Research Article

Automatic Error Detection Method of Embedded English Speech Teaching Recognition System under the Background of Artificial Intelligence

Wenjuan He

College of Foreign Languages, Hengyang Normal University, Hengyang 421002, Hunan, China

Correspondence should be addressed to Wenjuan He; fayewei@hnu.edu.cn

Received 2 June 2022; Revised 27 July 2022; Accepted 4 August 2022; Published 12 September 2022

Copyright © 2022 Wenjuan He. 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.

The embedded English speech teaching recognition system is a technology that writes the English speech recognition device control program into the chip and embeds the chip into the device, so that the human-like chip controls the English speech device to complete the speech recognition operation. Applying the embedded technology to the English speech recognition system can improve the recognition accuracy of the system to a higher level in terms of recognizing the English speech of a specific person. The purpose of this paper is to research and design an automatic error detection method for embedded English speech teaching recognition system in the context of artificial intelligence. This paper first gives a general introduction to the overview of artificial intelligence, then analyzes the speech recognition algorithm, uses MatLab software to obtain the correct number of recognition system words and the correct rate, and then implements the embedded English teaching recognition system in different environments. The experimental results of comparison through multiple test analysis show that in a quiet environment, the error rate of the embedded English speech teaching recognition system is very low, and the correct recognition rate can reach more than 90%. In a noisy environment affected by various noises, the correct recognition rate of the embedded English speech teaching recognition system is basically above 60%.

1. Introduction

With the rapid development of computer and artificial intelligence technology, the application of intelligent mechanical equipment is becoming more and more extensive. As one of the key technologies of intelligent control, the ability of voice recognition control directly determines the degree of intelligence of mechanical equipment. The rapid development of artificial intelligence has promoted the rapid development of China’s education industry, making China’s English phonetic teaching more mature. In order to improve the teaching quality and teaching level, the English education industry actively introduces the embedded English pronunciation teaching system. The embedded English pronunciation teaching system has high English pronunciation recognition and strong control ability, and is suitable for popularization and application in the field of intelligent control. Because the embedded English pronunciation teaching system has high English pronunciation recognition and strong control ability. It is suitable for promotion and application in the field of intelligent control. In order to improve the teaching quality and teaching level, the English education industry has also actively introduced the embedded English pronunciation teaching system.

Today, with the rapid development of artificial intelligence technology, the implementation of embedded English phonetic teaching system teaching in the field of education can improve the overall teaching quality and teaching level, which also has far-reaching significance to the education field. The application range of artificial intelligence is very wide. In recent years, scholars have carried out research on English teaching in the context of artificial intelligence. However, there are relatively few research designs on the automatic error checking method of embedded English speech teaching...
recognition system. Therefore, this paper will carry out a research and design on the automatic detection method of the recognition error of the embedded English speech teaching recognition system under the background of artificial intelligence. This has both theoretical and practical significance.

With the advancement of science and technology in China, more and more people have conducted research on artificial intelligence English teaching. Among them, Misirov proposed that in the context of artificial intelligence, English pronunciation teaching in primary and secondary schools has changed a lot [1]. However, his lack of careful analysis of the data resulted in a less rigorous article. However, Vong and Kaewurai pointed out that the educational community can develop, implement, and evaluate a teaching model based on a cognitive approach. This enhanced students’ critical thinking and developed their ability to teach critical thinking to learners [2]. But his research is not based on real-world conditions, so the applicability is not high. Subsequently, Yin proposed to study the English embedded grammar-assisted teaching method and applied it to the design of the microclassroom model. The experimental results showed that embedded oral English teaching was in line with students’ academic conditions and teaching rules in senior high school English teaching [3]. However, the research and investigation he used was relatively widespread and cannot be specific to each regional teaching at present. Gashaw conducted acoustic analysis of speech recordings of four Amharic learners and two native English-speaking Canadians. He concluded that native samples of Amharic showed actual peaks on almost all words, requiring longer pronunciation times [4]. However, the examples he studied were too few, and the individuals were never representative of the group. Later, Chika et al. pointed out in his research that the use of multiple indicators for systematic phonetic evaluation was not only conducive to a deeper and broader understanding of Japanese and English, but also to the development of phonetics teaching [5]. However, his research did not elaborate on the elaboration of the speech system, so the article is not complete. After reviewing the above studies, Isaacs and Harding pointed out that the number of postgraduate students studying second language pronunciation and doing academic work in international universities had reached a record high. This offers broad prospects for second language pronunciation research and teacher training [6]. However, he did not cite strong data in the article to prove his theory, so the theory was not very successful. Later, Thompson showed in his research that through the analysis of key events in English classroom activities, he found that English dialogues can help improve English pronunciation [7]. But there are many articles on the subject similar to this kind of research, so his research is not novel enough.

The innovations of this paper are: In the context of artificial intelligence, it applies the embedded English pronunciation teaching system to English education research. In other research reports, the embedded English phonetic teaching system is often described as a new type of educational method. And this paper is devoted to dig deep into the research and design of the automatic error detection method of embedded English speech teaching recognition system, and test the speech recognition and processing ability of intelligent mechanical equipment. This has a certain reference value for the research of artificial intelligence-related technologies. This paper is devoted to the research and design of the automatic error detection method of embedded English speech teaching recognition system. It also uses MatLab software to conduct multiple sets of comparative tests on the recognition system to test the voice recognition and processing capabilities of intelligent mechanical devices. This has a certain reference value for the research of artificial intelligence-related technologies.

2. The Method of Automatic Detection of Errors in the Embedded English Teaching Recognition System

2.1. Overview of Artificial Intelligence. Artificial Intelligence' English abbreviation is AI. Artificial intelligence means talking about how to use computers to mimic human brain activity, and brain activities, such as reasoning, proving, identifying, understanding, involving, learning, thinking, planning, and problem solving. And it uses it to deal with complex problems that people can only solve. This special brain activity was only attainable by humans before, but now AI shows the technical level that artificial intelligence can basically reach based on knowledge, which realizes artificial intelligence imitating human brain activity.. These technologies are acquired through human intelligence but do better than human intelligence. This is also the level standard that AI needs to achieve in the future development [8].

AI is a very popular subject, which breaks people’s traditional concepts. It has changed the common way of thinking and promoted the absorption of human knowledge and the development of human education. In the field of education, emerging educational technology has always provided a strong force for educational reform. Teaching work has also become more convenient and efficient because of artificial intelligence. AI makes education fairer, more accessible, and more democratized. Teaching software with artificial intelligence can see, hear, speak, learn, and understand and help users’ emotions or moods similar to humans. It allows users to communicate with computers naturally and smoothly through language, text, gestures, expressions, and other methods to achieve human–computer interaction. In other words, artificial intelligence, space technology, and atomic energy technology are known as the three major scientific and technological achievements of the twentieth century. Because of the changes brought about by the continuous deepening of technology in the field of modern education and teaching, artificial intelligence will be widely used in the field of education. And it will have a profound impact on educational philosophy, teaching process, and teaching management.

2.2. Embedded Speech Recognition Algorithms

2.2.1. Algorithm Structure Model. By consulting relevant literature, we can obtain the mathematical model of the algorithm and the corresponding mathematical model of the
speech recognition algorithm [9]. Its structure is shown in Figure 1. It can be seen that the sound frequency enters from the impulse mode along with the noise. In order to ensure the successful conversion of the speech signal, the model must solve the output problem in the communication channel.

It can be drawn from the figure that the important function of this model is to ensure the output function $K(M)$ required when the speech signal is converted and solved, such as

$$H(M) = Q\cdot W(M)C(M)R(M). \quad (1)$$

Then, the analog signal of voice pronunciation is changed into a digital signal that can be used for communication, and it needs to be preparatively processed, as shown in the following formula:

$$H(M) = 1 - UM^{-1}. \quad (2)$$

Assuming that $Y(i)$ is the value compared with a certain time node $m$, then formula (3) can be obtained after preprocessing.

$$A(i) = Y(i) - LY(i - 1). \quad (3)$$

In the model, it first defines a set of state quantities $Q_1, Q_2, \ldots, Q_M$, at time point $M$, and the state of the model can only be one of these $Y_M$. At the starting point $M = 0$ time point, the initialization probability $\pi$ is

$$\pi_l = W[Y = Q_l] \quad (l = 1, \ldots, m). \quad (4)$$

It is followed by the transition probability matrix, so we get

$$B_l, h = W[Y_M = Q_l | Y_{M-1} = Q_l]. \quad (5)$$

But other time points $Y_M$ except $M = 0$ are hidden, so the formula can be obtained as

$$W_{YM} = Q_l[T_M] = W[T_M | Q_l]. \quad (6)$$

Fitting it again, we get

$$W[Y_M = Q_l, G_M = Y] = \sum_{n=1}^{N} D_nM(\mu_{n, \sigma_{n}}). \quad (7)$$

Assuming that there are $u$ English words, model $R_u$ is established for each word, and we can get

$$R_u = (E_u, F_u, \pi_u). \quad (8)$$

The final formula is

$$R^* = arg\ max_{0 \leq s \leq R} \{W(G | R)\}. \quad (9)$$

Here, the weighted combination of several Gaussian components $I_l$ of the speech signal is

$$J(y) = \prod_{l=1}^{m} F_l I_l(y). \quad (10)$$

The probability density function of a Gaussian mixture distribution containing $n$ components is a weighted combination of $n$ Gaussian probability density distribution functions, which is defined as

$$W(y | S) = \prod_{l=1}^{n} A_l W_t(y). \quad (11)$$

For the output probability of the input signal $E$ in the SCHMM state $K$, the formula is

$$v_l(E) = \prod_{g=1}^{q} A_{K,H,L}M(E_{K,H,L}, U_{K,H,L}, \sigma_{K,H,L}). \quad (12)$$

The probability of a sentence appearing in the model can also be expressed by the following formula:

$$Q(A_1, A_2, A_3) \cdots Q(A_M | A_j, A_{j+1}, A_{j+2}) = \prod_{l=1}^{n} A(A_{j+1} | A_j, A_{j+2}). \quad (13)$$

Through the model, the optimal starting point formula for the pronunciation error collection of the embedded English speech recognition system can be obtained:

$$H^*(n) = \min_t \{H^*(t) + S_n(t) + A\} \quad (14)$$

Through processing, the transfer formula can be obtained as

$$Y(M) = Q(M) + U(M)$$

$$= \frac{Z_{j-1}^{(l)} W(X_j | Y_{j-1}, X_{j-1}) Q(Y_{j-1} | Y_{j-1}, X_{j-1})}{P(Y_{j-1} | X_{j-1})}. \quad (15)$$

It decomposes the model and estimates the pronunciation vector to obtain the following formula:

$$Q(Y_{j} | \alpha) = \prod_{l=1}^{L} Q(X_j | \alpha, W, I). \quad (16)$$

Then, it can be obtained by eigen decomposition:

$$D_K = \sum_{l=1}^{M} D_{K,L}, \quad (17)$$

$$Q_{K,L} = D_{K,L}/D_K. \quad (18)$$

The pronunciation feature spectrum $A S_L$ of the embedded English recognition system can be obtained by using the time-frequency analysis method to locate

$$AS_L = - \sum_{k} Q_{K,L} ln(Q_K, L). \quad (19)$$

Finally, by analyzing the tone of the speech recognition system, it can be concluded that the output feature quantity is
By consulting the relevant information, it can be known that the chip used by the core processor of the system is very efficient in processing data. The execution of the dispatched instructions and the operation of the chip are enhanced by the instruction set in its architecture with increased efficiency and performance. The implementation efficiency of the issued instructions and the running performance of the chip are widely used in various industries [13].

2.2.2. Two-Stage Recognition Algorithm. In the embedded English speech recognition system, a single word template is not enough because of its sensitivity, and the detection time and space occupied will increase due to the increase of the number of candidate words, so the waiting time will increase. Therefore, it is rarely used in the embedded English speech recognition system [10]. The Monophone model is significantly different. It has the advantages of convenient templates, less waiting states, high reading efficiency, less space occupied, and has nothing to do with the number of words read. Even so, it is not accurate enough in pronunciation-related descriptions, resulting in a low subsequent recognition rate. Triphone and Syllable models are relatively accurate in terms of pronunciation-related performance, but there are many models, many states, slow recognition speed, and large space occupation [11]. In order to deal with both the space occupancy and the reading speed, a two-stage search algorithm is used in this paper. The basic process is shown in Figure 2.

It can be seen that in the first stage of reading, the model and static recognition network used have obtained a lot of candidate word entries. In the second stage of reading, another recognition network is built according to the multiple candidate entries given in the previous stage, and another model is used to perform accurate reading and obtain the final recognition result [12]. Because the number of entries read in the latter stage is relatively small, compared with the recognition in the previous stage, it is found that the reading speed is very fast. At the same time, the second-stage reading can reuse the space resources of the first stage, which also reduces the memory footprint of the recognition system.

2.3. Design Requirements of Embedded English Speech Recognition System

2.3.1. System Hardware Part Design. By consulting the relevant literature, the hardware structure of the embedded English speech recognition control system can be obtained as shown in Figure 3.
teaching resources other than textbooks, and complete the input of the text information to be converted. It also completes the input of the text information to be converted, and the English voice can be converted into the corresponding text information by using the embedded English speech recognition system. As for the embedded English speech recognition system technology, it can be converted into corresponding text information. The audio pronunciation standard recognized by the embedded English speech recognition system technology can accurately present the English pronunciation and make the learners feel the pure English pronunciation. This creates a good English listening learning environment for learners, and solves the problems of non-standard and insufficient voice input in current listening teaching, and some teachers "dialect English."

In addition, the multitimbral synthesis provided by the embedded English speech recognition system technology can meet the needs of various scenarios. This creates realistic situational teaching scenarios for learners. This can stimulate learners’ motivation to learn English listening and improve their learning initiative. It helps teachers to better carry out English listening teaching and helps to improve the quality of English listening teaching. Whether the quality of the English audio synthesized by the embedded English speech recognition system technology really reaches the level that can be used as a teaching demonstration sound in English listening teaching has not been confirmed by the data in previous scholars’-related research. Also, how to apply the functional advantages of the technology and the English listening teaching properly needs further research and analysis. Based on this, this research aims to use scientific research methods and means to analyze the practical application mode of embedded English speech recognition system technology. This not only stems from realistic demands but also has certain practical feasibility.
3. Simulation Experiment Analysis of Embedded English Teaching Recognition System Error Automatic Inspection

3.1. Embedded English Speech Recognition System Test. According to the actual demand, in order to test the sensitivity of the embedded English recognition system and test its error automatic checking function, we used MatLab software to conduct a simulation test. It is used to verify the systematic and accurate efficiency of the embedded English teaching recognition system [16]. Two experiments were carried out in this paper. In the speech recognition error accuracy test, 6 speeches were tested 30 times, and the results obtained are given in Table 1. From the first experiment in Table 1, it can be seen that the recognition rate of the embedded speech recognition system is relatively high, basically above 80%, which meets the design requirements of the intelligent recognition system. And within 30 tests, each word was correctly recognized more than 25 times. It can be shown that the performance of the system is very superior, the error value is low, and it is very suitable for English phonetic teaching research.

There may be insufficient rigor in an experiment. It uses MatLab software to carry out the second simulation experiment, which is the result of the second experiment, as presented in Table 2.

It can be seen from this that the difference between the results of the second experiment and the first experiment is not very big, and the accuracy of basically all words is above 80%. This well illustrates the feasibility of the embedded English speech recognition system, and proves the scientficity of the two experiments.

3.2. Test of Embedded English Speech Recognition System in Different Environments. In order to take into account the influence of various factors on the deviation value of the embedded English recognition system as much as possible, two experiments were carried out in quiet and noisy environments. It conducts 30 tests on the above six words to achieve the scientficity, accuracy, and authenticity of the experiment [17]. The tool used is also MatLab software. Figure 5 shows the results of the first experiment.

It can be seen from the figure that in a quiet environment, the embedded English speech teaching recognition system has a higher accuracy rate of recognizing words, basically reaching more than 80%. In a noisy environment, the embedded English speech teaching recognition system has a low accuracy rate of word recognition, but basically it has reached more than 66%. It can be known from this that if there are surrounding environmental factors, it will also lead to an increase in the error rate of the embedded English teaching recognition system. However, it is undeniable that the accuracy of the recognition system is still very high. And maybe in the future, the system will have better design and research to minimize the error value of the system.

A second experiment followed, in order to add more complexity and possibilities to the experiment, we added a different noise factor to the experiment in a noisy environment. And it was experimented with the previous fixed energy threshold method and two-stage detection method [18]. The results are given in Table 3.

It can be seen from the table that no matter in a quiet environment or under the influence of different noise factors, the two-stage detection method is obviously better than the endpoint detection method with a fixed energy threshold, and the recognition rate is higher [19]. And the two-stage detection method is also more scientific and accurate, which is very suitable for similar research.

Due to the long search time required for English speech, it is not possible to read the information after all the speech information has been received [20]. Because real-time reading is required, the system needs to implement reading features and first-stage recognition when acquiring voice signals. Because the vocabulary to be searched has not been completed, the final score against the vocabulary cannot be obtained. So, it is necessary to save the fitness score of each node for each vocabulary at the time of retrieval, which will bring another memory footprint [21]. For example, taking the three-column angle Gaussian model as an example, the voice database can be set as the command words of 10 girls. It can use these 10 command words to make a comparison of window width and recognition rate. The comparison of window width and recognition rate is depicted in Table 4. It is obvious that the window width data is 15, and the recognition rate is not reduced. Therefore, it can be concluded that the recognition rate is related to the candidate entries.

If these data are compared with the experiments of the first stage and the second stage, as given in Table 5. Compared with the first-stage reading, the second-stage identification uses a two-stage endpoint test method, feature selection algorithm, and beam search algorithm, which greatly increases the recognition rate and reduces space occupation and reading time [22].

Then, new data are obtained by comparing the identification system errors of different noise factors in a quiet environment. As shown in Figure 6.

Table 1: The first experiment accuracy test results.

| Identify the statement | Correct recognition times | Accuracy (%) |
|------------------------|---------------------------|--------------|
| Happy                  | 25                        | 83           |
| Cry                    | 26                        | 86           |
| Yes                    | 25                        | 83           |
| Fly                    | 28                        | 93           |
| Dog                    | 29                        | 96           |
| Like                   | 25                        | 83           |

Table 2: The second experiment accuracy test results.

| Identify the statement | Correct recognition times | Accuracy (%) |
|------------------------|---------------------------|--------------|
| Happy                  | 26                        | 86           |
| Cry                    | 25                        | 83           |
| Yes                    | 26                        | 86           |
| Fly                    | 28                        | 93           |
| Dog                    | 25                        | 83           |
| Like                   | 27                        | 90           |
It can be seen from the data graph obtained by the experiment that the error value of the recognition system in a quiet environment is the lowest, that is, the recognition rate is the highest. Under the influence of different noise factors, the error value of the recognition system is also different [23]. This shows that the impact of noise on the embedded recognition system is still a bit large. But even under the influence of different intensities of noise, the recognition system can still recognize words with an accuracy of more than 55%. It can be concluded that the feasibility of the embedded English speech recognition system in different environments is still worthy of optimism.

In order to increase the difficulty and reliability of the experiment, we believe that the recognition degree of only detecting words is not very comprehensive, so the error value of using the embedded English speech recognition system to recognize sentences is also added to this experiment. Similarly, the sentence experiment is also tested twice to ensure the accuracy of the experiment. The results obtained from the first test are shown in Figure 7.

It can be seen from Figure 7 that the recognition rate of the embedded English speech recognition system for sentences is still very high. The accuracy rate in quiet environments has also reached more than 80%, and each sentence has reached 25 correct times in 30 tests. This clearly demonstrates the power of the system’s performance. Although there is a slight difference with the recognition rate of words, the difference is not fundamentally different. This shows that the recognition system’s feasibility of automatic error detection on sentences is also very good.

| Test the voice                                      | Fixed threshold (%) | Two phase (%) |
|-----------------------------------------------------|---------------------|--------------|
| The original voice                                  | 92.75               | 93.24        |
| Add car noise (15 dB)                               | 88.73               | 89.78        |
| Add car noise (10 dB)                               | 67.99               | 74.78        |
| Add white noise (15 dB)                             | 86.67               | 88.67        |
| Add white noise (10 dB)                             | 63.56               | 75.34        |
| Add the sound of people talking (15 dB)             | 87.65               | 89.54        |
| Add the sound of people talking (10 dB)             | 78.34               | 79.23        |

| Table 4: Comparison of window width and recognition rate. |
|----------------------------------------------------------|
| Nbest | Do not add window (%) | Window width is 15 (%) | Window width is 10 (%) | Window width is 8 (%) | Window width is 5 (%) | Window width is 2 (%) |
|-------|-----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| 1     | 96.62                 | 96.62                  | 96.45                  | 95.50                  | 91.67                  | 65.45                  |
| 2     | 98.56                 | 98.59                  | 98.94                  | 97.45                  | 93.12                  | 67.28                  |
| 3     | 99.42                 | 99.12                  | 98.23                  | 98.34                  | 94.23                  | 68.09                  |
| 4     | 99.54                 | 99.45                  | 98.34                  | 98.12                  | 94.34                  | 68.55                  |
| 5     | 99.56                 | 99.34                  | 99.23                  | 98.24                  | 94.23                  | 68.87                  |

| Table 5: One-stage identification and two-stage identification. |
|---------------------------------------------------------------|
| Sentence recognition rate (%) | Recognition time (double real-time) | Memory usage (kB) |
| One-stage identification                          | 97.03            | 2.16             | 282              |
| Two-stage viterbi decoding                          | 96.45            | 0.45             | 62               |
| Two-stage beam search                                | 96.32            | 0.14             | 45               |
The second experiment is carried out below, and different noise influencing factors are added to the experiment, and the comparison diagram is shown in Figure 8.

As can be seen from Figure 8, the embedded English speech recognition system is not very accurate for sentence recognition under the influence of different noise factors, and does not have a high recognition rate for each word. But it has basically reached more than 53%. And it was tested 30 times, and each time it reached the correct number of more than 15 times. This also corresponds to a ratio of the error value of the test intelligence system. Moreover, in this experiment, we tested the recognition rate of the system in a quiet environment. The results are obvious, in quiet environment, the system recognition rate is still very high. It has been repeatedly tested, which reflects the very good performance of the system in automatic error detection.

3.3. Automatic Error Detection Test Analysis of Different Identification Systems. Due to the design of the error method of the recognition system studied in this paper, it can be seen whether the system can accurately recognize English speech in practical applications [24]. We use six English voices to compare and test the previous speech recognition system and the embedded system we need to study in different environments, namely, a quiet environment and a noisy environment. It uses a different system for each English voice and it is in a different environment. Each test was conducted 20 times, and the number of times the system successfully recognized English speech was recorded. The test results are shown in Figures 9 and 10.

From the test results in Figures 9 and 10, it can be seen that the recognition rate of the embedded English teaching recognition system is significantly higher than that of the traditional recognition system. Through calculation, it can be obtained that the correct recognition rate of the embedded recognition system can reach 90%, while the traditional recognition system can only reach 60%. It shows that the performance of automatic error detection of embedded English teaching recognition system is still very good.

4. Discussion

This paper is devoted to researching and designing the automatic error detection method of embedded English speech teaching recognition system under the background of artificial intelligence, and applying it to analysis and
| number | The first sentence | The second sentence | The third sentence | The fourth sentence | The fifth sentence | The sixth sentence |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 5      | Identification    | Statement 1       | Identification    | Statement 2       | Identification    | Statement 3       |
| 10     | Identification    | Statement 4       | Identification    | Statement 5       | Identification    | Statement 6       |
| 15     | Identification    | Statement 1       | Identification    | Statement 2       | Identification    | Statement 3       |
| 20     | Identification    | Statement 4       | Identification    | Statement 5       | Identification    | Statement 6       |
| 25     | Identification    | Statement 1       | Identification    | Statement 2       | Identification    | Statement 3       |
| 30     | Identification    | Statement 4       | Identification    | Statement 5       | Identification    | Statement 6       |

Figure 8: Test results of sentences under the influence of quiet environment and different noise factors. (a) Sentence recognition correct times in different environments, (b) sentence recognition accuracy in different environments.

Figure 9: Test results of embedded English speech teaching recognition system.

Figure 10: Traditional recognition system test results.
processing. It not only expands the application scope of embedded English pronunciation teaching, but also is a new attempt to automatically detect errors in the embedded English pronunciation teaching recognition system. In different environments (respectively, a quiet environment and an environment affected by different noise factors), it analyzes the accurate recognition times and correct recognition rates of multiple groups of words and sentences. It also compares the recognition rates of different recognition systems, so as to mine and study the automatic error detection method of the embedded English speech teaching recognition system. This research method has certain potential in the research of complexity of intelligent system. This has certain potential in the study of the complexity of intelligent systems. For the research on automatic error detection method of embedded English speech teaching recognition system, this paper starts from the basic introduction of artificial intelligence. It analyzes embedded speech recognition algorithms and two-stage recognition algorithms as well as the design requirements of embedded speech systems. It successfully outlines the main structure of the embedded English speech teaching recognition system clearly. It then uses MatLab software in the experimental analysis stage to obtain multiple sets of speech recognition rate comparison combinations, and analyzes these combinations in two aspects in different environments and different systems. The results show that the obtained results are in line with the actual situation.

Through the analysis of this case, it is shown that the automatic error detection function of the embedded English speech teaching recognition system is very effective and the feasibility is very high. In the specific practice test, it can be found that the error value of the embedded English speech teaching recognition system in a quiet environment is very low, and the accuracy rate is very high.

This paper takes a case study of the recognition rate of embedded English speech teaching recognition system in different environments. Firstly, through the introduction of speech recognition algorithm, the structure of embedded English speech recognition system is obtained by consulting data. By comparing the recognition rates of different environments and analyzing the data, it is concluded that in a quiet environment, the feasibility of automatic error detection of the embedded English speech teaching recognition system is very high.

5. Conclusions

Under the influence of different environmental factors, the error values of the embedded English speech teaching recognition system are different. Among them, in a quiet environment, whether it is a word or a sentence recognition test, the system has the highest recognition rate. Under the influence of noise factors, it should be determined according to the type of noise and the size of decibels. But basically, the recognition rate of the system under the influence of noise will be a little lower than that in a quiet environment, but this is not absolute. After all, with the rapid development of artificial intelligence technology today, perhaps there will be better improvements to the embedded English speech teaching recognition system in the future, and the embedded English speech teaching recognition system can have better performance in identifying errors.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest in this study.

Acknowledgments

This work was supported by Scientific Research Fund of Hunan Provincial Education Department (the Project Number: 20C0294).

References

[1] S. Misirov, “The peculiarities of teaching English pronunciation in elementary classes (GRADE)[J],” Scientific Bulletin of Namangan State University, vol. 1, no. 2, p. 63, 2019.
[2] S. A. Vong and W. Kaewurai, “Instructional model development to enhance critical thinking and critical thinking teaching ability of trainee students at regional teaching training center in Takeo province, Cambodia,” Kasetart Journal of Social Sciences, vol. 38, no. 1, pp. 88–95, 2017.
[3] Y. Yin, “Microclassroom design based on English embedded grammar compensation teaching,” Mathematical Problems in Engineering, vol. 2021, no. 1, pp. 1–9, 2021.
[4] A. Ghawsh, “Rhythm in Ethiopian English: implications for the teaching of English prosody[J],” International Journal of Education and Literacy Studies, vol. 5, no. 1, pp. 2202–9478, 2017.
[5] F. Chika, E. Sayoko, and A.-Y. Reiko, “An evaluation of English pronunciation of Japanese EFL learners using multiple metrics[J],” Journal of the Phonetic Society of Japan, vol. 22, no. 2, pp. 39–43, 2018.
[6] T. Isaacs and L. Harding, “Pronunciation assessment,” Language Teaching, vol. 50, no. 3, pp. 347–366, 2017.
[7] I. Thompson, “Stimulating reluctant writers: a Vygotskian approach to teaching writing in Secondary Schools,” English in Education, vol. 46, no. 1, pp. 85–100, 2018.
[8] A. Haleem, M. Javaid, and I. H. Khan, “Artificial Intelligence (AI) applications in dentistry,” Current Medicine Research and Practice, vol. 10, no. 1, pp. 36–38, 2020.
[9] E. M. Weisberg, L. C. Chu, and E. K. Fishman, “The first use of artificial intelligence (AI) in the ER: triage not diagnosis,” Emergency Radiology, vol. 27, no. 4, pp. 361–366, 2020.
[10] D. Bedzow Ira Leeds, “Artificial intelligence (AI) and halakhic responsibility for physicians to consult data[J],” Studies in Judaism, Humanities, and the Social Sciences, vol. 3, no. 1, pp. 90–103, 2020.
[11] S. Bi, R. Chen, K. Zhang, Y. Xiang, and H. Wang Lin Yang, “Differentiate cavernous hemangioma from schwannoma with artificial intelligence (AI),” Annals of Translational Medicine, vol. 8, no. 11, p. 710, 2020.
[12] D. Calaprice-Whitty, K. Galil, W. Salloum, and B. Jimenez, “Improving clinical trial participant prescreening with artificial intelligence (AI): a comparison of the results of AI-
assisted vs. standard methods in 3 oncology trials[J],” Therapeutic Innovation and Regulatory Science, vol. 54, no. 1, pp. 1–6, 2020.

[13] J. Gregory, S. Welliver, and J. Chong, “Top 10 reviewer critiques of radiology artificial intelligence (AI) articles: qualitative thematic analysis of reviewer critiques of machine learning/deep learning manuscripts submitted to JMRI,” Journal of Magnetic Resonance Imaging, vol. 52, no. 1, pp. 248–254, 2020.

[14] F. Schuur, M. Rezazade, M. H. Rezazade Mehrizi, and E. Ranschaert, “Training opportunities of artificial intelligence (AI) in radiology: a systematic review,” European Radiology, vol. 31, no. 8, pp. 6021–6029, 2021.

[15] C. Koo, Z. Xiang, U. Gretzel, and M. Sigala, “Artificial intelligence (AI) and robotics in travel, hospitality and leisure,” Electronic Markets, vol. 31, no. 3, pp. 473–476, 2021.

[16] E. Kotter and E. Ranschaert, "Challenges and solutions for introducing artificial intelligence (AI) in daily clinical workflow," European Radiology, vol. 31, no. 1, pp. 5–7, 2021.

[17] K. H. Ng and J. H. D. Wong, "A clarion call to introduce artificial intelligence (AI) in postgraduate medical physics curriculum," Physical and Engineering Sciences in Medicine, vol. 45, no. 1, pp. 1-2, 2022.

[18] M. K. Tripathi, A. Nath, T. P. Singh, A. S. Ethayathulla, and P. Kaur, "Evolving scenario of big data and Artificial Intelligence (AI) in drug discovery," Molecular Diversity, vol. 25, no. 3, pp. 1439–1460, 2021.

[19] V. Arvind, D. Kaji, J. Kim, J. M. Caridi, S. K. Cho, and K. Samuel, "Artificial intelligence (AI) can predict postoperative complications better than traditional statistical testing following anterior cervical discectomy and fusion (ACDF),” The Spine Journal, vol. 17, no. 10, pp. S145–S146, 2017.

[20] M. F. Byrne, N. Chapados, F. Soudan et al., "Su1614 artificial intelligence (AI) in endoscopy--deep learning for optical biopsy of colorectal polyps in real-time on unaltered endoscopic videos," Gastrointestinal Endoscopy, vol. 85, no. 5, pp. AB364–AB365, 2017.

[21] M. KauKau, "Der Aufstieg von Artificial Intelligence (AI) und seine Auswirkungen auf die Rechtsanwaltschaft - existenzenbedrohung oder willkommene Entlastung?” Computer und Recht, vol. 37, no. 7, pp. 498–504, 2021.

[22] S. Makridakis, "The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms,” Futures, vol. 90, no. jun, pp. 46–60, 2017.

[23] J. A. Belfor, I. S. Sena, D. K. B. d Silva, and B. E. Fd LopesKoga JúniorSantos, "Competências pedagógicas docentes sob a percepção de alunos de medicina de universidade da Amazônia brasileira," Ciência e Saúde Coletiva, vol. 23, no. 1, pp. 73–82, 2018.

[24] M. C. Pennington and P. Rogerson-Revell, "Using technology for pronunciation teaching, learning, and assessment," English Pronunciation Teaching and Research, vol. 2019, pp. 235–286, Chapter 5, 2019.