WILEY | Hindawi

Retraction

Retracted: Optimization of Intelligent English Pronunciation Training System Based on Android Platform

Complexity

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Cao and H. Hao, "Optimization of Intelligent English Pronunciation Training System Based on Android Platform," *Complexity*, vol. 2021, Article ID 5537101, 11 pages, 2021.

WILEY | Hindawi

Research Article

Optimization of Intelligent English Pronunciation Training System Based on Android Platform

Qianyu Cao [1] and **Hanmei Hao** [2]

[1]School of Foreign Languages, Chengdu University of Information Technology, Chengdu 610036, China
[2]Chengdu Angke Technologies Co., Ltd., Chengdu 610000, China

Correspondence should be addressed to Qianyu Cao; cqy@cuit.edu.cn

Oral English, as a language tool, is not only an important part of English learning but also an essential part. For nonnative English learners, effective and meaningful voice feedback is very important. At present, most of the traditional recognition and error correction systems for oral English training are still in the theoretical stage. At the same time, the corresponding high-end experimental prototype also has the disadvantages of large and complex system. In the speech recognition technology, the traditional speech recognition technology is not perfect in recognition ability and recognition accuracy, and it relies too much on the recognition of speech content, which is easily affected by the noise environment. Based on this, this paper will develop and design a spoken English assistant pronunciation training system based on Android smartphone platform. Based on the in-depth study and analysis of spoken English speech correction algorithm and speech feedback mechanism, this paper proposes a lip motion judgment algorithm based on ultrasonic detection, which is used to assist the traditional speech recognition algorithm in double feedback judgment. In the feedback mechanism of intelligent speech training, a double benchmark scoring mechanism is introduced to comprehensively evaluate the speech of the speech trainer and correct the speaker's speech in time. The experimental results show that the speech accuracy of the system reaches 85%, which improves the level of oral English trainers to a certain extent.

1. Introduction

As the most mature language in globalization, English has become an indispensable communication tool in people's daily life. Oral English, as an important part of English learning, its corresponding pronunciation, the language environment of nonnative English-speaking countries, and the teachers of oral English are the factors that restrict the development of oral English [1, 2]. With the rapid development of intelligent speech recognition technology, professional machines and corresponding learning software related to oral English training emerge in endlessly. However, such software often has only some single functions, which can only assist oral English practitioners to carry out some simple oral pronunciation training, thus lacking effective feedback on some learners' oral English pronunciation, so its corresponding functions are limited. The training effect is often unsatisfactory [3–5]. At the same time,

the traditional speech recognition technology itself is vulnerable to the interference of the surrounding environment and other problems. At the same time, the professional spoken English training machine is too complex and huge and it does not have the characteristics of portable, so it is difficult to achieve civil and lightweight [6]. Therefore, how to design a piece of intelligent oral English training equipment and design a reasonable and efficient training algorithm is very important and meaningful.

Based on this, a large number of scientists and related research institutions have carried out research on the oral English training algorithm. Traditional oral English training mainly focuses on the research of speech recognition algorithm, which mainly depends on the rapid development of computer technology, artificial intelligence technology, and information and communication technology [7–9]. In the continuous development of speech recognition technology, there are mainly speech signal linear prediction coding

technology, dynamic time planning adjustment technology, linear prediction cepstrum technology, and dynamic time warping technology [10–12]. In the application of speech technology in computer-aided language learning, it is mainly the application of information technology to combine speech recognition technology with oral English training courses, so as to create a real oral English learning environment for oral English learners, and on this basis, in order to promote the virtuous circle of oral English learning, we should increase the corresponding multimedia technology and improve the interest of English learners [13–15]. Relevant scientific research institutions have designed different language training algorithms and equipment based on mature speech recognition systems, such as the language assisted learning system of Carnegie Mellon University in the United States, the education assisted learning system of relevant countries in Asia, and the language assisted teaching system of player in relevant European and American countries [16–19]. Based on the above research, the traditional oral English pronunciation practice relies too much on the realization of certain speech recognition technology, and it also depends on the development of related multimedia technology. Based on this, this paper will develop and design the spoken English assisted pronunciation training system based on Android smartphone platform. On the basis of in-depth research and analysis of spoken English pronunciation correction algorithm and speech feedback mechanism, this paper proposes a lip movement judgment algorithm based on ultrasonic detection, which is used to assist the traditional speech recognition algorithm to make double feedback judgment. In the feedback mechanism of intelligent speech training, a double benchmark scoring mechanism is introduced to evaluate the pronunciation of speech trainers in an all-round way and correct the pronunciation of speakers in time. The experimental results show that the pronunciation correction rate of the proposed system reaches 85%, which improves the level of oral English trainers to a certain extent.

Based on the above research background and significance analysis, the structure of this paper is as follows: in Section 2, we will analyze the related research literature of intelligent English pronunciation training system and give the disadvantages of the current related system; in Section 3, we will focus on the research and analysis of the lip movement judgment algorithm based on ultrasonic detection under the technology of speech recognition algorithm and give some suggestions. Section 4 will be based on an Android mobile phone development test and analysis of the corresponding experimental results; finally, this paper is summarized in Section 5.

2. Related Work Analysis: Research and Analysis of Intelligent Oral English Pronunciation Training

The main difficulty of intelligent spoken English pronunciation training system lies in the realization of recognition algorithm and corresponding feedback evaluation mechanism algorithm [20–22]. In the aspect of speaker speech recognition algorithm, the main recognition algorithms are focused on the speaker speech recognition algorithm. A large number of researchers and research institutions have studied and analyzed this kind of algorithm. The fluency pronunciation training system designed by researchers from European and American countries adopts the automatic speech recognition technology based on Sphinx engine. The system can correct the errors of syllables and prosody of oral English learners. However, the system points out that the errors of relevant syllables are too limited, and the whole system is too mechanical to analyze the learners' pronunciation based on words. Japanese research institutions have developed an oral English assisted pronunciation training system, which is mainly designed for Asian nonnative English-speaking countries. It has its own limitations. It only accepts nonnative accents and only provides feedback and visualization for the corresponding accurate information. The model used in the system is also a specific model; the relevant research and development institutions in the United States have designed a software development kit for oral English pronunciation training, which can realize the oral evaluation of English speakers and also provide the pronunciation spectrum and the corresponding duration scoring function, but the system depends on professional auxiliary training equipment. It is not portable and light. At the level of computer-aided pronunciation, computer-aided pronunciation assessment technology can make learners know their pronunciation level and ability at any time, so as to learn more targeted and train in the right direction. On the basis of speech evaluation technology based on statistical speech recognition, related research institutions and linguists have studied the core algorithm of speech evaluation, the adaptive method of acoustic model of speech evaluation, the application of duration and speed in speech evaluation, and the scoring mapping model of speech evaluation system, and many reliable algorithms are given.

3. Analysis and Research on Lip Motion Judgment Algorithm of Ultrasonic Detection Based on Speech Recognition Technology

This section will focus on the analysis of spoken English pronunciation training core algorithm design. This paper focuses on the original speech recognition algorithm based on the addition of ultrasonic detection lip motion judgment algorithm, in order to assist the accurate recognition of the whole system. The corresponding algorithm core architecture of speech recognition technology is shown in Figure 1, from which we can clearly see the operation mechanism of the whole core algorithm.

3.1. Analysis of Dual Detection Algorithm for Spoken English Pronunciation. In this paper, the core algorithm of spoken English speech recognition algorithm is the hybrid detection method which combines the conventional speech recognition algorithm with the auxiliary lip detection algorithm. The
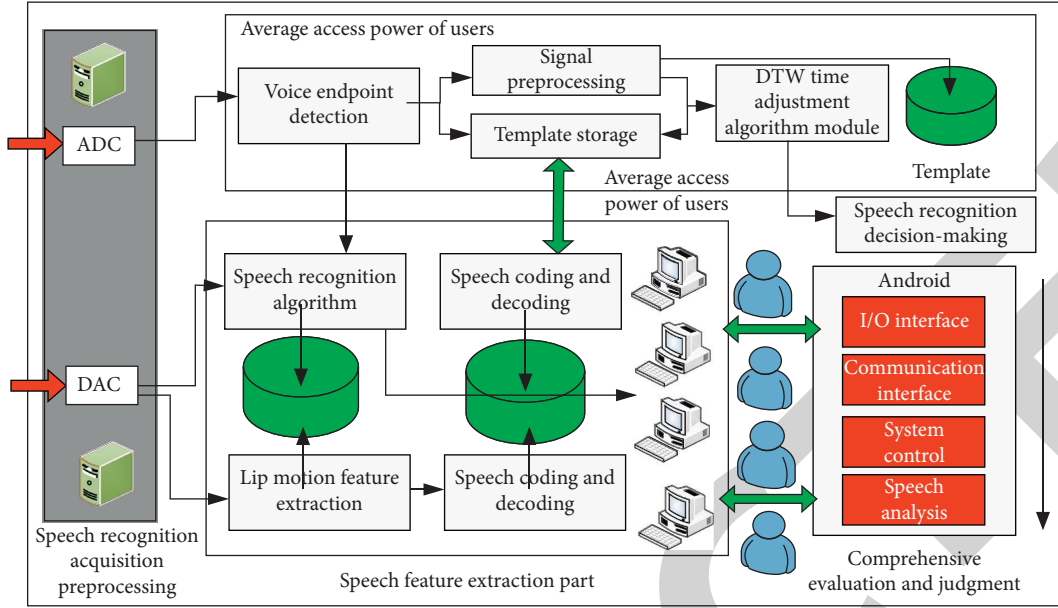
FIGURE 1: Framework of the speech recognition technology core algorithm of spoken English pronunciation training system.

algorithm is mainly composed of speech signal preprocessing module, speech feature and speaker lip feature extraction module, speech dynamic time adjustment recognition algorithm, and speech recognition feedback evaluation mechanism algorithm. The corresponding recognition principle block diagram is shown in Figure 2, and the technical details of the corresponding modules are as follows.

Speech signal preprocessing module: when the speaker carries out oral English training, the corresponding voice is processed through the speech preprocessing module. The corresponding processing steps mainly include digital processing of speech signal, speech endpoint detection, speech framing detection, speech windowing detection, and speech pre-emphasis detection.

In the preprocessing part of speech digitization, the analog signal corresponding to speech is mainly converted into the corresponding digital signal, and the corresponding digitization process includes speech analog sampling and digital quantization processing. In the actual system design, this paper uses the unique platform of Android mobile phone platform for sampling and reasonably analyzes the corresponding digital sampling rate and corresponding quantization in the digitization process to achieve the current input voice configuration requirements. After sampling and digital quantization, it is necessary to preemphasize the current voice file, so as to enhance the highfrequency signal of the input voice signal and filter out the low-frequency signal in the corresponding voice signal, so as to make the signal spectrum more flat. In this module, the corresponding digital filter is installed. The core calculation formula of the corresponding digital filter is shown as formula (1), where W is the corresponding cutoff frequency and I is the order of the filter. After the completion of digital filtering, the corresponding digital voice file needs to be processed by frame and window. In the process of

processing, the corresponding voice input occurred in a very short time is defined as a steady-state input signal, and it is regarded as a fixed signal. For the long-time signal, the short-time signal will be used as the frame reference for segmenting processing, so as to achieve the goal of long-time voice input reasonable analysis and preprocessing; based on this, the framing function used in this paper corresponds to formula (2), in which the corresponding frequency is set to 50 Hz, the corresponding sampling point is designed to about 5000 points, the corresponding window function in the formula is fn, the corresponding voice signal length is set to N, the corresponding "inc" represents the adjustment length, and the corresponding "overlap" represents the overlap part.

$$H(i)^2 = \frac{1}{\left(1 + (w/w_i)^{2n}\right)} = \frac{1}{\left(1 + e^2 (w/w_i)^{2n}\right)}, \tag{1}$$

$$fn = \left\{ \begin{array}{ll} \dfrac{(N - \text{overlap})}{\text{inc}}, & \{0 < n < N - 1\} \\ \\ 0, & (\text{other}) \end{array} \right\}. \tag{2}$$

In the aspect of voice signal endpoint detection technology, there are mainly two kinds of algorithms. The voice endpoint detection part is mainly based on the short-term energy detection technology to process the voice part, while the action detection algorithm is mainly used to judge the lip action components corresponding to the speaker. In the actual judgment, the "and" calculation results will be combined with the judgment results of the two components, only when the results are not correct. When there is a problem, make a speech recognition judgment. In the short-term energy detection part of the speech part, it is mainly based on the rule that the speech energy changes with time. The corresponding core calculation formula is shown as formula (3), where the corresponding n represents the number of
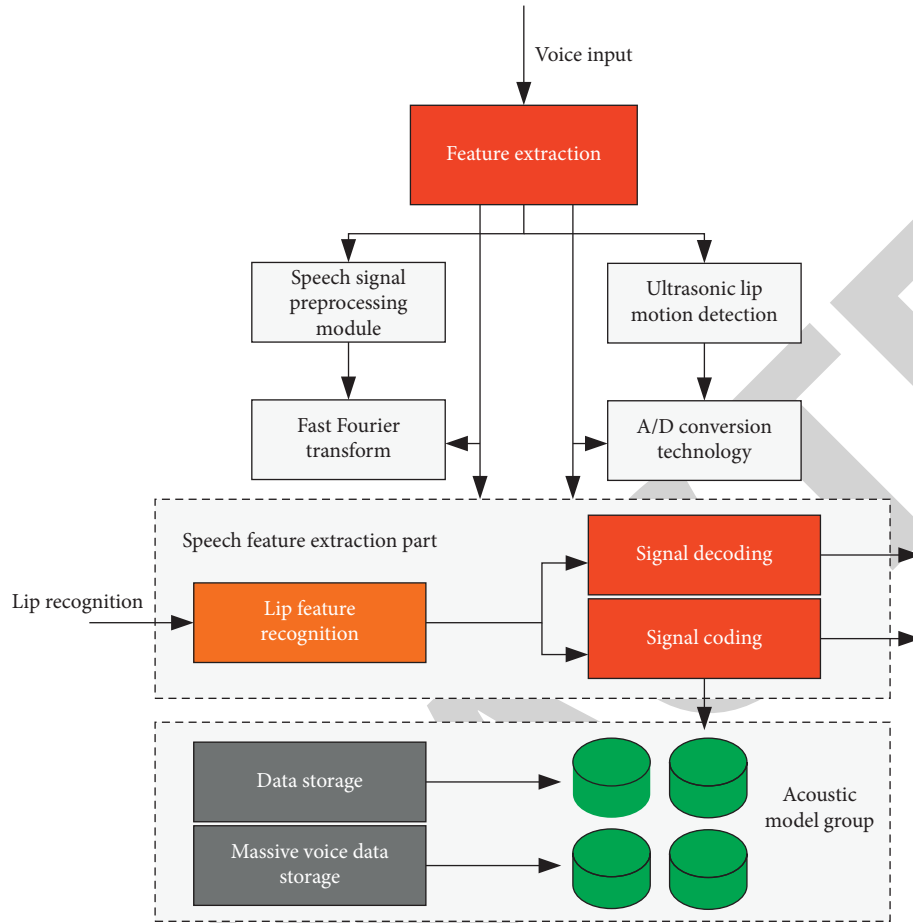
Figure 2: Schematic diagram of speech recognition technology for oral English pronunciation training.

frames of the speech signal, the corresponding energy is represented by W_n, and the corresponding $b(n)$ represents the window function:

$$W_n = \sum_{m=-\infty}^{\infty} |a(m)\,b(n-m)|^2 = a(n)^2 * f(n), \quad 0 < a < N - 1. \tag{3}$$

In the motion detection algorithm, it is mainly based on the speaker's lip frequency shift detection technology, which mainly includes two detection steps, corresponding to the boundary scan and the corresponding secondary frequency peak retrieval. In the corresponding boundary scan technology level, it is mainly to determine the range of the frequency spectrum of the corresponding lip reflection, realize the fast Fourier transform of the signal, and quickly transform the corresponding time-domain signal. After the corresponding spectrum signal is processed, the corresponding main frequency peak is selected as the center point of the transmission frequency, and the frequency points in the positive and negative directions are scanned. In the corresponding secondary frequency peak retrieval part, the main purpose is to determine the secondary frequency peak of the corresponding spectrum result.

In the aspect of speech signal feature extraction, it mainly depends on two aspects, namely, speech feature extraction and lip motion feature extraction. In the corresponding speech feature extraction level, it mainly includes pre-processing technology, fast Fourier transform technology, spectral line energy processing technology, and DCT cepstrum algorithm. In the corresponding preprocessing stage, it is essentially similar to the preprocessing module of the whole system. In the fast Fourier transform level, it is mainly to quickly extract the energy information of speech signal and quickly determine the corresponding energy distribution. In addition, the corresponding spectral line energy is calculated based on the corresponding fast Fourier transform results, and finally the speech features are extracted by the DCT cepstrum algorithm (the main speech features include voiceprint, volume, and tone). For lip motion feature recognition, the main task is to cut the lip reflected signal, extract the mouth features of the corresponding unit based on the unit, and query the corresponding pronunciation table (based on the pronunciation table of 10 basic mouth types) to segment the corresponding signal, extract the corresponding frequency shift features. The schematic diagram of speech signal feature extraction is shown in Figure 3, from which we can clearly see the principle process diagram
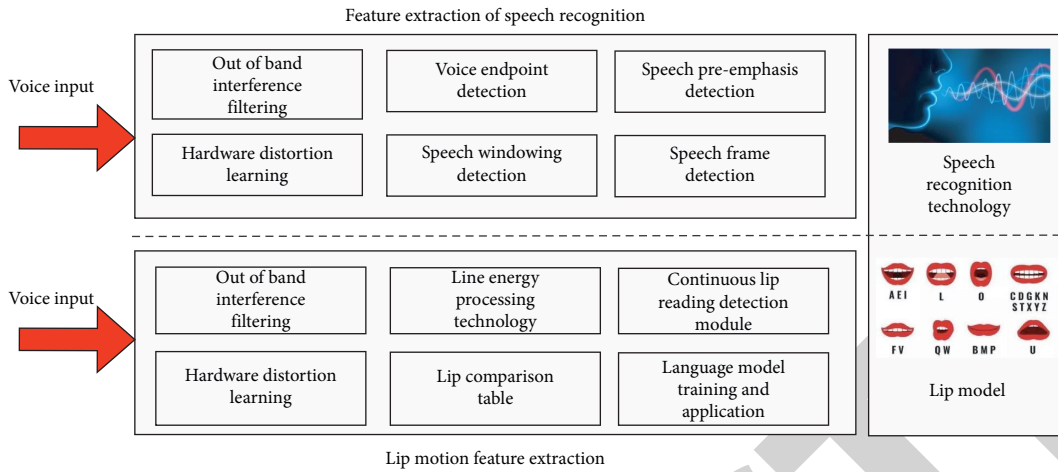
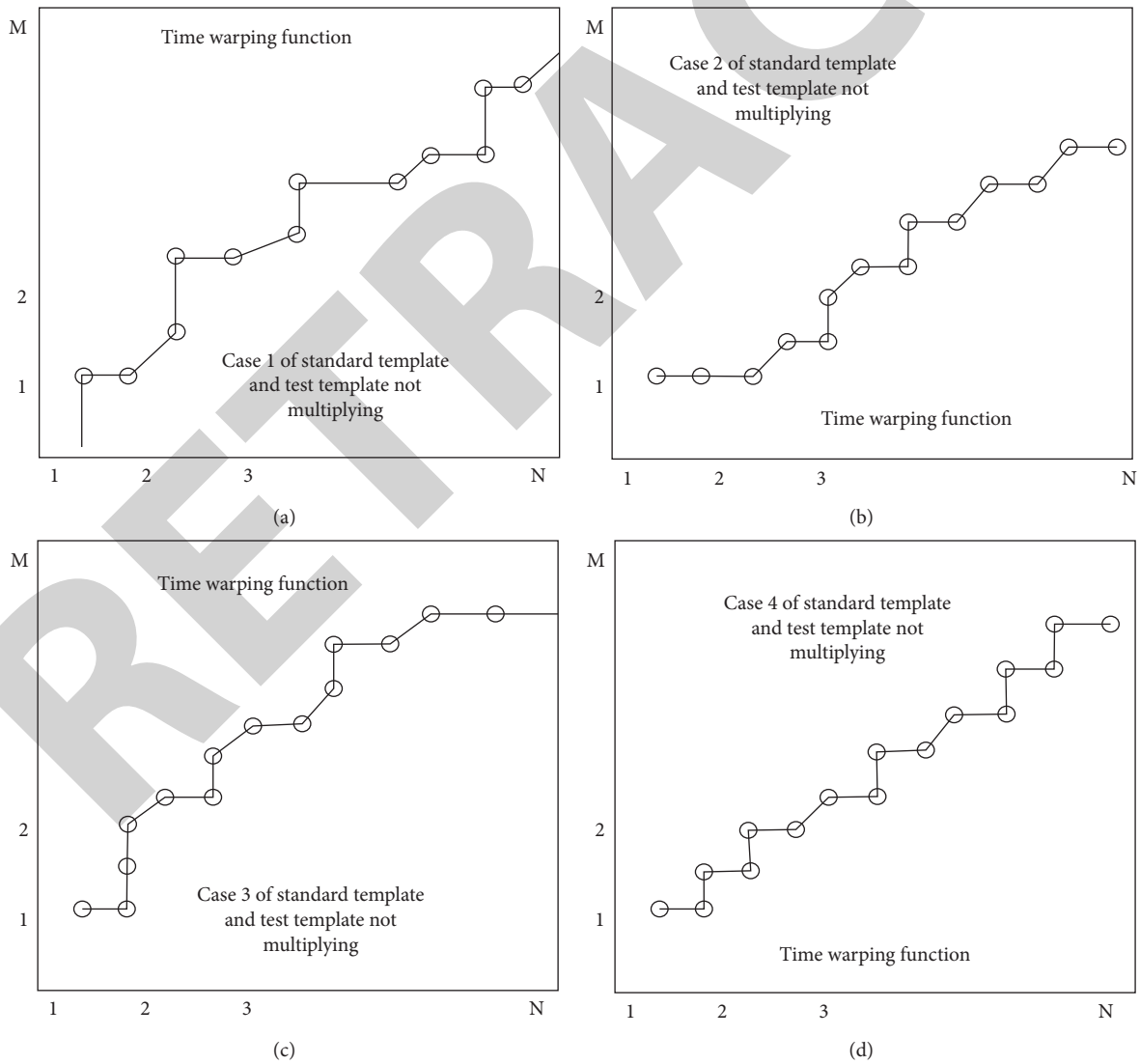FIGURE 3: Schematic diagram of feature extraction of the speech recognition and lip recognition algorithm.

FIGURE 4: Schematic diagram of the dynamic time adjustment algorithm.

$$d1 = \min(d12, d13)$$
$$d2 = \min(d13, d14)$$
$$d3 = \min(d23, d24)$$
$$\ldots\ldots\ldots\ldots$$
$$dn = \min(d1n-1, d1n)$$

$$D = (d1 + d2 + d3 \ldots\ldots + dn)/n$$

Double benchmark computer system

Frame matching distance

Average distance calculation

Evaluation score conversion mechanism

Judgment analysis

Benchmark strategy: it mainly compares the corresponding frame matching distances of the corresponding feature parameters. The corresponding frame matching distances in the figure are d1, d2 … dn, respectively. The frame distance between the corresponding learner and the corresponding standard pronunciation can be represented by the average distance.
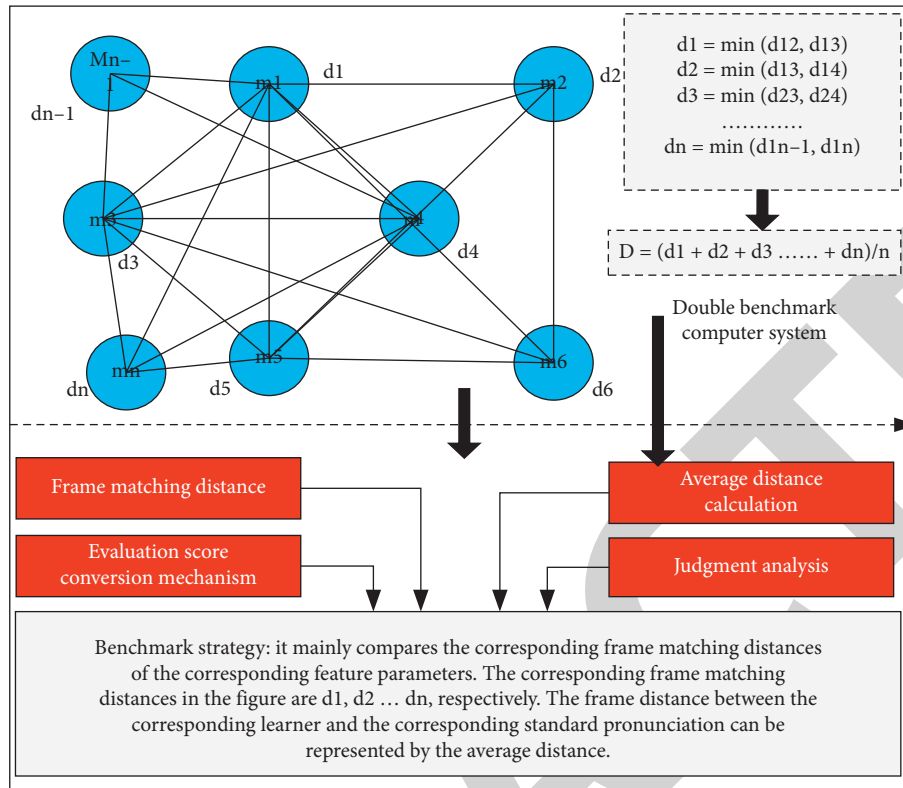
FIGURE 5: Evaluation mechanism of oral English pronunciation training based on the double standard evaluation mechanism.

of speech recognition and lip recognition algorithm feature extraction.

The dynamic time adjustment algorithm is mainly to solve the difference between the actual input speech and the speech evaluation system. In this paper, the dynamic time algorithm is used to stretch or compress the corresponding speech input until it is consistent with the reference. The core principle of the dynamic time adjustment algorithm used in this paper is shown in Figure 4. It can be seen from the figure that when the dynamic time adjustment algorithm is running, it mainly selects the corresponding reference speech template and the corresponding test speech template for score analysis. It directly projects the corresponding frame number into the corresponding coordinate axis x and the corresponding coordinate axis y; then the corresponding frame number that corresponds to the each intersection point in the corresponding grid can be defined as the corresponding matching degree. At this time, the corresponding dynamic time warping algorithm is to find the best path for the starting point and the endpoint to pass through each intersection point at the same time and ensure that the corresponding measure of the frame distance of each intersection point on the corresponding path is the minimum.

In the corresponding speech evaluation mechanism, two different standard pronunciation templates (corresponding to template 1 and template 2) are used for feedback comparison. The corresponding schematic diagram is shown in Figure 5. It can be seen from the figure that the spoken English speaker's corresponding pronunciation is test 3. In the actual feedback comparison, the frame matching

distance of the corresponding characteristic parameters between the two is mainly compared, and the corresponding frame matching distance is compared. The frame matching distances in the graph are D1, D2, and D3, respectively. The frame distance between the corresponding learner and the corresponding standard pronunciation can be expressed by the average distance. Based on the corresponding average distance, it is converted into the corresponding evaluation score, and the corresponding pronunciation level of the speaker is calculated.

Based on the analysis and research of the above principles, the core algorithm of oral English pronunciation training in this paper is basically constructed. At the same time, the problems of poor anti-interference ability of traditional speech recognition algorithms are solved. The corresponding dual speech feature extraction technology uses a hybrid algorithm of conventional speech recognition and lip motion recognition, which improves the accuracy of the whole speech recognition. The dual benchmark evaluation mechanism can significantly improve the judgment of the speaker's speech level, and the corresponding judgment standard can also help to improve the speaker's own oral level.

3.2. Research and Design of Oral English Pronunciation Training Based on Android Smartphone Platform. In order to verify the realizability of the algorithm, this paper designs and verifies the algorithm based on Android mobile platform. The main purpose is to realize the intelligent English
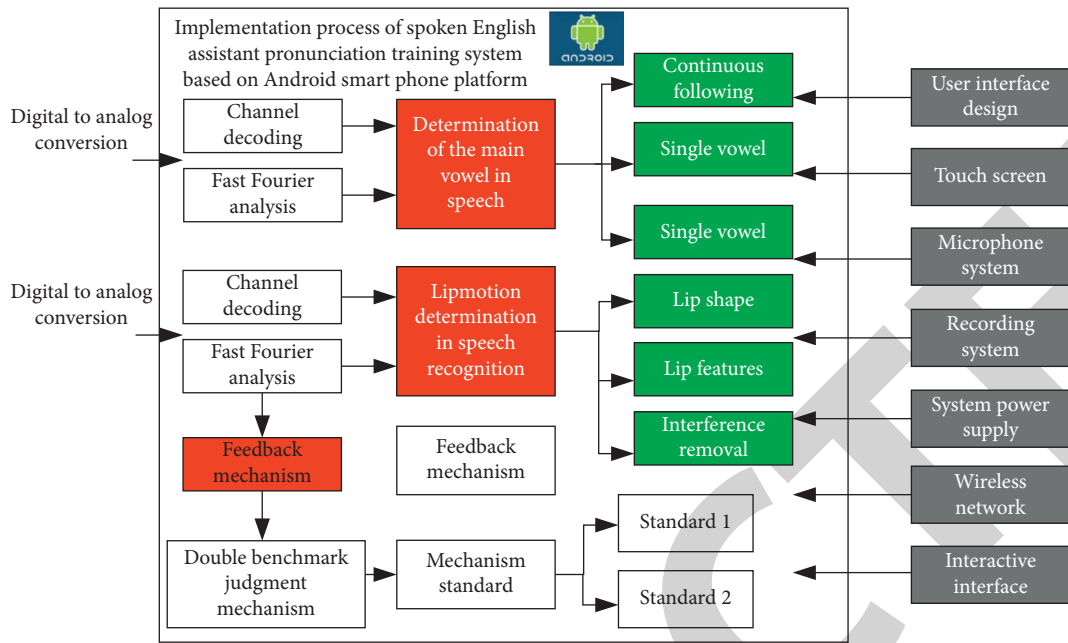
FIGURE 6: Implementation block diagram of spoken English assistant pronunciation training system based on Android smartphone platform.
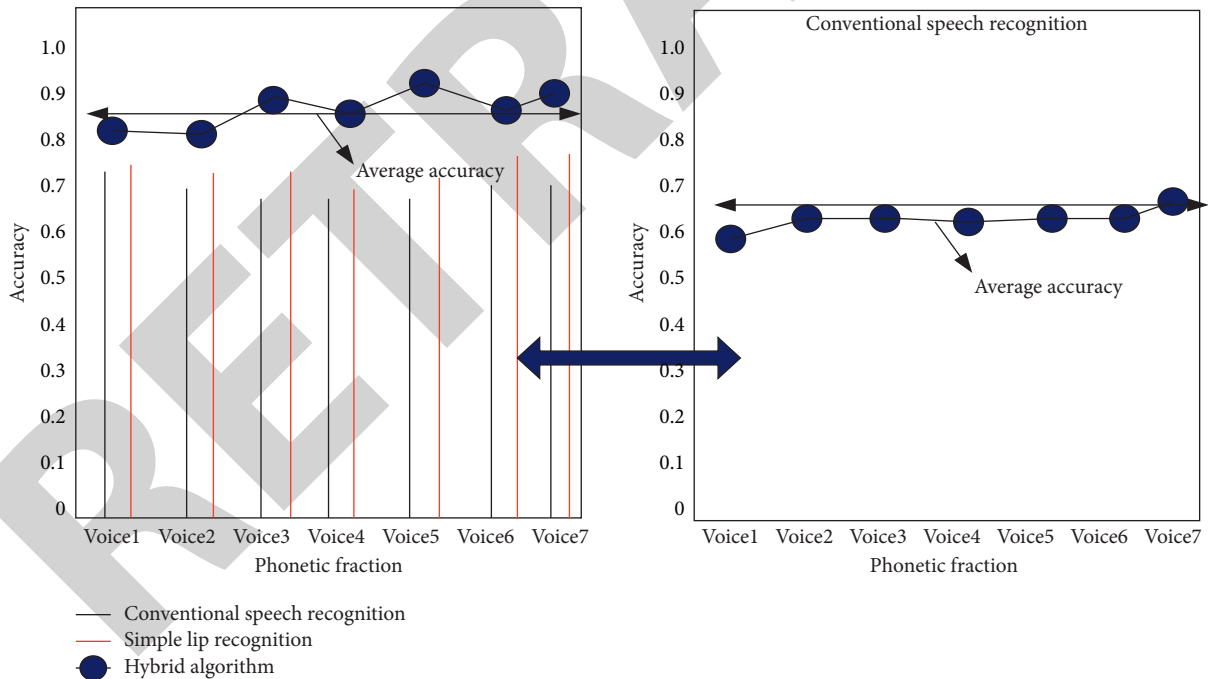

FIGURE 7: Comparison of average speech recognition accuracy of test samples.

pronunciation training system in the form of animation, sound assistance, pictures, and text through the Android mobile platform. The corresponding system organization module diagram is shown in Figure 6. It can be seen from the figure that the corresponding hardware design level mainly includes I/O module design, scoring module design, feedback module design, and user interface design. The

corresponding I/O module design mainly uses the built-in microphone and headset of Android mobile phone system to realize the input, sampling, and processing of voice signal; the corresponding scoring module design level mainly includes the double benchmark scoring mechanism, in which the corresponding scoring parameter generation and the corresponding score conversion mechanism are very
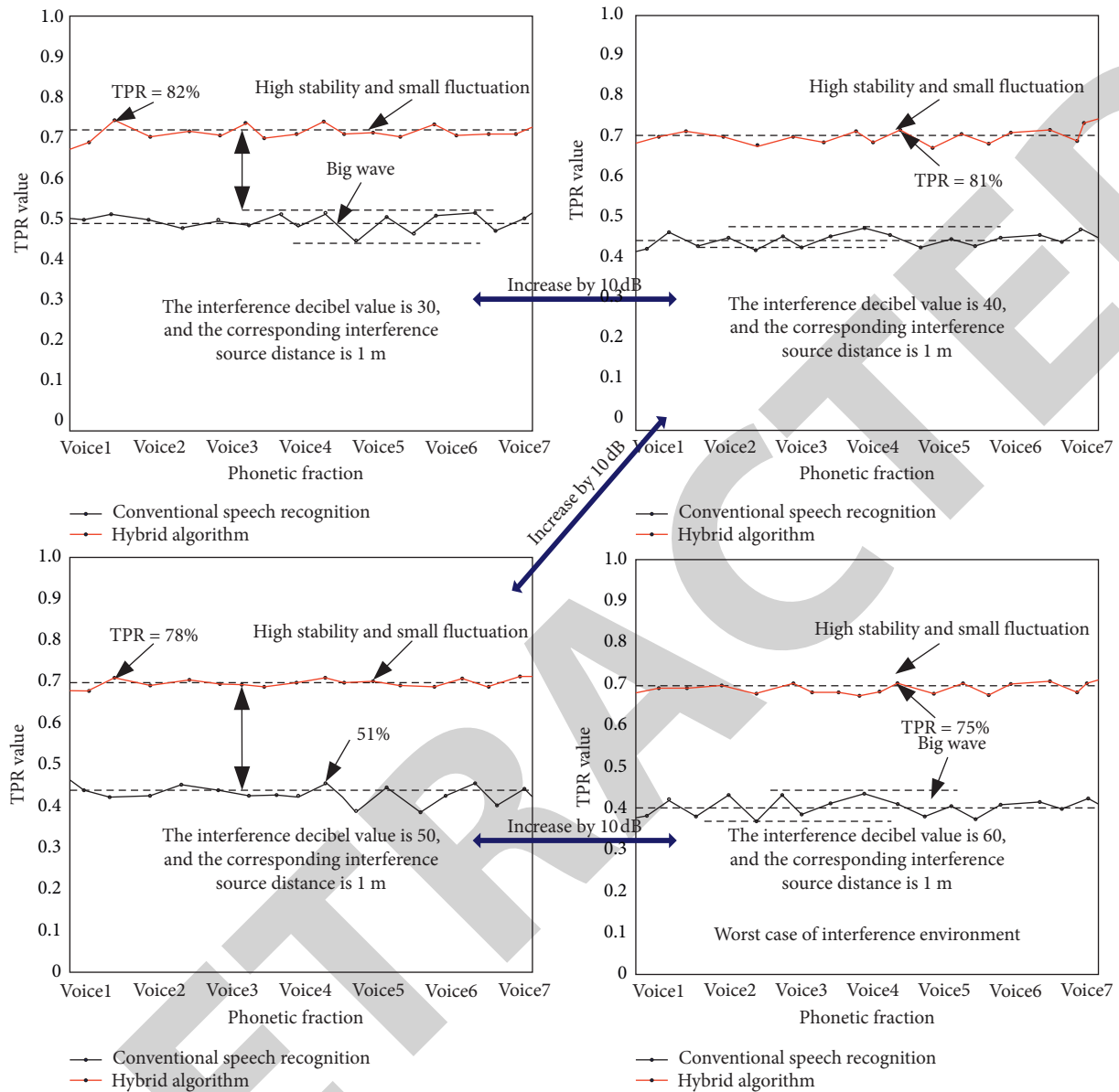
Figure 8: Comparison of anti-interference performance of test samples.

important; in the corresponding feedback module design level, two hardware processing modules are mainly used, which are, respectively, the Fourier transform processing of speech signal and the lip movement processing of the speaker, and the comprehensive signal is processed in the corresponding way through the specific processing mechanism. Graphical form is sent to Android mobile platform for display; at the same time, based on this, the formant contrast diagram of mixed processing signal is generated by using chart engine open source software; the corresponding user interface is mainly developed based on Android platform, and the corresponding function keys of voice assistant system are mainly set on the corresponding functions.

At the software level, the corresponding development environment is mainly implemented in the eclipse integrated environment. The corresponding development

environment is as follows: the PC uses Windows XP system; the development components use JDK 6; the platform hardware environment uses a brand of Android smartphone; the software platform uses Android OS2.2; and the programming language uses Java. Taking the development of image reference function in the corresponding scoring mechanism as the template, the corresponding core code is as follows, and the corresponding chart activity is the graphic drawing class function.

Intent intent = new Intent(FunctionTraining, this, ChartActivity. class);

Bundle bundle = new Bundle();

Bundle. putSerializable{"input_p_date"},

Factory.get_input_p_data();

TABLE 1: Comparison of average scores of oral English before and after training.

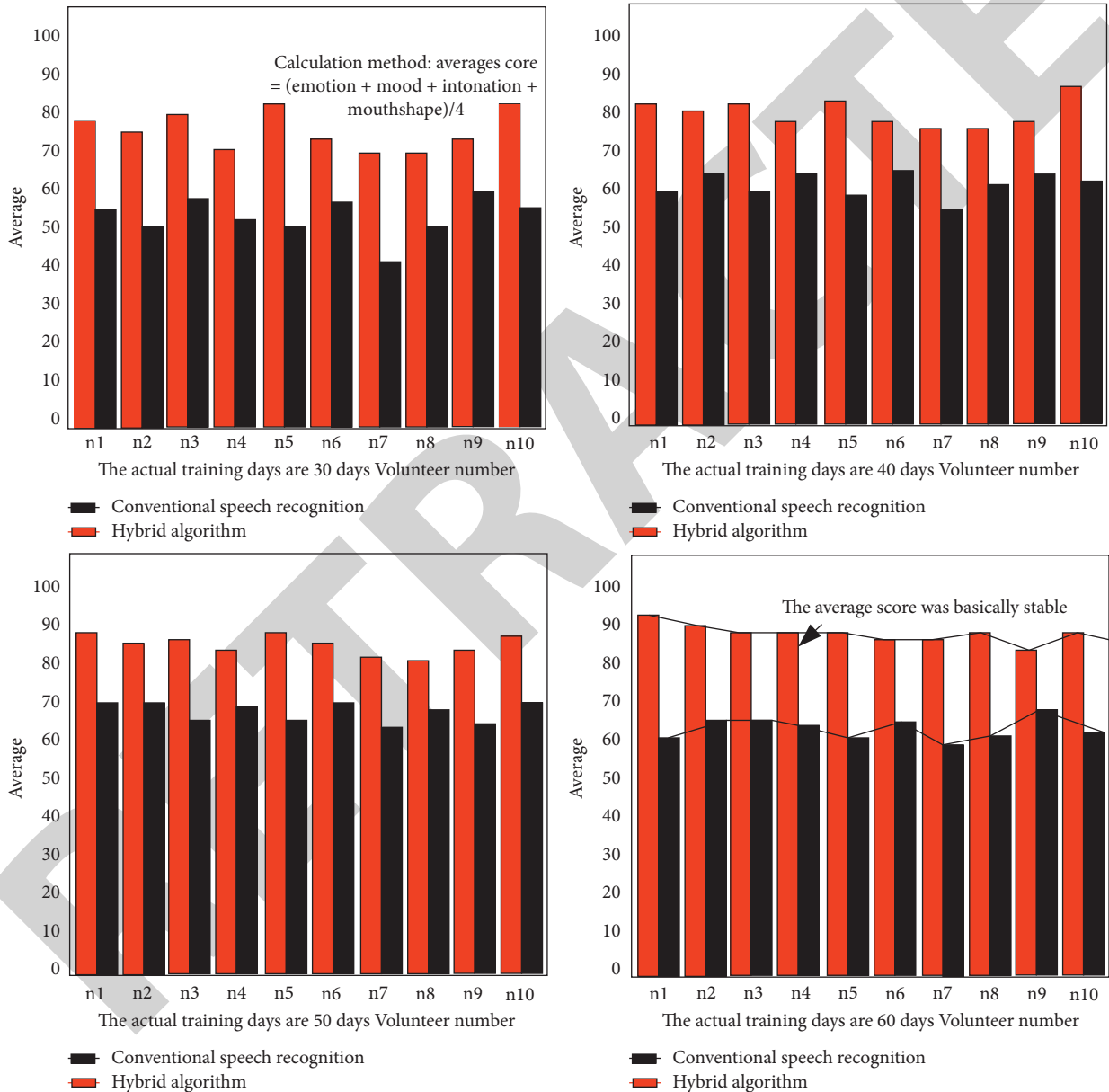| Test situation	Various indicators of oral English evaluation				Divide equally
Before training	Intonation 90	Emotion 85	Mouth shape 90	Tone 85	87.5
After training	Intonation 70	Emotion 75	Mouth shape 76	Tone 70	72.75

FIGURE 9: Comparison of the oral English pronunciation training effect test.

```
Bundle.putSerializable{"file_p_data",factor-
y.get_f_p_data{}};
Bundle.putInt{"input_frames",factory.get_input_-
frames{}};
```

```
Bundle.putInt{"file_frames",factory.get_input_frames
{}};
Intent.putExtra{"b",bundle};
startActivity(intent);
```

Based on the above core code, other similar functions are developed.

4. Experiment and Analysis

In order to verify the advantages of this algorithm compared with the traditional spoken English pronunciation training algorithm, this paper carries out a comparative implementation, and the corresponding experimental platform is an Android mobile phone. The corresponding experimental level is mainly from the three levels of comparative analysis that are speech recognition rate test, environment anti-interference test, and oral English pronunciation auxiliary training effect test.

4.1. Speech Recognition Rate Test Experiment. Before the experiment, the recorded standard speech database was put into the system, and 10 volunteers were used for speech test. In this test, the corresponding environmental variables were controlled (the environmental noise level was controlled to keep the same level). The average speech recognition accuracy of the test samples is shown in Figure 7. It can be seen from the figure that the corresponding speech recognition accuracy of this algorithm is significantly higher than that of traditional speech recognition, which shows that the speech recognition algorithm combining speech recognition with lip motion detection has obvious advantages.

4.2. Environmental Anti-Interference Performance Test. In this experiment, 60 dB music is selected as the environmental noise, and the corresponding noise source distance is consistent in the corresponding environment. Based on this environment, the environmental antijamming performance of this algorithm and the traditional algorithm is tested, and the evaluation index is the TPR value of the system. The contrast curve of the environment anti-interference performance corresponding to the test samples is shown in Figure 8. It can be seen from the figure that the TPR value of the corresponding system under the proposed algorithm is about 75%, which is far higher than the traditional algorithm.

4.3. Oral English Pronunciation Training Test. The purpose of this test is to test the effect of oral English improvement for oral English speakers. The corresponding training effect is transformed into the corresponding oral English test scores for quantitative evaluation. Ten volunteers from nonnative English-speaking countries are selected for the test. The experiment is based on the improvement of oral English scores of the trainees before and after the test and the corresponding experimental week. The training period is set at 30, 40, 50, and 60 days. The average oral English scores of the volunteers before and after the training are shown in Table 1. It can be seen from the table that the volunteers have made great progress in different indicators of oral English after the training. Figure 9 provides a detailed comparison chart of the effect of oral English pronunciation training in

four cycles. From the chart, it can be seen that the algorithm in this paper has obvious advantages, and these advantages will become more and more obvious with the growth of training time.

Comprehensive analysis shows that the proposed algorithm and the designed system have obvious advantages in improving the oral pronunciation effect compared with the traditional system.

5. Summary

This paper mainly analyzes the disadvantages of traditional oral English pronunciation practice and points out that the corresponding core algorithm has serious noise problems. Based on the analysis of the current research status of oral English pronunciation assistance, this paper develops and designs oral English pronunciation assistance training system based on Android smartphone platform. Based on the in-depth research and analysis of oral English pronunciation correction algorithm and voice feedback mechanism, this paper proposes a lip movement judgment algorithm based on ultrasonic detection, which is used to assist oral English pronunciation training. The traditional speech recognition algorithm makes double feedback judgment. In the feedback mechanism of intelligent speech training, a double benchmark scoring mechanism is introduced to evaluate the pronunciation of speech trainers in an all-round way and correct the pronunciation of speakers in time. The experimental results show that the pronunciation correction rate of the proposed system reaches 85%, which improves the level of oral English trainers to a certain extent. In the follow-up research, this paper will focus on the application of the speech recognition hybrid algorithm in other language learning and further promote the use of it.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] X. Y. Liu and J. Shen, "Design of English pronunciation training system for android," *Applied Mechanics & Materials*, vol. 631-632, no. 4, pp. 427–430, 2014.

[2] Y. Jin, "Design of students' spoken English pronunciation training system based on computer VB platform,"

International Journal of Emerging Technologies in Learning (iJET), vol. 14, no. 6, pp. 41–52, 2019.

[3] T. Li, J. Sun, and L. Wang, "An intelligent optimization method of motion management system based on BP neural network," *Neural Computing and Applications*, vol. 10, no. 6, pp. 41–52, 2020.

[4] K. Z. Chen, N. Johnson, S. Dai et al., "Contextual policy enforcement in android applications with permission event graphs," *Heredity*, vol. 110, no. 6, p. 586, 2013.

[5] Z. Yin, "Training & evaluation system of intelligent oral phonics based on speech recognition technology," *International Journal of Emerging Technologies in Learning*, vol. 13, no. 4, pp. 11–23, 2018.

[6] S. Lee, H. Noh, J. Lee, K. Lee, and G. G. Lee, "Foreign language tutoring in oral conversations using spoken dialog systems," *IEICE Transactions on Information and Systems*, vol. E95.D, no. 5, pp. 1216–1228, 2012.

[7] J. Zhou, "Construction of artificial intelligence-based interactive oral English teaching platform based on application problems of present intelligent products," *IOP Conference Series: Materials Science and Engineering*, vol. 569, no. 4, p. 052055, 2019.

[8] S. Lee, H. Noh, J. Lee et al., "Foreign language tutoring in oral conversations using spoken dialog systems," *IEICE Transactions on Information & Systems*, vol. 95, no. 5, pp. 1216–1228, 2017.

[9] H. Nakawala, G. Ferrigno, and E. De Momi, "Development of an intelligent surgical training system for Thoracentesis," *Artificial Intelligence in Medicine*, vol. 84, no. 4, pp. 50–63, 2018.

[10] W. Yang, L. Jin, and M. Liu, "DeepWriterID: an end-to-end online text-independent writer identification system," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 45–53, 2016.

[11] J. J. Castro-Schez, C. Glez-Morcillo, J. Albusac et al., "An intelligent tutoring system for supporting active learning: a case study on predictive parsing learning," *Information Sciences*, vol. 544, no. 5, pp. 23–31, 2020.

[12] C. Yang, Z. Deng, K.-S. Choi, and S. Wang, "Takagi-sugeno-kang transfer learning fuzzy logic system for the adaptive recognition of epileptic electroencephalogram signals," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1079–1094, 2016.

[13] H. Akkermans, "Intelligent e-business: from technology to value," *IEEE Intelligent Systems*, vol. 16, no. 4, pp. 8–10, 2001.

[14] J. Ruby, S. Daenke, Y. Yuan et al., "Automatic speech recognition and machine learning for robotic arm in surgery," *The American Journal of Surgery*, vol. 2, no. 1, pp. 10–18, 2020.

[15] G. R. Kidd and L. E. Humes, "Speech recognition for multiple bands: implications for the speech intelligibility index," *The Journal of the Acoustical Society of America*, vol. 11, no. 2, pp. 112–123, 2016.

[16] J. A. Beyea, K. P. Mcmullen, M. S. Harris et al., "Cochlear implants in adults: effects of age and duration of deafness on speech recognition," *Otology & Neurotology*, vol. 37, no. 9, pp. 1–12, 2016.

[17] J. Shehorn, O. Strelcyk, and P. Zahorik, "Associations between speech recognition at high levels, the middle ear muscle reflex and noise exposure in individuals with normal audiograms," *Hearing Research*, vol. 392, no. 4, p. 107982, 2020.

[18] F. De Wet, N. Kleynhans, D. Van Compernolle et al., "Speech recognition for under-resourced languages: data sharing in hidden Markov model systems," *South African Journal of Science*, vol. 113, no. 1, pp. 11–23, 2017.

[19] A. Sarma and K. S. Kandarpa, "Machine learning based sample extraction for automatic speech recognition using dialectal Assamese speech," *Neural Networks: The Official Journal of the International Neural Network Society*, vol. 78, no. 4, pp. 97–111, 2016.

[20] W. Wei, Z. Wang, C. Fu et al., "Intelligent recommendation of related items based on naive bayes and collaborative filtering combination model," *Journal of Physics: Conference Series. IOP Publishing*, vol. 1682, no. 1, p. 012043, 2020.

[21] A. Zielonka, A. Sikora, M. Wozniak, W. Wei, Q. Ke, and Z. Bai, "Intelligent internet of things system for smart home optimal convection," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, p. 4308, 2021.

[22] W. Wei, M. Guizani, S. H. Ahmed, and C. Zhu, "Guest editorial: special section on integration of big data and artificial intelligence for internet of things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2562–2565, 2020.