Research Article

Emotion Recognition Based on Type-2 Recurrent Wavelet Fuzzy Brain Emotion Learning Network Model

Di Li¹ and Xiangjian Chen²

¹China Shipbuilding Industry Corporation 723, Yangzhou, China
²Yangzhou University, Yangzhou, China

Correspondence should be addressed to Xiangjian Chen; yz_cxj831209@163.com

Received 22 March 2021; Revised 27 June 2021; Accepted 24 July 2021; Published 9 August 2021

Copyright © 2021 Di Li and Xiangjian Chen. 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.

Emotion recognition plays a crucial role in human-robot emotional interaction applications, and the brain emotional learning model is one of several emotion recognition methods, but the learning rules of original brain emotional learning model play poor adaptation and do not work very well. In fact, existing facial emotion recognition methods do not have high accuracy and are not sufficiently practical in real-time applications. In order to solve this problem, this paper introduces an optimal model, which merges interval type-2 recurrent wavelet fuzzy system and brain emotional learning network for emotion recognition. The proposed model takes advantage of type-2 recurrent wavelet fuzzy theory and brain emotional neural network. There are no rules initially, and then the structure and parameters of model are tuning online simultaneously by the gradient approach and Lyapunov function. The system input data streams are directly imported into the neural network through a type-2 recurrent wavelet fuzzy inference system; then, the results are subsequently piped into sensory and emotional channels which jointly produce the final outputs of the network. The proposed model could reduce the uncertainty in terms of vagueness by using type-2 recurrent wavelet fuzzy theory and removing noise samples. Finally, the superior performance of the proposed method is demonstrated by its comparison with some emotion recognition methods on five emotion databases.

1. Introduction

Emotion recognition is one of the most effective methods to obtain human information to improve the human-robot interaction, which includes body language emotion [1], facial expression recognition [2], and speech emotion recognition [3]. So it is necessary to recognize humans’ emotional states as accurate as possible for robot.

One important part of providing effective and natural interaction between human and computers is to enable computers to understand the emotional states expressed by the human subjects. Rahul et al. [4] proposed ont method to put this theory’s practical applications. Also, Cen et al. [5] indicated a new measure of authentic auditory emotion recognition, and put this application to patients with schizophrenia. Bhandari and Pal [6] aim to check if an explicit use of edges can help in emotion recognition from images using convolutional neural network (CNN).

For the speech emotion recognition, speaker-dependent (SD) and speaker-independent (SI) have been introduced in recognition processing [7, 8], in which the support vector machine [8], neural network [9], deep conventional neural networks, and deep brief network have been used for speech emotion recognition [10–13]. For the facial expression recognition, which has been successfully applied in many application scenarios, such as human-computer interaction systems and games, the machine could interpret user emotions through emoticon recognition, which makes the algorithm more intelligent and humanized. Single-tag learning or multitag learning paradigms have been used for facial expression recognition. Yang et al. [14] expanded the conditional probability neural network to a fuzzy form, which is used to predict the emotions expressed by the
picture and obtain better performance than the original conditional probability neural network. There are several research studies focused on the conventional neural network for FER [15–18]. Hu et al. [19] combined fuzzy rough set theory with conventional neural network for FER and so on.

However, there exist two problems: firstly, these neural networks mentioned above have large computation and complexity; for this problem, brain emotional learning model is presented in [20, 21], which has lower computation and complexity compared with the other neural networks, but the original BEL cannot achieve desirable performance for speech emotion recognition, so the fuzzy logic system was introduced in the BEL model to improve the recognition accuracy. The second problem is that the fuzzy logic systems used in the above references are all based on type-1, which does not handle the uncertainties more flexibly. In order to solve this problem, we choose type-2 as the better scheme, which has been applied in various applications [22, 23].

A fuzzy wavelet neural network combines neural networks and wavelet theory with fuzzy logic. Several researchers have used fuzzy wavelet neural network for solving signal processing and control problems. In [24] the wavelet network model of a fuzzy inference system is proposed. A fuzzy wavelet neural network structure constructed on the base of a set of fuzzy rules is proposed in [25] and used for the approximation nonlinear functions. Other wavelet based approaches include fuzzy wavelet neural network structures developed for system identification and control [26] and prediction of electricity consumption [27]. In [28] the combination of wavelet technology and type-2 fuzzy system is proposed. A type-2 fuzzy wavelet neural system is proposed for the prediction of stock market prices in [29], in which the convergence and learning of the proposed system are not discussed. In [30], a novel, type-2 fuzzy wavelet neural network (type-2 FWNN) structure combines the advantages of type-2 fuzzy systems and wavelet neural networks for identification and control of nonlinear uncertain systems by the use of clustering and gradient techniques for the updating of the parameters.

Furthermore, the combination of type-2 fuzzy logic sets and brain emotional learning network is proposed in [31, 32], in which the authors used this model to control robots. In [33], the wavelet fuzzy brain emotional learning network is used to control MIMO unconcern system. In [34], the type-2 recurrent fuzzy brain emotional learning network is used to achieve adaptive filter for the active noise cancellation. In [35], the data-driven control method is combined with the interval type-2 intuition fuzzy brain emotional learning network for the multiple degree-of-freedom rehabilitation robot.

According to the above analysis, this paper adopts interval type-2 recurrent wavelet fuzzy logic system combined with brain emotional learning network to construct one model for emotion recognition. The main contributions of this paper include the following: (1) the novel self-organizing type-2 recurrent wavelet fuzzy brain emotional learning network model is proposed firstly; (2) the parameters can be tuned online by adaptive laws; (3) the structure of interval type-2 recurrent wavelet fuzzy brain emotional learning network can be constructed automatically from the empty initial rule; (4) numerical simulations have been made to demonstrate the performance of the proposed method for emotion recognition.

This paper is organized as follows: Section 2 presents the structure of novel interval type-2 recurrent wavelet fuzzy brain emotional learning network model, the parameters learning, and self-organizing structure learning algorithms. Section 3 shows the simulation results. And finally, the conclusions are given in Section 4.

2. Framework of the IT2RWFBELN

The IT2RWFBELN is composed of two parts: one of them is amygdala network, which is suitable for emotional judgment, and the other one is orbitofrontal cortex network, which is responsible for emotional control. The fuzzy inference part of IT2RWFBELN adopts interval type-2 recurrent wavelet system; then these two parts can be described as

$$\text{if } x_1 = \mu_{11}^a \text{ and } \ldots \text{ and } x_i = \mu_{ij}^a \text{ and } \ldots \text{ and } x_n = \mu_{nj}^a \text{ then } y_k^a = \omega_{jk},$$

$$\text{if } x_1 = \mu_{11}^a \text{ and } \ldots \text{ and } x_i = \mu_{ij}^a \text{ and } \ldots \text{ and } x_n = \mu_{nj}^a \text{ then } y_k^a = \nu_{jk},$$

where the membership function grades and the weights for the two network could be described as $\mu_{ij}^o = [\mu_{ij1}, \mu_{ij2}]$ and $\mu_{ij}^o = [\mu_{ij1}, \mu_{ij2}]$ and $\omega_{jk} = [\omega_{j1}, \omega_{j2}]$ and $\nu_{jk} = [\nu_{j1}, \nu_{j2}]$, respectively.

2.1. Structure of IT2RWFBELN

Figure 1 shows the structure of IT2RWFBELN, which includes the amygdala network, the orbitofrontal cortex network, and the interval type-2 recurrent wavelet fuzzy sets. The proposed IT2RWFBELN is constructed with six layers, which includes an input layer, a MF layer, a spatial firing layer, a weight memory layer, a summarily layer, and an output layer:

(1) Input space: these nodes in this space are given as $X = [x_1, x_2, \ldots, x_n]$, where $n_i$ represents the number of the input signals, and all data from this layer will be transmit to the next space without any computation.

(2) MF space: in this layer, Gaussian activation function is adopted to finish the fuzzification by using interval type-2 wavelet membership functions, which are adopted as the basis function. Then the approximation ability by using wavelet functions than triangle or Gaussian basis functions, so the learning speed could be increased. Furthermore, the recurrent term with previous information is inserted in this layer; therefore, the network performance can be further improved. The wavelet membership functions can be represented as
\[ p_{ik} = \begin{cases} 
\left( \frac{x_{ri} - c_{ik}^l}{2\sigma_{ik}} \right) \left( \frac{x_{ri} - c_{ik}^l}{2\sigma_{ik}} \right)^2, & x_{ri}(t) \leq c_{ik}^l, \\
1, & c_{ik}^l \leq x_{ri}(t) \leq c_{ik}^r, \\
\left( \frac{x_{ri} - c_{ik}^r}{2\sigma_{ik}} \right) \left( \frac{x_{ri} - c_{ik}^r}{2\sigma_{ik}} \right)^2, & x_{ri}(t) \geq c_{ik}^r, 
\end{cases} \]

where the parameters \( c_{ik}^l, c_{ik}^r, \) and \( \sigma_{ik} \) represent the center and variance of the fuzzy rules of type-2 wavelet membership functions. So the output of \( i \)th input feature and \( k \)th rule can be represented by an interval MF \( [\mu_{ik}, \bar{\mu}_{ik}] \), and \( \bar{x}_{ri} = [\underline{x}_{ri}, \overline{x}_{ri}] \) represents the recurrent inputs, which could be given as

**Figure 1: Structure of IT2RWFBE LN.**
(3) Spatial firing layer: the rule in this layer is composed of both upper and lower firing strength of MFs as well as non-MFs. The firing strength of interval type-2 fuzzy rule is interval and can be calculated as follows:

\[
F_k = [f_k^u, f_k^l],
\]

\[
f_k = \sum_{j=1}^{m} \mu_{jk}, \quad k = 1, 2, \ldots, M, \tag{5}
\]

\[
\bar{f}_k = \sum_{j=1}^{m} \bar{\mu}_{jk},
\]

where \(\bar{f}_k\) denotes the intervals of the firing strength of MFs, the \(m\), and \(M\) represent number of input signals and fuzzy rules, respectively.

(4) Weight memory space: this layer contains two memory spaces: they are amygdala memory and an orbitofrontal memory, which are interval values because the firing space is interval. So the weight of orbitofrontal memory, which are interval values and fuzzy rules, respectively.

(5) Defuzzification space: the output of this defuzzification space is calculated by the output of the firing space and the weight space. So the left- and right-most point values for amygdala and orbitofrontal network output are as follows:

\[
y'^a_k = \frac{\alpha'_k + \alpha''_k}{2} = \frac{1}{2} \left( \frac{\sum_{j=1}^{n} (f_{jk} \mu_{jk})}{\sum_{j=1}^{n} f_{jk}} + \frac{\sum_{j=1}^{n} (\bar{f}_{jk} \bar{\mu}_{jk})}{\sum_{j=1}^{n} \bar{f}_{jk}} \right),
\]

\[
y'^o_k = \frac{\alpha'_k + \alpha''_k}{2} = \frac{1}{2} \left( \frac{\sum_{j=1}^{n} (f_{jk} \mu_{jk})}{\sum_{j=1}^{n} f_{jk}} + \frac{\sum_{j=1}^{n} (\bar{f}_{jk} \bar{\mu}_{jk})}{\sum_{j=1}^{n} \bar{f}_{jk}} \right). \tag{8}
\]

(6) Output space: the output of defuzzification space is an interval value, so the average operations are used to obtain the amygdala network and e orbitofrontal cortex network outputs are

\[
j_{IT2RWFBELN} = y'^a_k - y'^o_k. \tag{9}
\]

2.2. Self-Organization of IT2RWFBELN. According to the theory of nuclear fuzzy rough set, fuzzy upper approximation is used to indicate the possibility of a sample belonging to a certain emotion, and fuzzy lower approximation is used to indicate the inevitability of a sample belonging to a certain emotion. In general, the classification decision of a sample has less uncertainty for the strong ability to distinguish between feature spaces, which means that the closer the upper fuzzy approximation is to the lower fuzzy approximation, the better.

Achieving the better structure of the IT2RWFBELN requires deciding the appropriate rules by using the corresponding self-organizing algorithm. If the number of rules is large, this will cost long time for computational loading, and also if the numbers are small, this cannot reflect all the cases especially with the data with large ranges. Initially, there are no rules and MFs in the first space; when the first input data stream comes, the first MF then will be created. Then the self-organizing algorithm will be used to determine whether to generate new rules and MFs or to delete inappropriate rules and MFs. In this paper, the (interval type-2 fuzzy c-means) IT2FCM is used here to choose the cluster centers of the membership functions for fuzzy rules of the RT2WFNN. The IT2FCM [32] is an iterative optimization algorithm that minimizes the objective function as

\[
J_m (U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} \mu_{ik} (m) d^2_{ik},
\]

\[
\hat{\mu}_k (k) = \max \left( \frac{1}{\sum_{j=1}^{c} (d_{ik}/d_{jk})^{2/(m-1)}} \right),
\]

\[
\hat{\nu}_k (k) = \min \left( \frac{1}{\sum_{j=1}^{c} (d_{ik}/d_{jk})^{2/(m-1)}} \right),
\]

\[
\left( \begin{array}{c}
\mu_k (k) \\
\nu_k (k)
\end{array} \right) = \left( \begin{array}{c}
\frac{1}{\sum_{j=1}^{c} (d_{ik}/d_{jk})^{2/(m-1)}} \\
\frac{1}{\sum_{j=1}^{c} (d_{ik}/d_{jk})^{2/(m-1)}}
\end{array} \right).
\]
where $d_{ik}^t = \|x_k - v_i\|$ denotes the distance between the cluster centers $v_i$ and an input pattern $x_k$. The main steps of the IT2FCM can be shown as follows:

1. Set the fuzzifiers and the number $c$ of the cluster prototypes, and initialize the cluster center $V$ using GA algorithm.
2. Calculate the distance between the cluster centers $v_i$ and input pattern $x_k$; the lower and upper partition functions can be calculated by (2).
3. Update the cluster center $V'$, and the interval type-1 fuzzy set $[c_l,c_u]$ is obtained during the iterative process and optimal improved EKM algorithm, which is adopted to estimate both ends of the interval fuzzy set.
4. The new cluster center $V'$ is updated by a defuzzification method as $V' = (c_l + c_u)/2$; go to the next step; otherwise, set $V = V'$.
5. Finally, the type-reduction of the type-2 fuzzy partition matrix is set as $\mu_{ik} = (\mu^l_{ik} + \mu^r_{ik})/2$.

The output of IT2FCM algorithm is an interval type-2 FS that cannot be transformed to crisp set by defuzzifier directly. Hence, the type-reduction process is needed. The aim of type-reduction is to compute the centroid of a type-2 fuzzy set. At present, iterative Karnik–Mendel (KM) algorithm and the enhanced Karnik–Mendel (EKM) algorithm can compute the centroid of an interval type-2 fuzzy set efficiently. The improved EKM is used here, which could change the initialization conditions of switch points and improve the searching method for switch points.

2.3. Parameters Learning Algorithm of IT2RWFBELN

Defining one Lyapunov cost function as $V(E(t)) = (1/2)E(t)$, then $V(E(t)) = E(t)\dot{E}(t)$ and using the gradient descent method, the online tuning laws for the parameters of IT2RWFBELN are given as

$$
\dot{v}_{jk}(t + 1) = \ddot{v}_{jk}(t) - \eta^o \frac{\partial E(t)}{\partial \dot{v}_{jk}},
$$

$$
\dot{\omega}_{jk}(t + 1) = \ddot{\omega}_{jk}(t) - \eta^u \frac{\partial E(t)}{\partial \dot{\omega}_{jk}}.
$$

where $\eta^o$ and $\eta^u$ represent the learning rates for updating the weight of orbitofrontal cortex and an amygdala network, respectively, and $\eta^o$, $\eta^u$, $\eta^a$, and $\eta^b$ represent the learning rates for updating the means and the variances of type-2 wavelet MFs, respectively. $\eta^a$ and $\eta^b$ represent the learning rates for updating the IF-indices, indicating the hesitation level of recurring $t$. By applying the chain rule for the derivation of above terms, then we have

$$
\frac{\partial E(k)}{\partial \omega_{jk}} = \frac{\partial E}{\partial \dot{y}_{IT2RWFBELN}} \frac{\partial \dot{y}_{IT2RWFBELN}}{\partial y_k} \frac{\partial y_k}{\partial \omega_{jk}} \frac{\partial \omega_{jk}}{\partial \omega_{jk}}
$$

$$
= \frac{1}{2} e(t) \frac{\sum_{j=1}^{n} f^a_{jk}}{\sum_{j=1}^{n} f^a_{jk}},
$$

$$
\frac{\partial E(k)}{\partial \omega_{jk}} = \frac{\partial E}{\partial \dot{y}_{IT2RWFBELN}} \frac{\partial \dot{y}_{IT2RWFBELN}}{\partial y_k} \frac{\partial y_k}{\partial \omega_{jk}} \frac{\partial \omega_{jk}}{\partial \omega_{jk}}
$$

where $\eta^a$ and $\eta^u$ represent the learning rates for updating the weight of orbitofrontal cortex and an amygdala network, respectively, and $\eta^a$, $\eta^u$, $\eta^a$, and $\eta^b$ represent the learning rates for updating the means and the variances of type-2 wavelet MFs, respectively. $\eta^a$ and $\eta^b$ represent the learning rates for updating the IF-indices, indicating the hesitation level of recurring $t$. By applying the chain rule for the derivation of above terms, then we have
\[ \frac{\partial E(k)}{\partial \tilde{v}_{jk}} = \frac{\partial E(z)}{\partial \tilde{y}_{IT2RFWBELN}} \frac{\partial \tilde{y}_{IT2RFWBELN}}{\partial y_k} \frac{\partial y_k}{\partial \tilde{v}_{jk}} = \frac{1}{2} e(t) \frac{\tilde{f}_{jk}}{\sum_{j=1}^{n_j} \tilde{f}_{jk}} \]

\[ \frac{\partial E(k)}{\partial \tilde{v}_{jk}} = \frac{\partial E(z)}{\partial \tilde{y}_{IT2RFWBELN}} \frac{\partial \tilde{y}_{IT2RFWBELN}}{\partial y_k} \frac{\partial y_k}{\partial \tilde{v}_{jk}} = \frac{1}{2} e(t) \frac{\tilde{f}_{jk}}{\sum_{j=1}^{n_j} \tilde{f}_{jk}} \]

\[ \frac{\partial E(k)}{\partial \tilde{v}_{jk}} = \frac{\partial E(z)}{\partial \tilde{y}_{IT2RFWBELN}} \frac{\partial \tilde{y}_{IT2RFWBELN}}{\partial y_k} \frac{\partial y_k}{\partial \tilde{v}_{jk}} = \frac{1}{2} e(t) \frac{\tilde{f}_{jk}}{\sum_{j=1}^{n_j} \tilde{f}_{jk}} \]

2.4. Convergence Analyses

**Theorem 1.** Let \( \lambda_m \) and \( \lambda_o \) be the learning rates for the parameter of FBELC \( m_{ij} \) and \( \sigma_{ij} \), respectively. Then, the stable convergence is guaranteed if \( \lambda_m \) and \( \lambda_o \) are chosen as

\[ 0 < \lambda_m < \frac{2}{y_m} \]

\[ 0 < \lambda_o < \frac{2}{y_o} \]
Proof. A Lyapunov function is selected as
\[ L(k) = \frac{1}{2} e^2(k). \] (15)

The change of the Lyapunov function is
\[ \Delta L(k) = L(k+1) - L(k) = \frac{1}{2} \left[ e^2(k+1) - e^2(k) \right]. \] (16)

The predicted error can be represented by
\[ e(k+1) = e(k) + \Delta e(k) \equiv e(k) + \left[ \frac{\partial e(k)}{\partial m_{ij}} \right] \Delta m_{ij}, \] (17)

where \( \Delta m_{ij} \) denotes the change of \( m_{ij} \).

Using (11), it is obtained that
\[ \frac{\partial e(k)}{\partial m_{ij}} = -y_o(1 - y_o) \cdot (A - O) \cdot S_{ij} \cdot \frac{2(I_i - m_{ij})}{\sigma^2_{ij}}. \] (18)

Substituting (11) and (18) into (17) yields
\[ e(k+1) = e(k) - z_{my} \cdot (1 - y_o) \cdot (A - O) \cdot S_{ij} \cdot \frac{2(I_i - m_{ij})}{\sigma^2_{ij}} + e(k) - \lambda_m e(k) y_m^2. \] (19)

Thus,
\[ \Delta L(k) = \frac{1}{2} \cdot \left[ e^2(k+1) - e^2(k) \right] = \frac{1}{2} e^2(k) \left( 1 - \lambda_m y_m^2 \right)^2 - 1. \] (20)

If \( \lambda_m \) is chosen as (13), \( \Delta L(k) \) in (20) is less than 0. Therefore, the Lyapunov stability of \( L(k) > 0 \) and \( L(k) < 0 \) is guaranteed. The proof for \( \lambda_m \) can be derived in similar method, which should be chosen as (14). This completes the proof. \( \square \)

3. Experiments and validation

This paper implements the IT2RWFBLN model on Python 2.7. The initialization parameters of the IT2RWFBLN are derived from the ImageNet data set. The calculation of the model is done on the GPU. They were based on Windows OS i7-8700 CPU with a clock speed of 3.20 GHz, RAM of 8 GB, and GTX 1070 GPU.

In this part, in order to verify the performance of the IT2RWFBLN in the facial expression recognition, two groups of experiments have been conducted on five public expression data sets (Jaffe, BU-3DFE, CASIA, SAVEE, and FAU). The first one is tested on Chinese corpus of CASIA [36], English corpus of SAVEE [37], and FAU emotion corpus [38], in which both speaker-dependent (SD) and speaker-independent (SI) speech emotion recognition are performed; the second one is tested on Jaffe [39] and BU-3DFE [40], in which six different metrics are introduced to test. To verify the robustness and suitability of the proposed model, some conventional methods are used for comparison.

3.1. Emotion Databases. The five data sets include CASIA database, SAVEE database, FAU database, JAFFE database, and BU-3DFE database, which are described as CASIA Chinese emotion corpus provided with the data set including 300 emotional short utterances, which contain six basic emotions: surprise, happy, sad, angry, fear, and neutral. SAVEE data set is conducted by seven different emotions on recording from four males and seven basic emotions: surprise, happy, sad, angry, fear, disgust, and neutral. The FAU database is conducted by recording 30 females and 21 males and contains five emotional states: angry, emphatic, positive, neutral, and rest. The JAFFE database is conducted by 10 different people with 213 facial images, which include six expressions: anger, disgust, fear, happy, sad, surprise, and neutral. The BU-3DFE multiuser facial expression database is conducted by 56 females and 44 males, which includes six facial expressions: anger, disgust, fear, happiness, sadness, and surprise.

3.2. Experiments on JAFEE and BU-3DFE Databases. The first experiments focused on emotional recognition for JAFFE and BU-3DFE database. BU-3DFE database has six emotions: anger, disgust, fear, happiness, sadness, and surprise. The JAFFE database has seven emotions: fear, happy, sadness, anger, disgust, surprise, and neutral. Furthermore, six different metrics including chebychev, kldist, cosine, canberra, clark, and intersection are used here to validate the effectiveness of the proposed method by comparing with other methods, including Local Binary Patterns (LBP), “AN(FC1)” and “AN(FC2),” which are the first two layers of fully connected layer output of Alex Net pretrained, “AN-FT(MSE)” and “AN-FT(KL),” respectively, and represent the fuzzy classification results obtained by Alex Net using the mean square error and KL divergence as loss functions and fuzzy rough conventional neural network (FRCNN).

Table 1 gives the experimental results for the BU-3DFE data set. The features are obtained in the type-2 fuzzy sets convolutional neural network training task, and the fuzzy classification results are obtained in the fuzzy expression recognition task based on Algorithm Adaption k-Nearest-Neighbors classification. It can be considered that type-2 fuzzy convolutional neural network effectively learns
relevant knowledge from fuzzy multilabels. Therefore, the fuzzy classification effect is better than “AN(FC1)” and “AN(FC2). As shown in Table 2, using Algorithm Adaption k-Nearest-Neighbors as the fuzzy classification algorithm, under the six metrics, emotional learning networks extract features are better than “LBP.” Compared with the other features, type-2 fuzzy wavelet emotional learning neural network achieves good fuzzy classification accuracy under various indicators. Since the performance of the Algorithm Adaption k-Nearest-Neighbors algorithm depends on the distinguishability of the feature space, it can be considered that, compared to other algorithms, type-2 fuzzy wavelet neural network model maps the original picture to a space more suitable for distinguishing facial expressions, that is, a face picture with a similar expression.

To evaluate the performance of the proposed algorithm more efficiently, Tables 3 and 4 list the accuracy of proposed model and other algorithms on the JAFEE and BU-3DFE databases. From the results, we can see that accuracy of our algorithm is superior to most of the other advanced algorithms.

3.3. Experiments on CASIA, SAVEE, and FAU Databases. To verify the performance of the IT2RWFBELN model for speech emotion recognition, the performance of the proposed model is compared with conventional brain emotional learning (BEL), support vector machine (SVM), extreme learning machine (ELM), genetic algorithm-brain emotional learning (GA-BEL), BELFIS, and BELBLA methods that have been made on CASIA, SAVEE, and FAU databases, as presented in Tables 5–7.

Table 1: Comparison results on BU-3DFE database.

| Method    | Canberra | Chebyshev | Clark | Cosine | Intersection | Kldist |
|-----------|----------|-----------|-------|--------|--------------|--------|
| LBP       | 0.8256   | 0.1265    | 0.3975| 0.9229 | 0.8496       | 0.0788 |
| AN-FT(MSE)| 0.5822   | 0.0883    | 0.2827| 0.9606 | 0.8948       | 0.0392 |
| AN-FT(KL) | 0.9811   | 0.1641    | 0.4962| 0.8638 | 0.8116       | 0.1463 |
| AN(FC1)   | 0.5102   | 0.0787    | 0.2486| 0.9692 | 0.9078       | 0.0303 |
| AN(FC2)   | 0.5264   | 0.0814    | 0.2566| 0.9674 | 0.9047       | 0.0321 |
| FRCCNN    | 0.4240   | 0.0554    | 0.1858| 0.9778 | 0.9237       | 0.0132 |
| IT2RWFBELN| 0.2240   | 0.0425    | 0.1638| 0.9078 | 0.9219       | 0.0115 |

Table 2: Comparison results on JAFEE database.

| Method    | Canberra | Chebyshev | Clark | Cosine | Intersection | Kldist |
|-----------|----------|-----------|-------|--------|--------------|--------|
| LBP       | 0.8603   | 0.1285    | 0.4018| 0.9260 | 0.8455       | 0.0757 |
| AN-FT(MSE)| 0.7139   | 0.8285    | 2.3493| 0.9866 | 0.1715       | 0.1659 |
| AN-FT(KL) | 0.9001   | 0.1349    | 0.4152| 0.9195 | 0.8388       | 0.0827 |
| AN(FC1)   | 0.6514   | 0.0830    | 0.3199| 0.9643 | 0.8928       | 0.0437 |
| AN(FC2)   | 0.4937   | 0.0624    | 0.2400| 0.9792 | 0.9179       | 0.0229 |
| FRCCNN    | 0.4531   | 0.0543    | 0.2031| 0.9572 | 0.9148       | 0.0202 |
| IT2RWFBELN| 0.4326   | 0.0498    | 0.1884| 0.9738 | 0.9246       | 0.0159 |

Table 3: The accuracy (%) on JAFEE data set.

| Methods     | Expression | Accuracy |
|-------------|------------|----------|
| LBP [41]    | 6 + neutral| 85.20    |
| HOG [39]    | 6 + neutral| 84.70    |
| AU-DNN [42] | 7          | 87.80    |
| JFDNN [43]  | 6          | 93.10    |
| CNN [40]    | 6 + neutral| 91.05    |
| C-CNN [44]  | 6          | 96.8     |
| DNN [45]    | 6          | 93.1     |
| IT2RWFBELN  | 7          | 93.56    |

Table 4: The accuracy (%) on BU-3DFE data set.

| Methods     | Poses | Accuracy |
|-------------|-------|----------|
| HOG         | 5     | 52.64    |
| LGBP [46]   | 7     | 74.1     |
| JFDNN [43]  | 5     | 76.5     |
| PCRF [47]   | 5     | 78.1     |
| CGPR [48]   | 5     | 79.5     |
| GSRRR [49]  | 5     | 79.90    |
| DNN [45]    | 5     | 80.10    |
| IT2RWFBELN  | 5     | 87.26    |

Table 5: Comparison results on CASIA database.

| Method     | SD     | SI     |
|------------|--------|--------|
| SVM [8]    | 89.60  | 47.32  |
| ELM [50]   | 89.08  | 47.50  |
| BEL [50]   | 83.28  | 38.55  |
| GA-BEL [51]| 88.41  | 42.24  |
| BELFIS [52]| 75.60  | 39.65  |
| BELBLA [53]| 76.19  | 40.21  |
| IT2RWFBELN | 90.40  | 47.18  |
The proposed model takes advantage of type-2 recurrent wavelet fuzzy theory and brain emotional neural network, and there are no rules initially; then the structure and parameters of model are tuning online simultaneously by gradient approach and Lyapunov function. The system input data streams are directly imported into the neural network through a type-2 recurrent wavelet fuzzy inference system, and then the results were subsequently piped into sensory and emotional channels which jointly produce the final outputs of the network. In order to demonstrate the performance of IT2RWFBELN model, two groups of experiments are conducted, which include facial expression recognition and speech emotion recognition. The results illustrated the effectiveness of the proposed recognition model.

Data Availability

The data used in this paper have been cited in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful for the support of the National Science Foundation of China (Nos. 61502211, 61572242, and 61702234).

Table 6: Comparison results on SAVEE database.

|       | SD   | SI   |
|-------|------|------|
| SVM [8] | 82.60 | 43.82 |
| ELM [50] | 83.08 | 43.50 |
| BEL [50] | 79.28 | 37.53 |
| GA-BEL [51] | 81.43 | 41.24 |
| BELFIS [52] | 70.60 | 38.61 |
| BELBLA [53] | 72.19 | 40.21 |
| IT2RWFBELN | **83.40** | **44.18** |

Table 7: Comparison results on FAU database.

|       | SD   | SI   |
|-------|------|------|
| SVM [8] | 90.28 | 49.02 |
| ELM [50] | 91.17 | 48.46 |
| BEL [50] | 88.47 | 44.19 |
| GA-BEL [51] | 89.91 | 45.24 |
| BELFIS [52] | 80.63 | 39.65 |
| BELBLA [53] | 81.19 | 40.21 |
| IT2RWFBELN | **91.40** | **49.23** |

models. Moreover, the SD recognition accuracy of IT2RWFBELN model is higher than SVM and ELM, but average accuracy of IT2RWFBELN model on SI is similar to ELM.

Table 6 shows the recognition accuracy of BEL model, SVM, ELM, GA-BEL, BELFIS, and BELBLA methods on SAVEE database, which shows that the IT2RWFBELN model has higher accuracy on SD and SI than BEL and GA-BEL models and which is also superior compared with SVM and ELM on SD, but the accuracy of IT2RWFBELN model is lower than the ELM on SI for some instance as well.

Table 7 shows the recognition accuracy of BEL model, SVM, ELM, GA-BEL, BELFIS, and BELBLA methods on FAU database, which shows that the IT2RWFBELN model has higher accuracy on SD and SI than BEL and GA-BEL models and which is also superior compared with SVM and ELM on SD, but the accuracy of IT2RWFBELN model is similar to ELM.

According to the above two groups of experiments, we can conclude that the proposed model could get better recognition effects compared with other models, owing to the combination of type-2 recurrent wavelet fuzzy system and brain emotion learning network. As a result, the proposed method is feasible for facial or speech emotion recognition.

4. Conclusions

This paper introduces one applicable model for emotion recognition, which is vital part in the communication between humans and machines. The proposed model is based on the combination of interval type-2 recurrent wavelet fuzzy system and brain emotional learning network (IT2RWFBELN), which takes advantage of dealing with uncertainties by interval type-2 recurrent wavelet fuzzy system and the lower computation by brain emotional learning network. The proposed model takes advantage of

References

[1] K. Rattanyu, M. Mizukawa, and M. Mizukawa, “Emotion recognition based on ECG signals for service robots in the intelligent space during daily life,” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 15, no. 5, pp. 582–591, 2011.

[2] W. Sun, H. Zhao, and Z. Jin, “An efficient unconstrained facial expression recognition algorithm based on stack binarized auto-encoders and binarized neural networks,” *Neurocomputing*, vol. 267, pp. 385–395, 2017.

[3] P. Song, W. Zheng, S. Ou et al., “Cross-corpus speech emotion recognition based on transfer non-negative matrix factorization,” *Speech Communication*, vol. 83, pp. 34–41, 2016.

[4] C. Rahul, K. Michael, and S. E. Keefe, “A new measure of authentic auditory emotion recognition: application to patients with schizophrenia,” *Schizophrenia Research*, vol. 222, pp. 450–454, 2020.

[5] L. Cen, L. Y. Zhu, F. Wu, Z. L. Yu, and F. Hu, “A real-time speech emotion recognition system and its application in online learning,” *Emotions, Technology, Design, and Learning*, pp. 27–46, 2016.

[6] A. Bhandari and N. R. Pal, “Can edges help convolution neural networks in emotion recognition?” *Neurocomputing*, vol. 433, no. 14, pp. 162–168, 2021.

[7] Y. Sun, G. Wen, and J. Wang, “Weighted spectral features based on local Hu moments for speech emotion recognition,” *Biomed. Signal Process. Control*, vol. 18, pp. 80–90, 2015.

[8] K. Wang, N. An, and B. N. Li, “Speech emotion recognition using Fourier parameters,” *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 69–75, 2017.

[9] A. Rawat and P. K. Mishra, “Emotion recognition through speech using neural network,” *International Journal of
[44] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order," *Pattern Recognition*, vol. 61, pp. 610–628, 2017.

[45] T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan, "A deep neural network-driven feature Learning method for multi-view facial expression recognition," *IEEE Transactions on Multimedia*, vol. 18, no. 12, pp. 2528–2536, 2016.

[46] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," *Computer Vision and Image Understanding*, vol. 115, no. 4, pp. 541–558, 2011.

[47] A. Dapogny, K. Bailly, and S. Dubuisson, "Dynamic pose-robust facial expression recognition by multi-view pairwise conditional random forests," *IEEE Transactions on Affective Computing*, vol. 99, pp. 1–16, 2016.

[48] O. Rudovic, I. Patras, and M. Pantic, "Coupled Gaussian process regression for pose-invariant facial expression recognition," in *Proceedings of the European Conference on Computer Vision*, pp. 350–363, Crete, Greece, September 2010.

[49] W. Zheng, "Multi-view facial expression recognition based on group sparse reduced-rank regression," *IEEE Transactions on Affective Computing*, vol. 5, no. 1, pp. 71–85, 2014.

[50] S. Motamed, S. Setayeshi, and A. Rabiee, "Speech emotion recognition based on a modified brain emotional learning model," *Biologically Inspired Cognitive Architectures*, vol. 19, pp. 32–38, 2017.

[51] Z.-T. Liu, Q. Xie, M. Wu, W.-H. Cao, Y. Mei, and J.-W. Mao, "Speech emotion recognition based on an improved brain emotion learning model," *Neurocomputing*, vol. 309, pp. 145–156, 2018.

[52] M. Parsapoor and U. Bilstrup, "Brain emotional learning based fuzzy inference system (BELFIS) for solar activity forecasting," in *Proceedings of the IEEE 24th International Conference on Tools with Artificial Intelligence*, vol. 1, pp. 532–539, Athens, Greece, November 2012.

[53] 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 Behaviour*, pp. 115–124, MIT Press, Cambridge, MA, USA, 2000.