Research Article

An Approach to Oral English Assessment Based on Intelligent Computing Model

Caihong Jing, Xiaoling Zhao, Haiyan Ren, Xuexia Chen, and Naren Gaowa

1 Nursing College, Inner Mongolia Medical University, Hohhot 010010, China
2 Law and Business Department, Inner Mongolia Chifeng College, Chifeng 010010, China

Correspondence should be addressed to Xiaoling Zhao; jingch@immu.edu.cn

Received 9 February 2022; Revised 28 April 2022; Accepted 3 May 2022; Published 8 June 2022

Copyright © 2022 Caihong Jing et al. 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.

“English craze” has become a key topic of concern to the majority of the people. Apart from setting English as a compulsory course like Chinese and Mathematics in schools, various English training institutions outside the school are also emerging one after another. Due to the English teaching mode in class, dumb English appears. In recent years, with the popularity of virtual electronic devices, more and more researchers try to use virtual reality (VR) to create an immersive English learning environment. Oral English teaching is an important part of the whole English teaching. In the traditional English classroom teaching practice, teachers’ pronunciation is not standard, and students are difficult to learn the correct pronunciation standard, which makes oral English very passive. The most important problem in oral English teaching is to improve students’ interest in oral English, make students willing to speak and realize English communication. Oral English teaching is an important link in both primary and secondary schools and universities. Teachers are afraid of nonstandard pronunciation in limited classrooms, and they are afraid to speak and unwilling to speak, which leads to passive oral English teaching. Therefore, this paper will set up an intelligent computing model to evaluate and analyze spoken English in a standard and accurate way. Artificial intelligence speech synthesis and imitation of voice change are typical applications of decoupling representation learning in speech, the oral evaluation is based on the proposition that speech is a dynamic and complex process. With the help of the rapidly developed computer speech synthesis and imitation technology, an oral evaluation path based on speech synthesis and imitation is proposed, that is, oral evaluation is carried out by using the network parameters and output of deep learning of computer speech imitation.

1. Introduction

With the gradual internationalization of society, English is one of the common languages in the world, and the use of spoken English has become the most basic way of communication in various social behaviors such as tourism, foreign trade, and learning. However, many English studies are too limited to the study on paper, while ignoring the oral English study which can be really used in practice. Xiao [1] mainly describes the formative systematic assessment of oral English in middle school classrooms, while Li et al. [2] make a dynamic assessment study of information-based oral English classrooms. Zhuang [3] studies the identification of nonlinear systems based on intelligent computing model, which can also be used for oral English evaluation. Li [4] expounds the value of intelligent computing in modeling. Based on the fuzzy logic theory in reference [5], an evaluation model of oral English is established. Wan [6] expands the application of oral English in international teaching. Based on the background of artificial intelligence in literature [7], APP is used for students’ oral learning mode. Based on the literature [8], the evaluation research of learning theory in oral English teaching class of higher vocational colleges is constructed. Through the artificial intelligence in literature [9], the artificial spoken pronunciation is recognized, and whether the pronunciation emits standard spoken language is judged. The results can be measured more conveniently by using the automatic speech recognition instrument in reference [10]. A study of oral English learning and assessment in SELL-Corpus and VR

1.1 The Key Conditions of Oral English Evaluation

The key conditions of oral English evaluation are the ability to pronounce correctly and the ability to express fluently. Pronunciation is the most important condition to evaluate oral English, while pronunciation is one of the key conditions of oral English evaluation. However, the pronunciation ability of students varies greatly, and the pronunciation ability of students in limited classrooms is not very good. Therefore, this paper will set up an intelligent computing model to evaluate and analyze spoken English in a standard and accurate way. Artificial intelligence speech synthesis and imitation of voice change are typical applications of decoupling representation learning in speech, the oral evaluation is based on the proposition that speech is a dynamic and complex process. With the help of the rapidly developed computer speech synthesis and imitation technology, an oral evaluation path based on speech synthesis and imitation is proposed, that is, oral evaluation is carried out by using the network parameters and output of deep learning of computer speech imitation.
2. The Development Status of Oral Evaluation

2.1. Development of Oral English Assessment Abroad. In recent years, with the rapid development of computer technology and the development of intelligent speech recognition technology, using intelligent computing model to evaluate the standardization of spoken language has become a research hotspot. However, foreign countries have developed rapidly in oral evaluation, and they have designed a language learning system VILTS for speech interaction to evaluate the speech spoken by users. The system scores intelligently from four different aspects of phonetic research: similarity, phonetic accuracy, phonetic emotion, and speech speed. There are many big differences between computer automatic evaluation and artificial evaluation, and people constantly improve the evaluation methods in constant exploration. Gina-Anne levow lists the difficulties faced by simulators in modeling artificial scores from two aspects: process and result, and points out that it is impossible to model the evaluation process comprehensively in the field of speech features and speech recognition at present. In addition, the recognizer adopts a wide range of grammatical forms, which enables it to accept different grammatical and semantic changes based on the target language. In the nonacoustic field, machine learning technology is applied to gather human recognition related features. At the same time, the network-based oral English learning is also developing rapidly, which is to solidify the ready-made speech evaluation technology into oral evaluator, which greatly facilitates oral practitioners. To sum up, foreign countries have carried out in-depth research and analysis on oral English assessment, and have involved a wider field of oral English features. Some systems have been recognized by some professional language experts for assessing spoken English pronunciation from different levels such as sentences, words, or phonemes.

2.2. Development of Oral English Assessment in China. The research on oral English assessment has just started in China, and only some research institutes in Taiwan have carried out some related research. The evaluation of speech is divided into three parts: the content of the uttered speech piece, the pronunciation standard, and the oral evaluation of the speech database. The first part is obtained by calculating the HMM probability of the given speech normalization. In the second part, the Viterbi decoded syllables are identified by GMM-based tone recognition. The third part is realized by the greedy search algorithm. It is not necessary to apply the whole process of speech recognition to speech evaluation, but only to build a linear grid of text-based speech model. Then, Viterbi decoding is used to align the mesh model that users said and built. Based on speech recognition and accent adaptation technology based on implicit Markov model, oral evaluation in mainland China studies the accuracy and fluency information of phoneme pronunciation, gives phoneme-level pronunciation quality scores, and further obtains the scoring results of the whole sentence. This method has gained most recognition from some specialized experts.

3. Oral Fuzzy Logic Combined with Neural Network Learning

3.1. The Concept of Fuzzy Relations. The fuzzy set on the direct product \( U \times N \) is defined, with \( U \) and \( N \) as computational domains, and \( R \) is defined as the fuzzy relation of \( U \) on \( N \); \( R(x,y) \) expresses the degree of connection between \( x \) and \( y \) for \( (x, y) \) and \( u(x) \). If \( R \) is a classical set on the direct product \( U \times N \), then the relation \( R \) is \( U \) and \( N \) is a general relation, so the fuzzy relation is the extension of the classical relation, and the classical relation is a special fuzzy relation. Suppose \( U \) is a finite computational domain composed of \( m \) elements and \( N \) is a finite universe composed of \( n \) elements. The fuzzy relation \( R \) from \( U \) to \( N \) can be expressed by a matrix of \( m \times n \), namely,

\[
\begin{bmatrix}
    r_{11} & \cdots & r_{1n} \\
    \vdots & \ddots & \vdots \\
    r_{m1} & \cdots & r_{mn}
\end{bmatrix}
\]  

(1)

It can also be expressed as \( R = (r_{ij})_{m \times n} \). It is not difficult to find that there is a certain mapping relationship between fuzzy matrix and fuzzy relation. Definition (fuzzy relation has composition): let \( U, N, W \) be the universe, \( R \) be a fuzzy relation from \( U \) to \( N \), and \( Q \) be a fuzzy relation from \( N \) to \( W \), then the composition \( T \) from \( R \) to \( Q \) is also a fuzzy relation.

\[
u_{RO}(X, Z) = N_{yxv} \left( u_{R}(x, y) \land u_{Q}(y, z) \right) (x \in U, z \in W).
\]  

(2)

Let \( R = (r_{ij})_{m \times n} \) and \( Q = (q_{jk})_{n \times m} \) be two matrices with fuzzy relations, and their composite \( R \times Q \) is a fuzzy matrix \( S \) with \( n \) rows and one column, and the elements of row \( i \) and column \( k \) of \( S \) are

\[
S_{ik} = N^{m}_{j=1} \left( r_{ij} \land q_{jk} \right).
\]  

(3)

3.2. Fuzzy Logical Reasoning Method

3.2.1. Zadeh Reasoning Method. Let \( F \) be a fuzzy set on \( U \), \( G \) be a fuzzy set on \( N \), and the fuzzy implication relation “if \( F \) then \( G \)” is expressed by \( F \rightarrow G \). Zadeh defined it as a fuzzy relation on \( U \times N \), namely,

\[
R = F \rightarrow G(F \times G)(F^{c} \times N).
\]  

(4)

It belongs to the function:

\[
R(x, y) = [F(x) \land G(y)] \lor [1 - A(x)].
\]  

(5)
When a fuzzy relation $R$ is set, a fuzzy relation transformation is determined. Using the synthesis of fuzzy relations, there are the following reasoning rules:

Given the fuzzy relation $R$ of the fuzzy implication relation $F \rightarrow G$, for a given $F' \in U$, we can infer $G'$, $G' \in N$, $G' = F' \times R$. That is, when $W$ is a finite universe:

$$G'(y) = W[F'(X) \land [F(x) \land G(y)W(1 - F(X))]]. \quad (6)$$

3.2.2. Mamdani Reasoning Method. In Mamdani’s reasoning method, the fuzzy implication relation $F \rightarrow G$ is expressed by the direct product of $F$ and $G$, that is, $R = F \rightarrow G = F \times G$, which can also be written as

$$R(x, y) = F(x) \times G(y). \quad (7)$$

3.2.3. Most of Them Are Input Fuzzy Reasoning Methods. It is known that the major preconditions of reasoning are “if $A$ and $B$ then $C$,” $A \in F(x)$, $B \in F(y)$, $C \in F(z)$, and the fuzzy implication relation is “if $A$ and $B$ then $C$’s”:

$$R = A \times B \times C = (A \times B) \rightarrow C,$$

$$R(x, y, z) = A(x) \land B(y) \land C(z). \quad (8)$$

3.3. Establishment Process of Fuzzy System

(1) Composition of fuzzy system: the fuzzy system refers to those systems associated with fuzzy essentials and fuzzy logic. It consists of four parts: fuzzy receiver, fuzzy speech library, fuzzy server, and defuzzy processor, as shown in Figure 1:

(2) Fuzzy speech: assuming that a variable has a value of “X”, $F$ is called X fuzzification if “X” is replaced by a set of functions. Commonly used fuzzification methods include single-factor fuzzification, Gaussian function fuzzification, and trigonometric function fuzzification. The blur receiver completes the blur operation.

(3) Establishing fuzzy speech database: fuzzy speech library is composed of a series of “if-then” fuzzy conditional sentences, in which the prelude is input and state, and the postlude is control variable. There are two common methods to establish fuzzy rule base, one is from the experience of experts, which is summarized into a group of rules according to the experience of experts. The other is through self-learning methods, such as neural network or genetic algorithm.

4) Fuzzy reasoning of speech: fuzzy reasoning is to transform the fuzzy “if-then” of fuzzy speech database into some kind of mapping according to fuzzy logic rules. Fuzzy reasoning generally includes three parts: aggregation, that is, the calculation of IF part of rules, composition, calculation of THEN part of rules.

(5) Defuzzification: after the fuzzy reasoning, the results are still expressed by data speech. We must convert linguistic variables to a certain value, that is, map a fuzzy set to a certain point. This stage is called defuzzification. The relationship between linguistic values and corresponding values is given by the definition of membership function. The general method of defuzzification includes two stages: first, calculate the linguistic value of each linguistic variable to get a "typical value," and the method of calculating each linguistic typical value is to find the maximum value of each membership function; then, the best compromise value of fuzzy logic reasoning is calculated.

3.4. Learning Methods of Neural Networks

(1) A learning algorithm for correcting computational errors. It continuously adjusts the neurons according to the increasing intensity of voice data and output error data.

Let $y_k(n)$ be equal to the actual output of neuron $k$ at time $n$ when the target input $x(n)$, $k$ is assumed, and $d_k(n)$ is the known output given by the training sample, then the output error can be expressed as

$$e_k(n) = d_k(n) - y_k(n). \quad (9)$$

The learning algorithm for correcting computational errors is to minimize the objective function value based on $e_k(n)$, so that the actual output of each output neuron in the network is close to the output in a certain statistical sense. This problem is also the problem of transforming to find the minimum value. The most commonly used error objective function is the mean of the sum of squares of errors, which is

$$E_n = \frac{1}{2} \sum_k e_k^2(n). \quad (10)$$

If the relation weight of neurons $i$ to $j$ is $w_{ij}$, the adjustment amount of weight is
\[ \Delta w_{ij} = \eta \delta_j v_i, \]  

(11)

In the formula, \( \eta \) is the learning rate, \( \delta_j \) is the partial derivative of the error function to the input of neuron \( j \), and \( v_i \) is the output of the \( i \)-th neuron.

(2) Hebb learning algorithm: according to the mechanism of conditioned reflex in biology, it belongs to unsupervised learning by psychologist in 1949. This rule means that if two neurons are activated simultaneously, the strength of the connection increases, and vice versa are described mathematically as follows:

\[ \Delta w_{kj} = f(x_i, n, y_k(n)), \]  

(12)

where \( x_i \) and \( y_k(n) \) are the states of neurons connected to \( w_{kj} \), respectively, and the most commonly used \( f \) function is as follows:

\[ \Delta w_{kj} = \eta x_i n \times y_k(n). \]  

(13)

(3) Random learning algorithm: the error learning algorithm usually adopts gradient descent method, but the problem of this algorithm is that it may lead to local optimum. The stochastic learning algorithm achieves global optimum by introducing unstable factors.

(4) Competitive learning algorithm: it means that the outputs of neural network compete with each other, and the strongest will be activated. The rules of competitive learning are

\[ \Delta w_{ij} = \begin{cases} \eta(x_i - w_{ij}), & \text{if neuron } j \text{ wins the competition,} \\ 0, & \text{if neuron } j \text{ fails to compete.} \end{cases} \]  

(14)

3.4.1. BP Neural Network Algorithm. The specific implementation steps of BP standard algorithm are as follows:

(1) Network initialization.

(2) Take any \( p \)-th input sample and its corresponding expected output

\[ d(p) = (d_1(p), d_2(p), ..., d_n(p)). \]  

(15)

(3) Calculate the hidden layer input and output. The formula is as follows:

\[ h_{ij}(p) = \sum_{i=1}^{n} x_i(p) - b_j, \quad j = 1, 2, 3, ..., n, \]  

(16)

\[ h_{oj} = f(h_{ij}(p)), \quad j = 1, 2, 3, ..., n. \]

(4) Calculate the input and output of the output layer.

\[ y_{ikp} = w_{ik}x_i(p) - b_k, \quad k = 1, 2, ..., m. \]  

(17)

(5) The partial derivative of the error function to each neuron in the output layer is calculated:

\[ \delta_k(p) = (d_k(p) - y_0_k(p))(1 - y_0_k(p)), \quad k = 1, 2, ..., m. \]  

(18)

(6) Using error correction algorithm, the partial derivatives of neurons in the hidden layer are obtained:

\[ \delta_j(p) = \left( \sum_{k=1}^{m} \beta_k(p)w_{jk} \right)(h_{oj}(p)(1 - h_{oj})), \quad j = 1, 2, ..., n. \]  

(19)

(7) Adjust the connection weight between the hidden layer and the output layer.

\[ w^N_{kj}(p) = w^N_{kj}(p) + \eta \delta_j(p)x_i(p), \]  

\[ b^N_{kj}(p) = b^N_{kj}(p) + \eta \delta_j(p). \]  

(20)

(8) Adjust the connection weight between the input layer and the hidden layer.

\[ v^N_{ij}(p) = v^N_{ij}(p) + \eta \delta_j(p)x_i(p), \]  

\[ b^N_{ij}(p) = b^N_{ij}(p) + \eta \delta_j(p). \]  

(21)

(9) Find the error \( E \) of the whole calculation process, and the calculation formula is as follows:

\[ E = \frac{1}{2m} \sum_{p=1}^{l} \sum_{k=1}^{m} (d_k(p) - y_0_k(p))^2. \]  

(22)

(10) Judge whether to continue training: when \( E < \varepsilon \) or the number of learning times is greater than the set maximum number of times \( M \), the training ends. Otherwise, randomly select another learning sample, return to step 3, and carry out circular training. The BP neural network algorithm is enough for the daily use of oral evaluation, and the improved BP algorithm can be superior to BP algorithm in complex environment.

4. Experiment

4.1. Sample Selection. In artificial neural network and adaptive fuzzy neural system, samples are the object of network training, and the difference degree of sample selection directly affects the evaluation results of the model, so the selection of samples must be representative. The samples used in this study were 5 men and 5 women of the same age and with the same oral literacy from different regions for oral evaluation. Because this experiment is to evaluate spoken English, the representative samples must first be people who can speak English but do not have spoken English in local dialects so that they can evaluate spoken English correctly. The relevant collection sources have been explained in the experiment. The speech information in this paper is processed data, and the main work of this paper is the research of spoken language recognition. At present, there are many speech processing tools, such as prefiltering,
Table 1: Hierarchical cross error of fuzzy logic combined with neural network.

<table>
<thead>
<tr>
<th>Classify samples correctly (%)</th>
<th>Error classification sample (%)</th>
<th>Kappa</th>
<th>Average absolute error</th>
<th>Root mean square error</th>
<th>Relative absolute error (%)</th>
<th>Relative square root error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>86.16</td>
<td>12.31</td>
<td>0.72</td>
<td>0.0092</td>
<td>0.212</td>
<td>23.1</td>
<td>43.21</td>
</tr>
</tbody>
</table>

Table 2: Accuracy rate of fuzzy logic combined with neural network.

<table>
<thead>
<tr>
<th>Correct rate</th>
<th>Error rate</th>
<th>Precision rate</th>
<th>Feedback rate</th>
<th>ROC area</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.882</td>
<td>0.024</td>
<td>0.843</td>
<td>0.882</td>
<td>0.992</td>
<td>Good</td>
</tr>
<tr>
<td>0.877</td>
<td>0.021</td>
<td>0.844</td>
<td>0.877</td>
<td>0.868</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.912</td>
<td>0.012</td>
<td>0.793</td>
<td>0.912</td>
<td>0.911</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.932</td>
<td>0.013</td>
<td>0.792</td>
<td>0.932</td>
<td>0.893</td>
<td>Good</td>
</tr>
<tr>
<td>0.878</td>
<td>0.022</td>
<td>0.851</td>
<td>0.878</td>
<td>0.994</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

Table 3: Stratified cross error of the ANFIS model.

<table>
<thead>
<tr>
<th>Classify samples correctly (%)</th>
<th>Error classification sample (%)</th>
<th>Kappa</th>
<th>Average absolute error</th>
<th>Root mean square error</th>
<th>Relative absolute error (%)</th>
<th>Relative square root error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>78.16</td>
<td>15.31</td>
<td>0.82</td>
<td>0.0098</td>
<td>0.202</td>
<td>33.1</td>
<td>63.21</td>
</tr>
</tbody>
</table>

Table 4: Accuracy of ANFIS model.

<table>
<thead>
<tr>
<th>Correct rate</th>
<th>Error rate</th>
<th>Precision rate</th>
<th>Feedback rate</th>
<th>ROC area</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.782</td>
<td>0.047</td>
<td>0.743</td>
<td>0.882</td>
<td>0.992</td>
<td>Good</td>
</tr>
<tr>
<td>0.677</td>
<td>0.081</td>
<td>0.644</td>
<td>0.877</td>
<td>0.868</td>
<td>General</td>
</tr>
<tr>
<td>0.812</td>
<td>0.022</td>
<td>0.893</td>
<td>0.912</td>
<td>0.911</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.732</td>
<td>0.043</td>
<td>0.792</td>
<td>0.932</td>
<td>0.893</td>
<td>Good</td>
</tr>
<tr>
<td>0.678</td>
<td>0.092</td>
<td>0.751</td>
<td>0.878</td>
<td>0.994</td>
<td>General</td>
</tr>
</tbody>
</table>

Table 5: Hierarchical cross error of bp neural network.

<table>
<thead>
<tr>
<th>Classify samples correctly (%)</th>
<th>Error classification sample (%)</th>
<th>Kappa</th>
<th>Average absolute error</th>
<th>Root mean square error</th>
<th>Relative absolute error (%)</th>
<th>Relative square root error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>68.16</td>
<td>20.31</td>
<td>0.83</td>
<td>0.0138</td>
<td>0.302</td>
<td>43.1</td>
<td>83.21</td>
</tr>
</tbody>
</table>

A/D conversion, pre-emphasis, framing, windowing, endpoint detection, and so on.

4.2. Experimental Testing. Based on fuzzy logic combined with neural network intelligent computing model, people from different regions are found for the oral evaluation test, and the relevant data we get are as follows:

Hierarchical cross-validation errors are shown in Table 1.

People from five different regions were found to take the oral evaluation test, and the accuracy rate is as shown in Table 2:

Based on the ANFIS model, we found the same batch of people for oral evaluation test, and the relevant data we got are as follows:

Hierarchical cross-validation errors are shown in Table 3.

The accuracy of the ANFIS model is shown in Table 4.

Based on the bp neural network, we found the same batch of people for oral evaluation test, and the relevant data we got are as follows:

Hierarchical cross-validation errors are shown in Table 5.

The accuracy of the bp neural network is shown in Table 6.

4.2.1. Model Comparison. We compare the fuzzy logic with the neural network intelligent computing model, ANFIS network model, and bp neural network model to evaluate the accuracy of spoken English as shown in Figure 2.

This work makes statistics on the error and accuracy of hierarchical cross-validation of models so that readers can intuitively know the error rate and accuracy of each model.

4.3. Experimental Analysis. Spoken language evaluation based on the fuzzy logic combined with the neural network intelligent computing model. In order to test the reliability of the model more concretely, we decided to add speech emotion index to test the comprehensive evaluation results of pronunciation quality. We found the same sample to evaluate oral English in four states: happiness, sadness,
anger, and surprise, and showed the mean and variance of speech emotion evaluation results in different manual ratings, as shown in Figures 3–6.

Through the oral evaluation charts of four different emotional states, we can see that the variance of the evaluation results is the smallest and the most stable in the surprised state. In sad mood, the variance is the largest, and the oral evaluation results are the most fluctuating.

### 4.4. Contrast Test

According to the experimental test in this paper, the fuzzy logic combined with the neural network intelligent computing model has achieved the most accurate accuracy for oral English evaluation. We are now conducting a comparative experiment between artificial evaluation and this model, as shown in Table 7.

Table 7 lists the average number of spoken language evaluation results of four kinds of speech emotion under...
Figure 4: Evaluation results when you are sad.

Figure 5: Evaluation results when angry.

Figure 6: Evaluation results when surprised.
different manual ratings, which shows that different ratings can be distinguished. However, whether the mean difference of different manual ratings is essential or random, this paper uses one-way ANOVA to test. By observing the $p$ values in the data results table, we can see that the $p$ values under the four emotions are all less than 0.05, which shows that the differences between different emotional pronunciation quality levels have statistical significance.

## 5. Conclusion

With the gradual internationalization of society, we cannot avoid oral English communication with foreign compatriots not only at home but also abroad. Also, designing a standardized model for testing spoken English has become a concrete problem that we need to solve now. Based on the intelligent computing model, this paper analyzes the standardized methods of export language evaluation more accurately. The research results are as follows:

1. By comparing the fuzzy logic combined with the neural network intelligent calculation model and ANFIS network model, it is known that the accuracy rate of fuzzy logic combined with neural network intelligent calculation model is as high as 80%, and the error rate is less than 2%.

2. This paper combines the fuzzy logic algorithm and BP neural network algorithm to form a computational method that can measure spoken language standardization.

3. The analysis of the experiment is always through the further specific test of the speech in four different emotional States. The accuracy and error rates in the four emotional states are inconsistent. The variance of people’s oral evaluation in the state of surprise is the smallest and the most stable.

4. In the comparative experiment, the intelligent computing model of fuzzy logic combined with neural network is compared with manual evaluation under four emotional states, and the intelligent computing model far exceeds the standardization of manual evaluation.

<table>
<thead>
<tr>
<th>Artificial evaluation</th>
<th>Intelligent computing model of fuzzy logic combined with neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manual evaluation</td>
</tr>
<tr>
<td></td>
<td>Happy</td>
</tr>
<tr>
<td>A</td>
<td>0.35</td>
</tr>
<tr>
<td>B</td>
<td>0.32</td>
</tr>
<tr>
<td>C</td>
<td>0.28</td>
</tr>
<tr>
<td>D</td>
<td>0.29</td>
</tr>
<tr>
<td>Univariate variance $E$</td>
<td>6.6</td>
</tr>
<tr>
<td>$F$ of one-way variance</td>
<td>0</td>
</tr>
</tbody>
</table>

### Data Availability

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

### Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

### References


