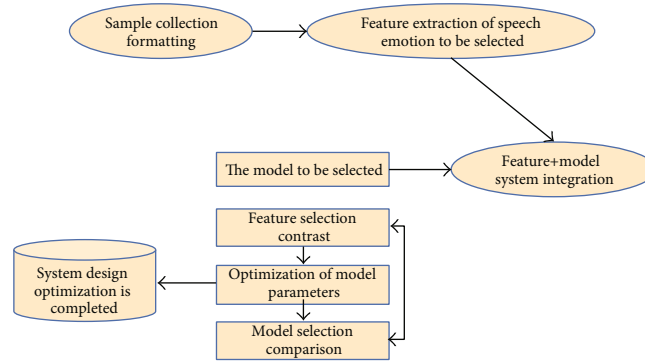


FIGURE 3: Design process of emotion perception system.

(3) *Key Feature Extraction.* The effect of machine recognition largely depends on the quality of features. The features used in this paper mainly include Mel cepstrum coefficient (MFCC) and spectrum. However, the characteristic effects under different parameters are different, which needs to be tested [14].

(4) *Model Training.* In deep learning, the focus on classification has changed from manual features to feature representation, and more attention is paid to expressing features through models. Therefore, it is necessary to build different models and constantly adjust the models to improve the classification effect.

(5) *Model Identification.* There is no doubt about this step. When the model is trained well and the data is input, the classifier can directly output the results.

In order to illustrate the effectiveness of the emotion perception model proposed in this paper, this paper makes a complete experimental comparison. At the same time, due to the strong engineering and reasonable writing expression of this experiment, this paper explains these experiments from the perspective of engineering by introducing the process of how to build a better voice emotion perception system.

Figure 3 shows the construction process of speech emotion recognition system. The process mainly includes sample collection and formatting, alternative speech feature extraction, alternative model construction, alternative feature model combination, feature selection and comparison, model optimization, and comparison [15].

The samples, features, and models to be selected in Figure 3 constitute the alternatives and experimental objects in this paper. The relationship between these alternatives and experimental objects is drawn as shown in Figure 4.

It can be clearly seen from Figure 4 that experiments are needed to determine the following:

- (1) Comparison between feature MFCC and spectrogram
- (2) Comparison of classifier SVM and soft Max
- (3) System model combination relationship

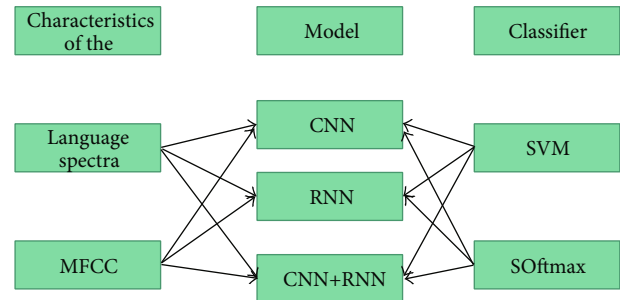


FIGURE 4: Relationship diagram of system alternative models.

- (4) Comparison of fuzzy neural network/cyclic neural network/fuzzy neural network + cyclic neural network
- (5) Effectiveness of the combination of fuzzy neural network/cyclic neural network/fuzzy neural network + cyclic neural network and SVM
- (6) Is the model universal on different data samples: Berlin emo, SAVEE, and CASIA
- (7) The above part is the basic idea and operation process of the emotion perception model proposed in this paper. This is a model proposed based on many experiments. In order to realize the subsequent model, this paper will first establish a traditional recognition model and then use this model to carry out feature comparison experiments and classifier comparison experiments

3.1.2. *Voice Emotion Data Collection and Formatting.* The collection of speech emotion data is the first step in the experiment of speech emotion perception system. This paper mainly collects three public speech emotion databases, CASIA, Berlin emo, and SAVEE. The reason why data formatting is needed is that these databases have different formats for labeling different types of audio data. Data formatting is to preprocess the way of labeling these data and label them according to a unified posture, which is convenient for the unified processing of subsequent programs. These corpora are used by different professional actors to read the same types of texts with different emotions, and

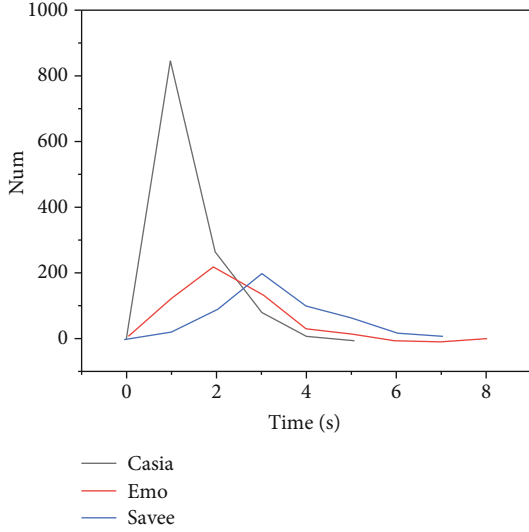


FIGURE 5: Duration distribution of speech emotion database.

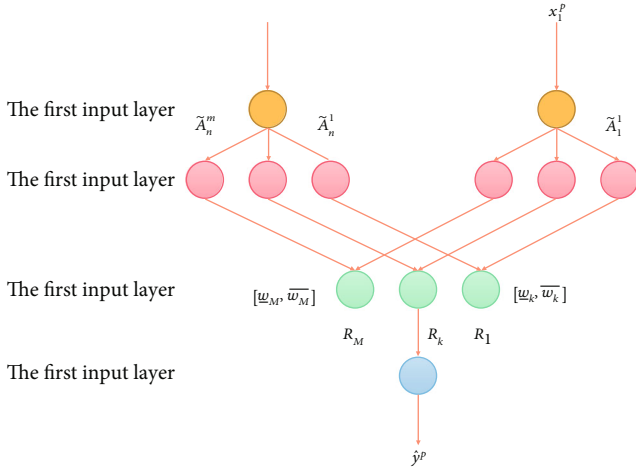


FIGURE 6: Structure diagram of type II fuzzy neural network.

some actors will interpret different versions of the same emotion for the same text [16]. Therefore, the audio file of each emotion database consists of several elements: actor number, text number, emotion number, and version number (there may be). Finally, the length of audio is counted, and the statistical results are shown in Figure 5.

As can be seen from Figure 5, most of CASIA's data length is between 1s and 2s, most of emo's data length is between 2S and 3S, and most of SAVEE's audio length is between 3S and 4S. By analyzing the duration, the duration distribution of the sample is understood, which provides a reference for the duration of subsequent system test recording.

**3.2. Interval Type II Fuzzy Neural Network.** The structure of general interval type II fuzzy neural network is shown in Figure 6.

Assuming that the type 2 fuzzy neural network has  $m$  rules, the  $k$ -th rule  $R^k$  can be expressed as follows.

Each layer of multi-input single output neural network is described as follows.

The first layer is the input layer: It contains  $n$  neurons, representing the input vector:  $x^p = [x_1^p, x_2^p, \dots, x_n^p]$ . The output of the  $j$ -th neuron is the  $j$ -th input variable  $x_j^p$  of the input vector [17].

The second layer is the membership function layer: Each neuron is the membership function of the input variable, the number of neurons is  $n \times M$ , and the output of each neuron represents the  $i$ -th input variable  $x_j^p$  in the rule. If the membership function in  $k$  is a traditional type-I fuzzy membership function, the Gaussian membership function is selected, which can be expressed as

$$N_{ki}(x_i^p : m_{ki}, \sigma_{ki}) = \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - m_{ki}}{\sigma_{ki}} \right)^2 \right]. \quad (1)$$

Either the membership function of Gauss type and type II can pass through the fuzzification center  $m_{ki}$  or the width  $\sigma_{ki}$  indicates that if the fuzzification center  $m_{ki}$  can be past through; thus, the upper and lower limits of the membership function can be expressed as

$$\bar{N}_{ki}(x_i^p : \bar{m}_{ki}, \underline{m}_{ki}, \sigma_{ki}) = \begin{cases} \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - \underline{m}_{ki}}{\sigma_{ki}} \right)^2 \right], & x_i^p \leq \underline{m}_{ki} \\ 1, & \underline{m}_{ki} < x_i^p \leq \bar{m}_{ki} \\ \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - \bar{m}_{ki}}{\sigma_{ki}} \right)^2 \right], & x_i^p > \bar{m}_{ki} \end{cases}, \quad (2)$$

$$\underline{N}_{ki}(x_i^p : \bar{m}_{ki}, \underline{m}_{ki}, \sigma_{ki}) = \begin{cases} \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - \underline{m}_{ki}}{\sigma_{ki}} \right)^2 \right], & x_i^p \leq \frac{(\bar{m}_{ki} + \underline{m}_{ki})}{2} \\ 1, & \underline{m}_{ki} < x_i^p \leq \bar{m}_{ki} \\ \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - \bar{m}_{ki}}{\sigma_{ki}} \right)^2 \right], & x_i^p > \frac{(\bar{m}_{ki} + \underline{m}_{ki})}{2} \end{cases}.$$

If the center  $\sigma_{ki}$  is blurred, the upper and lower limits of the membership function can be expressed as

$$\bar{N}_{ki}(x_i^p : m_{ki}, \bar{\sigma}_{ki}, \underline{\sigma}_{ki}) = \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - m_{ki}}{\bar{\sigma}_{ki}} \right)^2 \right], \quad (3)$$

$$\underline{N}_{ki}(x_i^p : m_{ki}, \bar{\sigma}_{ki}, \underline{\sigma}_{ki}) = \exp \left[ -\frac{1}{2} \left( \frac{x_i^p - m_{ki}}{\underline{\sigma}_{ki}} \right)^2 \right]. \quad (4)$$

Now, choose the second form of Gaussian membership function, namely, Formula (3) and Formula (4).

The third layer is the fuzzy reasoning layer: The number of neurons is  $m$ , and the output of each neuron can be expressed as

$$F_k = [\underline{f}_k, \bar{f}_k]. \quad (5)$$

$$\text{In Formula (5), } \underline{f}_k = \prod_{i=1}^n N_{ki}(x_i^p : m_{ki}, \bar{\sigma}_{ki}, \underline{\sigma}_{ki}) \\ \bar{f}_k = \prod_{i=1}^n \bar{N}_{ki}(x_i^p : m_{ki}, \bar{\sigma}_{ki}, \underline{\sigma}_{ki}).$$

The fourth layer is the output layer: If it is a single output system, the number of neurons is 1, of which the third and fourth layers have a connection parameter  $[\underline{w}_k, \bar{w}_k]$ , and the neuron output can be expressed as

$$\hat{y} = \frac{y_l + y_r}{2}. \quad (6)$$

In Formula (6),  $y_l = \sum_{k=L+1}^L \underline{f}_k \underline{w}_k + \sum_{k=L+1}^M \underline{f}_k \underline{w}_k / \sum_{k=1}^L \underline{f}_k + \sum_{k=L+1}^M \underline{f}_k$ ;  $y_r = \sum_{k=1}^R \bar{f}_k \bar{w}_k + \sum_{k=R+1}^M \bar{f}_k \bar{w}_k / \sum_{k=1}^R \bar{f}_k + \sum_{k=R+1}^M \bar{f}_k$ .

$y_l$  and  $y_r$  are obtained by the KM order reduction algorithm of interval type II fuzzy sets. The specific km order reduction algorithm can be found in the literature.

Fuzzy c-means clustering algorithm is an unsupervised learning algorithm, and its objective function is

$$J_m(U, P) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2. \quad (7) \\ \text{s.t. } U \in M_{fc}$$

In Formula (7),  $m$  is the weighted index, and  $m$  is a real number greater than 1;  $d_{ik}$  is the Euclidean distance between the data point and the cluster center. The result of clustering is to obtain the minimum value of objective function  $J_m(U, P)$ , which meets the following constraints:  $\sum_{i=1}^c u_{ik} = 1$ .

According to the Lagrange multiplier method, the specific implementation steps of FCM clustering algorithm are as follows.

Initialization given the number of cluster categories  $c$ ,  $2 \leq c \leq n$ ,  $n$  is the number of samples, set the given stop threshold  $\epsilon$ , the maximum number of iterations  $C_{\max}$ , initialize the cluster center  $P^0$ , and set the iteration counter  $r=0$  [18].

Step 1: Calculate the fuzzy partition matrix  $U^{(r+1)}$  of the  $R+1$ st iteration with Equation (8).

For  $\forall i, k$ , if  $\exists d_{ik}^r > 0$ , there are

$$u_{ik}^{r+1} = \frac{1}{\sum_{l=1}^c (d_{ik}^r / d_{il}^r)^{2/m-1}}. \quad (8)$$

If for  $\exists i, k$ ,  $d_{ik}^r = 0$ , then there is  $u_{ik}^{r+1} = 1$ , and for  $j \neq k$ ,  $u_{ij}^{r+1} = 0$ .

Step 2: Calculate the cluster center  $P^{(r+1)}$  of the  $r+1$ -th iteration with Equation (9):

$$p_i^{r+1} = \frac{\sum_{k=1}^n (u_{ik}^{r+1})^m x_k}{\sum_{k=1}^n (u_{ik}^{r+1})^m}. \quad (9)$$

In Step 3, if  $\|P^{(r+1)} - P^{(r)}\| < \epsilon$  or the number of iterations is greater than  $C_{\max}$ , the algorithm stops, and the value of the last iteration is the fuzzy partition matrix  $U$  and the cluster center  $P$ . Otherwise, make  $r=r+1$ , and turn to Step 1, where  $\|\bullet\|$  is an appropriate matrix norm.

The parameter training method of interval type II fuzzy neural network proposed in this paper uses BP algorithm, and the optimized objective function is

$$e^p = \frac{1}{2} (\hat{y}^p - y^p)^2. \quad (10)$$

Firstly, the direct order reduction algorithm is used to avoid the problems caused by the common km order reduction algorithm. The output of type II fuzzy neural network is  $\hat{y}^p = y_l + y_r/2$ , where  $y_l = \sum_{k=1}^M \underline{f}_k \underline{w}_k / \sum_{k=1}^M \underline{f}_k$ ;  $y_r = \sum_{k=1}^M \bar{f}_k \bar{w}_k / \sum_{k=1}^M \bar{f}_k$ .

The parameters of the BP algorithm here include: center  $m_{ki}$  and width  $\underline{\sigma}_{ki}, \bar{\sigma}_{ki}$  of Gaussian membership function; And the subsequent parameter  $\underline{w}_k, \bar{w}_k$ , the parameter learning algorithms are

$$\begin{aligned} m_{ki}(t+1) &= m_{ki}(t) - \eta_m \frac{\partial e^p}{\partial m_{ki}}, \\ \underline{w}_k(t+1) &= \underline{w}_k(t) - \eta \frac{\partial e^p}{\partial \underline{w}_k}, \\ \bar{w}_k(t+1) &= \bar{w}_k(t) - \eta_w \frac{\partial e^p}{\partial \bar{w}_k}, \\ \underline{\sigma}_{ki}(t+1) &= \underline{\sigma}_{ki}(t) - \eta_\sigma \frac{\partial e^p}{\partial \underline{\sigma}_{ki}}, \\ \bar{\sigma}_{ki}(t+1) &= \bar{\sigma}_{ki}(t) - \eta_\sigma \frac{\partial e^p}{\partial \bar{\sigma}_{ki}}, \end{aligned} \quad (11)$$

According to the derivation of composite function, the learning algorithms of each parameter are

$$\frac{\partial e^p}{\partial \bar{w}_k} = \frac{\partial e^p}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial y_l} \frac{\partial y_l}{\partial \bar{w}_k} = e^p \frac{1}{2} \frac{\underline{f}_i}{\sum_{k=1}^M \underline{f}_k}, \quad (12)$$

$$\frac{\partial e^p}{\partial \bar{w}_k} = \frac{\partial e^p}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial y_r} \frac{\partial y_r}{\partial \bar{w}_k} = e^p \frac{1}{2} \frac{\bar{f}_i}{\sum_{k=1}^M \bar{f}_k}, \quad (13)$$

$$\frac{\partial e^p}{\partial m_{ki}} = \frac{\partial e^p}{\partial \hat{y}} \left( \frac{\partial \hat{y}}{\partial y_l} \frac{\partial y_l}{\partial m_{ki}} + \frac{\partial \hat{y}}{\partial y_r} \frac{\partial y_r}{\partial m_{ki}} \right), \quad (14)$$

$$\frac{\partial e^p}{\partial \underline{\sigma}_{ki}} = \frac{\partial e^p}{\partial \bar{y}} \left( \frac{\partial \bar{y}}{\partial y_l} \frac{\partial y_l}{\partial \underline{\sigma}_{ki}} \right), \quad (15)$$

$$\frac{\partial e^p}{\partial \bar{\sigma}_{ki}} = \frac{\partial e^p}{\partial \bar{y}} \left( \frac{\partial \bar{y}}{\partial y_r} \frac{\partial y_r}{\partial \bar{\sigma}_{ki}} \right). \quad (16)$$

For  $N$  groups of training input and output data  $D^p(x^p, y^p)$ , ( $p = 1, 2, \dots, N$ ), set the maximum number of iterations  $I$  and cutoff error  $\varepsilon$ .  $m_{ki}$  Learning rate  $\eta_m$ ,  $\underline{\sigma}_{ki}$ ,  $\bar{\sigma}_{ki}$ , learning rate,  $\eta_\sigma$  and subsequent parameter  $\underline{w}_k$ ,  $\bar{w}_k$  learning rate  $\eta_w$ . The interval type II fuzzy neural network algorithm proposed in this paper is described as follows [19]:

(1) Set the number of fuzzy rules  $m$  of type II fuzzy neural network

(2) The results of FCM clustering algorithm are used to initialize the parameters of Gaussian membership function in neural network, that is,  $m_{ki}$  is initialized as the input  $x$  cluster center, and the subsequent parameter  $\bar{w}_k = \bar{w}_k$  of fuzzy rule is initialized as the output  $y$  cluster center

(3) Use Equation (17) to initialize the parameters of Gaussian membership function:

$$\underline{\sigma}_{ki} = \sqrt{\frac{\sum_{p=1}^N u_{pk} (x_i^p - m_{ki})^2}{\sum_{p=1}^N u_{pk}}} \bar{\sigma}_{ki} = 1.2 \underline{\sigma}_{ki}. \quad (17)$$

(4) Set the initial iteration counter  $t = 0$  for each training data  $D^p(x^p, y^p)$ , ( $p = 1, 2, \dots, N$ ); use Equations (12) to (16) to update parameters  $m_{ki}$ ,  $\underline{\sigma}_{ki}$ ,  $\bar{\sigma}_{ki}$ , and  $\underline{w}_k$ ,  $\bar{w}_k$ .

(5) If the sum  $J$  of root square error (RSE) of all training data, as shown in Equation (18), is less than the given cutoff error  $\varepsilon$  or the number of iterations reaches the maximum number of iterations  $I$ , the training process is ended; otherwise,  $t = t + 1$ ; go to Step (3):

$$J = \frac{\sum_{p=1}^N (e^p)^2}{N}. \quad (18)$$

#### 4. Experimental Methods and Results

In the selection of base classifier, this paper selects three groups of eight algorithms based on different principles as the base classifier.

Algorithms based on tree structure: lad tree, rep tree, and cart tree

Rule based algorithms: DTNB (decision tree naive Bayes) and part rule (some rules)

Linear: support vector machine SVM and linear regression LR

Multilayer perceptron: MLP algorithm with neural network structure

Input the training data into the above algorithm for training, and its accuracy is shown in Figure 7.

The above algorithms have been cross verified for 10 times, respectively, and the accuracy of each algorithm is given. It can be seen from the figure that the algorithm based on tree structure obviously has better performance, followed by the rule-based algorithm, and finally the MLP algorithm

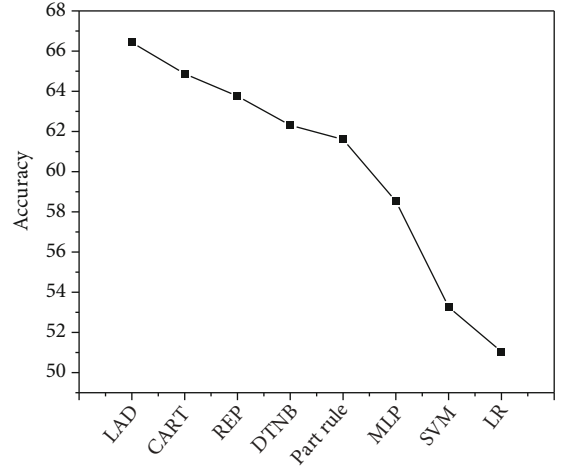


FIGURE 7: Base classifier accuracy rate.

based on neural network, while the accuracy of SVM and logical regression algorithm LR is lower [20].

Within the value range of 2 to 10 for  $K$ , the  $k$ -means algorithm is used to cluster the data samples and calculate the contour coefficient value, and the broken line diagram of contour coefficient and  $K$  value can be obtained, as shown in Figure 8.

As can be seen from the figure, the contour coefficient increases monotonically between the values of  $K$  from 2 to 3 and decreases monotonically between the values of  $K$  from 3 to 10. An inflection point is formed on the value of  $k = 3$ , that is, the maximum value is taken when  $k = 3$ , which means that the optimal cluster number is 3. Three base classifiers are selected for final classifier integration. At the same time, the difference measurement is calculated for the 10 base classifiers, as shown in Figure 9.

Firstly, in the difference measurement, this paper adopts the pairwise difference measurement method, that is, the difference between two classifiers is compared, and then it is summarized and averaged. However, as can be seen from Figure 9, this measurement method often cannot reveal the best difference due to its averaging characteristics. For example, for two classifiers with large differences, adding the third classifier will always reduce the overall average difference measure. Therefore, through experiments, it is found that the average difference measurement method, which summarizes the pairwise difference measurement, often cannot get the expected ideal results, that is, the curve is relatively flat without particularly prominent peaks and troughs. In this case, the different values between algorithms based on the same principle are close and clustered in a similar range [21]. At the same time, the algorithm with better accuracy also has a significant correct recognition rate for the samples misclassified by other classifiers, so the difference value is high. On the contrary, the algorithm with poor accuracy (such as LR and SVM) has a low difference value. To sum up, based on the experimental results (i.e., there is no particularly prominent peak and trough difference data), we choose the algorithm with relatively high difference and the algorithm based on different principles to fuse with each

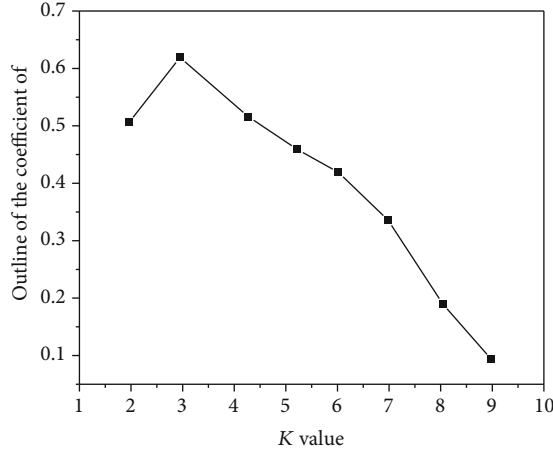


FIGURE 8: Corresponding diagram of contour coefficient and cluster number  $K$ .

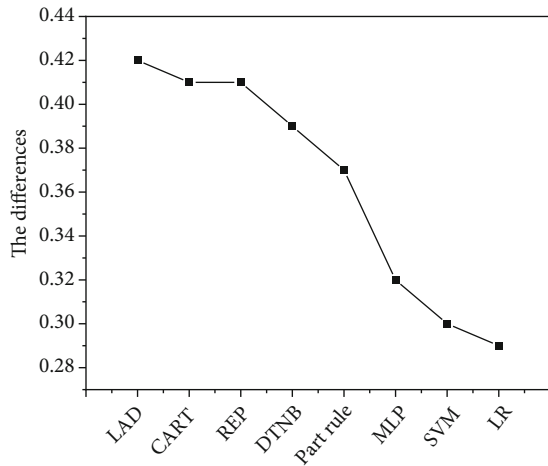


FIGURE 9: Differences of base classifiers.

other, so as to observe the final recognition effect. In this experiment, we choose different algorithms to verify the relationship between the fusion effect of different kinds of classifiers and their differences. At the same time, we use the voting method, the weight method based on recognition accuracy, and the adaptive weight method as the fusion strategy to observe the experimental effect. In this experiment, firstly, three base classifiers with different combinations are selected, and the three fusion strategies are adopted to compare their recognition effects and performance, as shown in Figure 10.

As shown in Figure 11, the analysis shows that the fusion accuracy of voting method is the lowest; using the fusion method based on accuracy, because the weight of the base classifier with the lowest accuracy is the smallest and the discourse power of the base classifier with the highest accuracy is the heaviest, its recognition performance is slightly improved compared with the voting method [22]; after using the adaptive weight method, its performance has been significantly improved compared with the accuracy method, espe-

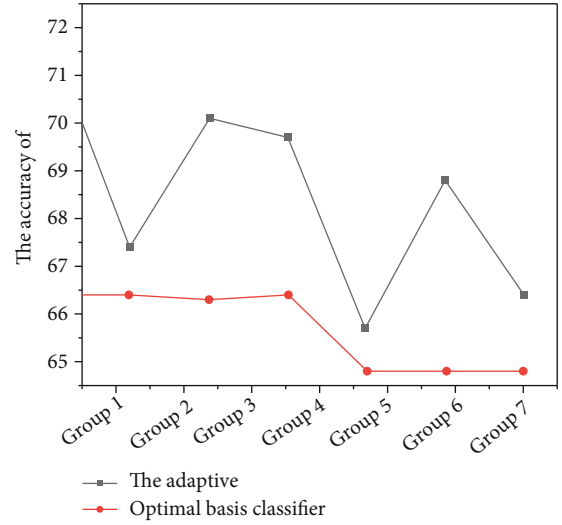


FIGURE 10: Fusion effect of the two classifiers.

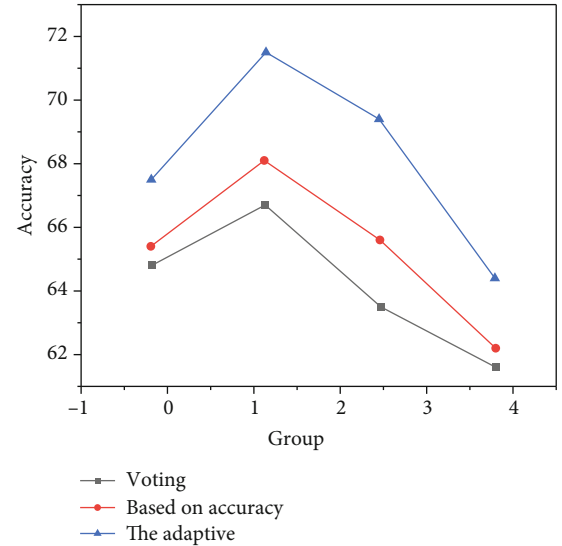


FIGURE 11: Linear graph of difference and fusion strategy effect.

cially in group 2 and group 3. Then, analyze each group in detail.

Group 1 adopts three tree algorithms with the highest accuracy and average difference for fusion. The fusion effect based on accuracy and adaptive weight method is slightly improved compared with the voting method, but the improvement effect is not obvious, indicating that the selection of the base classifier is not the optimal choice.

Group 2 replaces the rep algorithm with the rule-based DTNB algorithm, which has the worst performance in group 1, and obtains better results than group 1 in the final fusion effect. It is the combination with the highest final recognition rate in all groups, and the improvement range becomes larger. In group 3, the cart algorithm in group 2 is replaced by the MLP algorithm based on neural network. The accuracy of the algorithm is low, but the maximum improvement is unexpectedly obtained, and the final recognition rate is



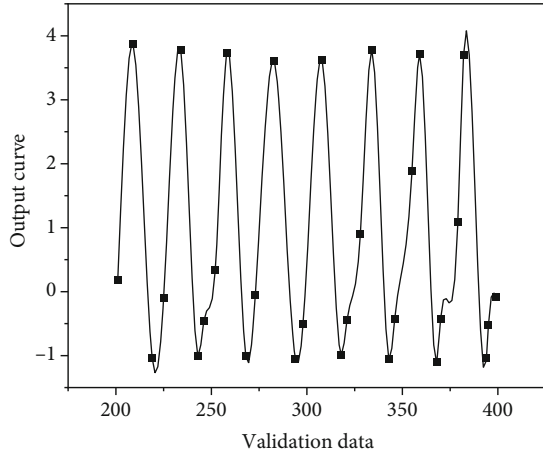


FIGURE 12: Output curve of the algorithm in this paper.

second only to group 2, slightly higher than group 1 [23]. Group 4 then replaced the MLP algorithm with SVM algorithm. Due to the ultra-low accuracy of SVM, its voting method obtained the worst recognition effect, and the performance based on accuracy and adaptive weight method was not ideal. Through the above comparative analysis, we can draw the following conclusion.

In the combination strategy, in most cases, the adaptive weight method can often achieve better decision performance than the optimal subclassifier and has a significant improvement in performance compared with the traditional method of determining the weight based on the accuracy of each classifier. In the difference and base classifier category selection strategy, the average difference measurement method based on pairwise reduces the prominent peak on the curve because of its average, which reduces the reference of its data. From the analysis of groups 1-5, in most cases, the combination of base classifiers based on different principles can achieve better combination improvement effect (e.g., group 3, the classifier integrating three principles has the largest improvement). Group 2 integrates a rule-based DTNB algorithm into the two tree algorithms and achieves the best fusion performance because of the high accuracy of its base classifier. Group 1 uses three tree algorithms, even though its base classifier has the best accuracy, and its improvement is not obvious. As for group 4, the linear SVM algorithm may not be complementary to other algorithms, so its fusion effect is also very limited.

In simulation, the number of fuzzy rules  $M = 4$ . 200 groups of data were used for training, and 200 groups of data were used for verification. Figure 12 shows the fuzzy neural network output curve of the algorithm in this paper, Figure 13 shows the error curve of the verification data, and Figure 14 shows the objective function curve of the number of iterations. In the output curve of Figure 12, the line represents the output  $\hat{y}(k)$  of neural network.

Table 1 shows the performance comparison results based on root mean square error between this algorithm and other identification algorithms. The comparison criterion is the root mean square error of training data and verification data.

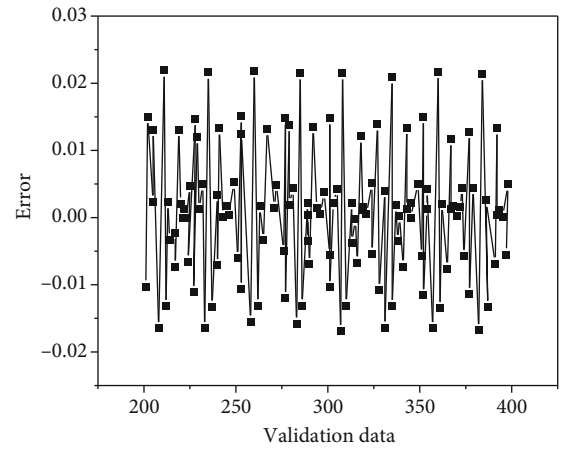


FIGURE 13: Error curve of the algorithm in this paper.

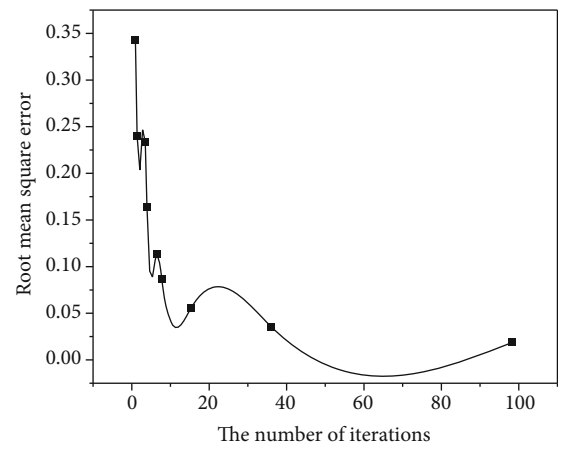


FIGURE 14: Performance curve of the proposed algorithm (root mean square error).

TABLE 1: Performance comparison of different identification algorithms.

Identification method	Rule number	Training data RMSE	Validate data RMSE
BP algorithm	4	0.0126	0.0125
Hybrid algorithm	4	0.128	0.0133
Traditional algorithm	6	0.0109	0.0041
In this paper, algorithm	4	0.0094	0.0087

In simulation, the number of fuzzy rules  $M = 4$ . Figure 15 shows the output curve of the fuzzy neural network for the validation data, Figure 16 shows the error curve of the validation data, and Figure 17 shows the objective function curve of the number of iterations.

Table 2 shows the performance comparison between this algorithm and other identification algorithms. The comparison criterion is the root mean square error of the validation data.

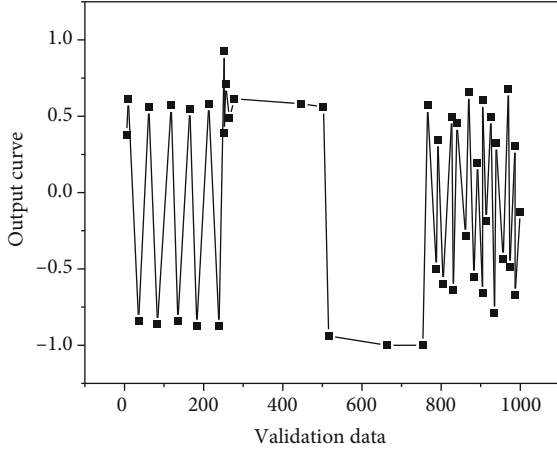


FIGURE 15: Output curve of the algorithm in this paper.

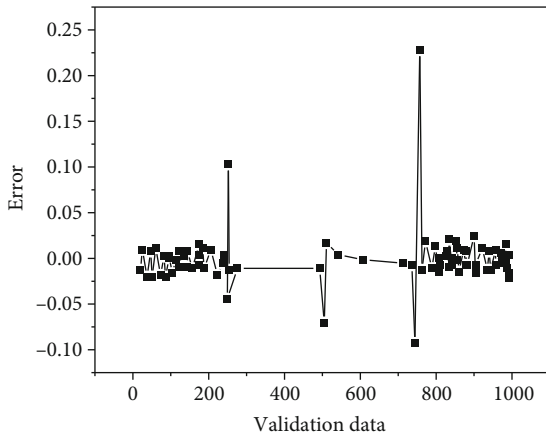


FIGURE 16: Error curve of the algorithm in this paper.

Next, for the test of the system, this paper simulates six emotional states of 50 students in English class, obtains a total of 300 data, and then tests it through the speech emotion recognition system of the mobile terminal. The model used by the mobile terminal is fuzzy neural network. The test results are shown in Figure 18 below. The data in the table represents the accuracy of input speech for emotion recognition.

The test results are not as accurate as expected in previous experiments. The possible reasons are as follows:

- (1) The training sample data is insufficient. In the experiment, the speaker independent and text independent recognition experiments of different speech emotional states are used. In this case, when the training sample size is small enough to include all the speaker's personal voice quality characteristics and all possible texts, it is almost impossible to realize speaker independent and text independent speech emotion recognition on the test set. The deep learning model needs a lot of data to support the training of the model

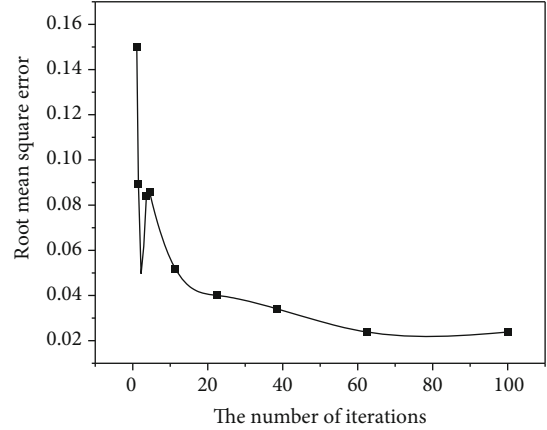


FIGURE 17: Performance curve of the proposed algorithm (root mean square error).

TABLE 2: Performance comparison of different identification algorithms.

Identification method	Rule number	Validate data RMSE
BP algorithm	5	0.0346
Hybrid algorithm	4	0.0397
Traditional algorithm	3	0.0282
In this paper, algorithm	4	0.0136

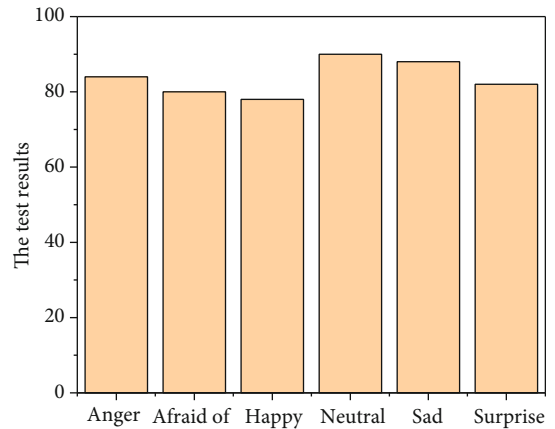


FIGURE 18: Statistical chart of test results.

- (2) The training sample data comes from professional actors, while the testers are ordinary people. In this paper, the main training samples come from the performance simulated emotions of professional actors, and ordinary people may not be the same as the simulated emotions of professional actors without professional training. Therefore, when the training set is inconsistent with the test set, the accuracy of test recognition will be reduced

## 5. Conclusion

Aiming at the complex structure identification of interval type-2 fuzzy neural network and the dependence of BP

parameters on initial values, an interval type-2 fuzzy neural network identification algorithm based on FCM clustering algorithm is proposed. Firstly, the center and membership degree of the training data are obtained by FCM clustering algorithm, and the initial values are assigned to the center and width of the Gaussian membership function and the subsequent parameters of the type 2 fuzzy system. Secondly, a direct type-2 fuzzy set reduction algorithm is used to simplify the process of training network parameters by BP algorithm. The simulation results show that the proposed algorithm has the advantages of simple learning process, good identification accuracy, and fast convergence speed.

## Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no competing interests.

## Acknowledgments

This work is supported by the Shenyang Urban Construction University.

## References

- [1] X. Chen, "Simulation of english speech emotion recognition based on transfer learning and cnn neural network," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 2, pp. 2349–2360, 2021.
- [2] G. Khodabandelou and M. M. Ebadzadeh, "Fuzzy neural network with support vector-based learning for classification and regression," *Soft Computing*, vol. 23, no. 23, pp. 12153–12168, 2019.
- [3] B. Hong, Y. Zhang, Anxi College of Tea Science, Fujian Agriculture and Forestry University, China, and Fine Art and Design College, Quanzhou Normal University, China, "Research on the influence of attention and emotion of tea drinkers based on artificial neural network," *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 3423–3434, 2021.
- [4] A. Nasuha, F. Arifin, A. S. Priambodo, N. Setiawan, and N. Ahwan, "Real time emotion classification based on convolution neural network and facial feature," *Journal of Physics Conference Series*, vol. 1737, no. 1, article 012008, 2021.
- [5] S. Kahaki, M. J. Nordin, N. S. Ahmad, M. Arzoky, and W. Ismail, "Deep convolutional neural network designed for age assessment based on orthopantomography data," *Neural Computing and Applications*, vol. 32, no. 13, pp. 9357–9368, 2020.
- [6] S. Fatahi, "An experimental study on an adaptive e-learning environment based on learner's personality and emotion," *Education and Information Technologies*, vol. 24, no. 4, pp. 2225–2241, 2019.
- [7] D. N. Kanya, P. S. Rani, D. S. Geetha, D. M. Rajkumar, and G. Sandhiya, "An efficient damage relief system based on image processing and deep learning techniques," *Revista Gestão Inovação e Tecnologias*, vol. 11, no. 2, pp. 2124–2131, 2021.
- [8] R. A. Naqvi, M. Arsalan, A. Rehman, A. U. Rehman, and A. Paul, "Deep learning-based drivers emotion classification system in time series data for remote applications," *Remote Sensing*, vol. 12, no. 3, p. 587, 2020.
- [9] M. S. Soria, V. Maldonado, G. D. C. Garcia, K. Patio, and O. Camacho, "Fuzzy control of temperature on SACI based on the emotion recognition," *Adv. Intell. Syst. Comput.*, vol. 1273, no. 1, pp. 94–106, 2020.
- [10] J. R. Jebadass and P. Balasubramaniam, "Low light enhancement algorithm for color images using intuitionistic fuzzy sets with histogram equalization," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8093–8106, 2022.
- [11] Y. Cui, S. Wang, and R. Zhao, "Machine learning-based student emotion recognition for business english class," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 12, p. 94, 2021.
- [12] N. Sabri, "Student emotion estimation based on facial application in e-learning during covid-19 pandemic," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 1.4, pp. 576–582, 2020.
- [13] V. Sreenivas, V. Namdeo, and E. V. Kumar, "Group based emotion recognition from video sequence with hybrid optimization based recurrent fuzzy neural network," *Journal Of Big Data*, vol. 7, no. 1, p. 56, 2020.
- [14] G. Zhou, J. Zhang, X. Liu, and Z. Han, "Study on reliability of disc brake system for mine hoist based on fuzzy dynamic fault tree," *Meitan Xuebao/Journal of the China Coal Society*, vol. 44, no. 2, pp. 639–646, 2019.
- [15] S. Sutiah, I. A. Zuhriyah, and Z. Arifin, "Expert system on core competency detection and student achievement based on fuzzy topsis method," *IOP Conference Series: Materials Science and Engineering*, vol. 846, no. 1, p. 012026, 2020.
- [16] A. Moria, "The keys to learning for university students with disabilities: motivation, emotion and faculty-student relationships," *PLoS One*, vol. 14, no. 5, p. e0215249, 2019.
- [17] R. Lumbrears and W. H. Rupley, "Educational experiences of ell educators: searching for instructional insights to promote ell students' learning," *Educational Research for Policy and Practice*, vol. 18, no. 1, pp. 17–38, 2019.
- [18] Y. Du, "Study on cultivating college students' english autonomous learning ability under the flipped classroom model," *English Language Teaching*, vol. 13, no. 6, p. 13, 2020.
- [19] S. A. Salmee and M. M. Arif, "A study on the use of humour in motivating students to learn english," *Asian Journal of University Education*, vol. 15, no. 3, p. 257, 2020.
- [20] N. M. Krishna, K. Sekaran, A. Venkata, N. Vamsi, and R. Damaevius, "An efficient mixture model approach in brain-machine interface systems for extracting the psychological status of mentally impaired persons using eeg signals," *IEEE Access*, vol. 7, no. 1, pp. 77905–77914, 2019.
- [21] M. A. Hasnul, N. Aziz, S. Alelyani, M. Mohana, and A. Aziz, "Electrocardiogram-based emotion recognition systems and their applications in healthcare—a review," *Sensors*, vol. 21, no. 15, p. 5015, 2021.
- [22] S. Anjani and S. Ch, "An efficient facial emotion recognition system using novel deep learning neural network-regression activation classifier," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 1–26, 2021.
- [23] W. C. Fang, K. Y. Wang, N. Fahier, Y. L. Ho, and Y. D. Huang, "Development and validation of an eeg-based real-time emotion recognition system using edge ai computing platform with convolutional neural network system-on-chip design," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 4, pp. 645–657, 2019.