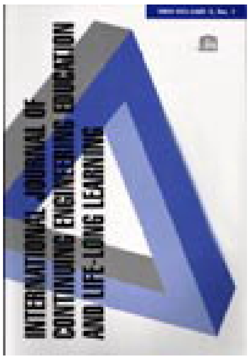


International Journal of Continuing Engineering Education and Life-Long Learning

ISSN online: 1741-5055 - ISSN print: 1560-4624
https://www.inderscience.com/ijceell

Artificial intelligence-based emotion recognition application of English teaching in smart learning

Cheng Huang

DOI: 10.1504/IJCEELL.2025.10073221

Article History:

Received:	18 January 2025
Last revised:	09 April 2025
Accepted:	27 June 2025
Published online:	10 October 2025

Artificial intelligence-based emotion recognition application of English teaching in smart learning

Cheng Huang
Department of Foreign Language Teaching,
Hainan Vocational University of Science and Technology,
Haikou 571100, Hainan, China
Email: hctotoo@outlook.com

Abstract: With the rise of the intelligent era, innovative learning has gained increasing attention, particularly regarding students' needs. While English teaching has moved away from the 'dumb English' approach, focusing more on integrating listening, speaking, reading, and writing, many still view English as a subject rather than a language, affecting teaching effectiveness. Due to spatial and temporal limitations, emotional interaction between teachers and students is lacking. This study explores an AI-supported emotion recognition teaching model, integrating relevance, originality, and impact (ROI) theory with innovative English education. An echo state network was constructed, and the algorithm was optimised. Emotion classification and speech signal preprocessing were implemented. Experimental results show improvements in students' performance in vocabulary (3.8%), listening (4.5%), reading (5.9%), and speaking (7.1%) compared to traditional methods, enhancing smart learning quality and classroom interaction.

Keywords: artificial intelligence; smart learning; English teaching; emotion recognition.

Reference to this paper should be made as follows: Huang, C. (2025) 'Artificial intelligence-based emotion recognition application of English teaching in smart learning', *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 35, No. 8, pp.113–128.

Biographical notes: Cheng Huang is a Lecturer in Hainan Vocational University of Science and Technology. She holds a Master's degree from Hainan University and a Doctor's degree from Dhurakij Pundit University in Thailand. Her research interests include English education.

1 Introduction

Emotion recognition is widely used in the new field of human-computer interaction, and it is suitable for the teaching of intelligent learning. Through the sensitive interactive voice teaching system, students' learning emotions can be identified. According to the students' learning mood and learning results, teachers can better understand the teaching effect and overall progress, to quickly and effectively adapt to teaching actions and improve teaching quality. Facial expressions are the most direct and effective way to express emotions in the process of communication in daily life. The automatic recognition of speech emotion is of great significance for the intelligentisation of human-computer interaction. Its development and implementation can effectively improve people's work efficiency, learning efficiency, and quality of life, and produce huge economic and social benefits. Therefore, emotion recognition has great potential for development and has important scientific significance, academic value, and practical application value. Emotion recognition research includes knowledge of speech, cognitive psychology, signal processing, artificial intelligence, and other disciplines. This is a new field of research. Although important research results have been achieved, the technology is still in the development process due to the short research time. It needs more researchers to further improve in theory and practical application.

At present, the good state of community informatisation development provides a rare development opportunity for the integration of learning resources. Students have a strong thirst for knowledge and are increasingly aware of the importance of learning resources. Learning resources are closely related to information resources. The constructed emotion model should not only correspond to facial expressions and emotions but also reflect the positional relationship between different emotions, to analyse and understand the current emotions of individuals by using the acquired expression information. Learning resources are a more specialised classification of information resources. The integration of information resources involves the integration of learning resources. Research on the integration of learning resources can also prevent the phenomenon of information silos. Therefore, the management and service of the integration of learning resources would enter a new stage and would be endowed with new ideas. This paper aims to explore the application mechanism and effect of artificial intelligence-supported emotion recognition English teaching screen in the smart learning environment, analyse the needs and goals of English teaching based on ROI theory, and build an academic emotion recognition model in combination with the echo state network in artificial intelligence, to optimise the feedback mechanism of the smart teaching screen and dynamically adjust the teaching content and interaction methods.

2 Related work

To improve the efficiency of cross-cultural English teaching, a cross-cultural O2O English teaching system with intelligent recognition and management functions was constructed with the support of artificial intelligence emotion recognition and neural network algorithms. In addition, the background model was used to detect and track objects, and the full recognition of students' emotions was achieved, enabling teachers to effectively control online teaching. In addition, according to actual needs, an overall O2O English teaching model was constructed and a corresponding teaching process was formulated. Lin et al. (2021) took high-speed rail signals and communication technology as the background. With the application of the new generation of railway mobile communication technology, some advanced communication technologies have appeared in the form of English teaching materials. Therefore, the bilingual teaching model is of great significance and helps to cultivate research talents with international competitiveness. Through the reform of bilingual mixed teaching, the proportion of research topics has changed from 6% to 15%. The wide range of application possibilities has led to an increase in the application of emotion recognition in the field of computer science. Using non-verbal cues such as gestures, body movements, and facial expressions to convey feelings and feedback to users can achieve basic human-computer interaction. He introduced various methods and techniques for emotion recognition and reviewed a dataset that was considered an algorithm for detecting emotions through facial expressions. The sensors he designed can play an important role in accurate detection by providing very high-quality input, thereby improving the efficiency and reliability of the system (Mehta et al., 2018). Coskun (2019) reported on the continued implementation of the facial emotion recognition and empathy test (FERET), a powerful and effective tool for assessing the ability of primary school children to recognise and empathise with facial emotions. The researchers developed human face images as response options for children, constructing a set of response options and the resulting interest scale from a two-factor coordinated emotion structure representing the emotional dimensions. The subjects were 422 primary school students, and the children were asked to recognise emotions and make appropriate empathetic responses. The data were analysed by item analysis, exploratory factor analysis (EFA), and item response theory (IRT). It can be seen that it is difficult to conduct research solely on the topic of student emotion recognition in English teaching, and the intelligent learning of artificial intelligence will simplify the experimental construction.

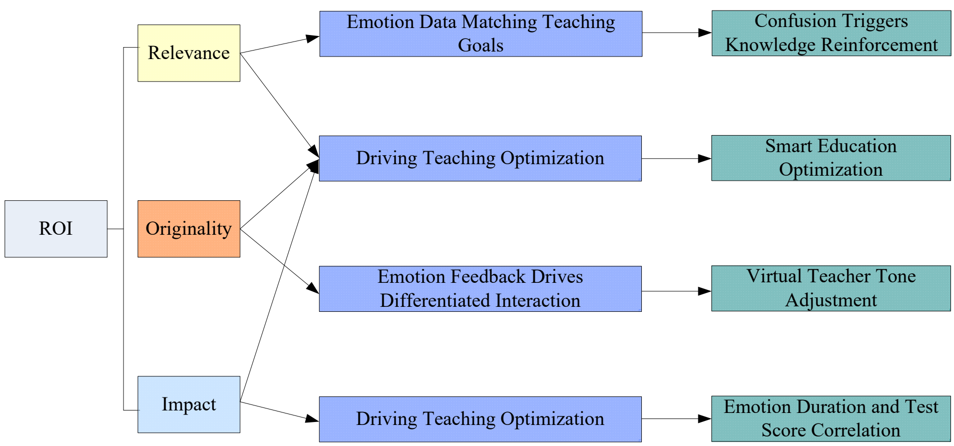
In terms of intelligent learning, Al-Nakhal and Naser (2017) proposed an intelligent teaching system designed to help students learn computer theory. He built the intelligent teaching system using the ITSB authoring tool. The system helps students learn finite automata, pushdown automata, and Turing machines. In addition, the relationship between these automata and formal languages, deterministic and non-deterministic machines, regular expressions, context-free grammars, undecidability, and complexity is studied. In the process, the intelligent tutoring system provides multiple types of help and feedback in an intelligent way based on the student's behaviour. The evaluation of the intelligent teaching system shows that its results are quite acceptable in terms of usability and learning ability. Berta and Moreno-Ger (2018) conducted a study on intelligent learning assessment in serious games. It is generally agreed that serious games (SG) have the potential to be an educational tool, and the focus of the discussion is no longer on whether it is necessary to use SG in education. It is a revolution in how to use its potential to effectively evaluate the learning progress of players. In this context, he aims to introduce some recent advances in the field of learning assessment. Ma et al. (2018) believed that with the popularisation of e-learning, learning support systems (LSS) play an increasingly important role in teaching and education reform. Although pure resource-based LSS is prevalent, it can no longer meet the needs of learning support in the era of big data and massive resources. He optimised LSS from five aspects. It includes a micro-resource package structure oriented to knowledge points, knowledge structure reorganisation of learning resources, learning behaviour monitoring and feedback guided by learning progress visualisation theory, resource recommendation based on learning progress and knowledge structure, and reuse of high-quality discussion posts and questions. Teaching practice has proved that the structured design of resource packages can stimulate imagination.

3 Methods of emotion recognition in the context of artificial intelligence

3.1 ROI theory

ROI theory plays an important guiding role in advertising creative design (Xu and Sun, 2017). That is to say, a good advertisement should emphasise the three basic conditions of creative design, namely relevance, originality, and impact (Won, 2018). In creative instructional design, there are higher requirements for relevance, originality, and impact (Razek and Frasson, 2017). If the presentation of knowledge content (images, words, and sounds) in intelligent learning English teaching has nothing to do with the knowledge itself and the learning needs of students, the importance of teaching would disappear. In the form in which knowledge is presented, teaching is not original but unified. It leads to the loss of the 'vigour' of teaching itself and its attractiveness to students (Reis et al., 2021). The constituent elements of a smart learning environment are six parts, namely resources, tools, learning communities, teaching communities, learning methods, and teaching methods. It is shown in Figure 1.

Figure 1 The ROI theory of English teaching creativity in smart learning (see online version for colours)



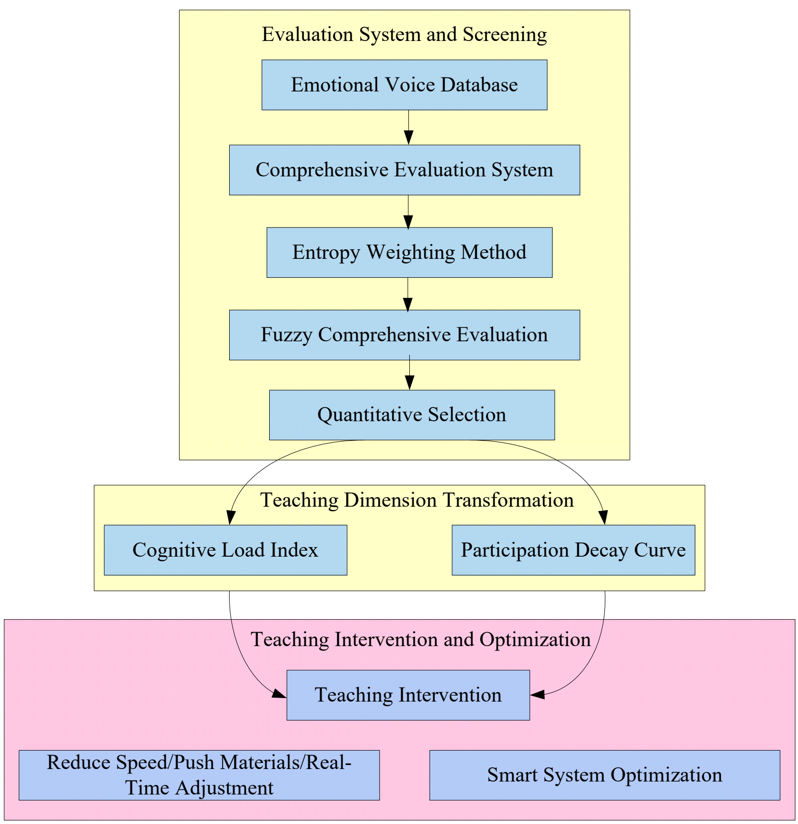
During the learning process, students are more likely to be infected by vivid and interesting content (Khan et al., 2021). To render students through video learning, in the content design and image composition of educational micro-videos, it is necessary to draw more creativity from the creative points of advertising and film and television (Troussas et al., 2021); it diversifies the presentation forms of teaching content (Li, 2020). The basic elements of teaching are teachers, learners, and teaching content (also known as teaching resources). This type of data collection is an information asset that requires new processing modes to have stronger decision-making power, insight discovery power, and process optimisation capabilities.

In the intelligent education scenario, the three dimensions of ROI theory need to be deeply integrated with emotion recognition: relevance requires accurate matching of emotion data with teaching objectives (such as identifying students' confused emotions to trigger knowledge point reinforcement); originality emphasises the design of differentiated teaching interactions through emotional feedback (such as expression-driven virtual teacher tone adjustment); and influence evaluates the effectiveness of teaching strategies through long-term emotional trajectory analysis (such as the correlation between the duration of positive emotions in class and test scores). This theoretical framework ensures that emotion recognition technology will not become an isolated data collection tool, but a core engine driving teaching optimisation.

3.2 Fuzzy comprehensive evaluation model for emotion recognition

Emotion recognition can be done in the form of speech databases (Sharma and Shekhar, 2020). A comprehensive and objective evaluation model can provide prerequisites and guarantees for a high-quality emotional speech database. In the initial stage, fuzzy comprehensive evaluation combined with weight and entropy method is used to identify and screen the speech database to ensure the validity of the database. First, a comprehensive indicator evaluation system is created, that is, a set of indicators is created for speech evaluation. Then, the entropy weight method is used to obtain the total weight of speech marks, and through fuzzy synthesis and fuzzy evaluation of the total weight matrix, a fully fuzzy speech evaluation matrix can be obtained (Tang, 2020). Finally, the comprehensive fuzzy evaluation matrix is quantified to obtain the overall score of the speech. According to the scores, the pros and cons of emotional speech can be judged, and logical and effective emotional speech can be selected. The emotion recognition flow chart in its intelligent learning is shown in Figure 2.

Figure 2 Emotion recognition flow chart in intelligent learning (see online version for colours)



Through Figure 2, the improved fuzzy comprehensive evaluation method is used to evaluate and control the emotional speech database from the two aspects of emotion accuracy and emotion recognition. Finally, the emotional speech database is installed. The controlled emotional speech database meets the research requirements of the above two indicators, making emotional expression more accurate and natural, closer to the teaching requirements of intelligent learning, and has no noise influence, which is conducive to further research on emotional speech.

The breakthrough of this model in the educational scenario is that it transforms traditional speech emotion classification (such as anger/happiness) into teaching-specific dimensions (such as the 'cognitive overload index' and 'engagement decay curve'). When the system detects high-frequency hesitation pauses in students' voices (corresponding to confused emotions), it can automatically trigger the teaching system to slow down the current knowledge point explanation speed and push visual auxiliary materials. This fuzzy judgement mechanism realises a seamless connection from 'emotion recognition' to 'teaching intervention', enabling the intelligent education system to have real-time adaptive adjustment capabilities.

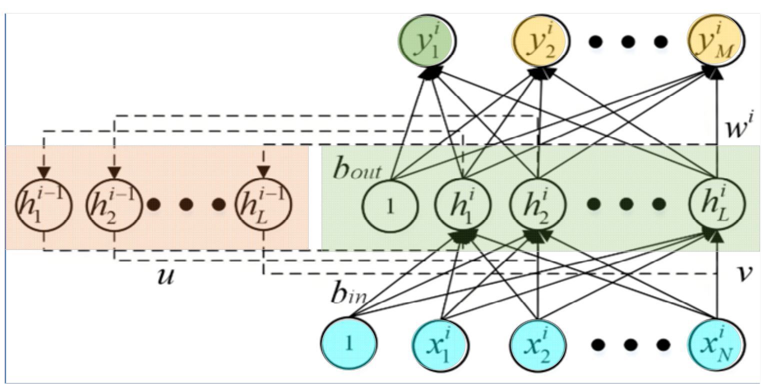
3.3 Echo state network

The selection of the SVM kernel function plays a crucial role in its performance, especially for those linearly inseparable data, so the selection of the SVM kernel function is very important. Echo state network (ESN) can effectively deal with nonlinear system identification and chaotic time series prediction problems. It provides structure and supervised learning specifications for recurrent neural networks. The state update formula of the classic sigmoid-based ESN is composed of N reserve pool units, K input layer units, and L output layer units:

$$x(n+1) = f\left(Wx(n) + W^{in}u(n+1)W^{fb}y(n)n\right) \tag{1}$$

Among them, $x(n)$ is the repository and f is a function. W is the internal matrix and $u(n)$ is the input signal. $y(n)$ is the output signal.

Figure 3 The basic structure of the Elman neural network (see online version for colours)



The output result can be expressed as:

$$y(n) = g\left(W^{out}z(n)\right) \tag{2}$$

Among them, g is the activation function and W^{out} is the output matrix. $Z(n)$ is a digital signal.

Together these actions activate the Elman neural network. Depending on the task situation, the above process may need to loop multiple times. Therefore, the error is obtained by comparing the output result with the expected value, and the connection weight is adjusted by the error back propagation rule. Figure 3 is the basic structure diagram of the Elman neural network.

f() is the activation function of the neurons in the hidden layer, and *h*() is the activation function of the successor layer. Net() represents the net input to a layer. A represents the input layer, *B* represents the successor layer, and t represents the iteration order. $v_i(t)$ represents the wave input and $\omega_i(t)$ represents the frequency. $x_n(t)$ represents the iterative function, and $y(t + 1)$ represents the output. The expressions between the layers are as follows:

The hidden layers are as formulas (3)–(6).

$$v_i(t) = \begin{cases} u_r(t), & i \in A \\ x_c(t), & i \in B \end{cases} \tag{3}$$

$$\omega^i(t) = \begin{cases} \omega^2, & i \in A \\ \omega^3, & i \in B \end{cases} \tag{4}$$

$$net_n(t+1) = \sum_{i \in A \cup B} \omega^i(t) v_i(t) \tag{5}$$

$$x_n(t+1) = f\left(net_n(t+1)\right) \tag{6}$$

The successor layers are as formulas (7)–(8).

$$net_n(t) = \sum_{i \in A \cup B} \omega^i(t-1) v_i(t-1) \tag{7}$$

$$x_c(t) = h\left(net_n(t)\right) \tag{8}$$

The output layer is as formulas (9)–(10).

$$net_m(t+1) = \sum \omega^1(t+1) x_n(t+1) \tag{9}$$

$$y(t+1) = g\left(net_m(t+1)\right) \tag{10}$$

At this point, the system energy changes. Define the acceptance probability of the system from $x(n)$ to $x(n + 1)$ as p:

$$p = \begin{cases} 1 \\ e^{\left(-\frac{E(n+1)-E(n)}{T}\right)} \end{cases} \tag{11}$$

An annealing rate calculation is performed, that is, the formula decreases:

$$T(n) = \lambda T(0) \tag{12}$$

The value is between 0.8 and 0.9. The other two descending methods are shown in formulas (13) and (14).

$$T(n) = T(0)/\log(1+t) \tag{13}$$

$$T(n) = T(0)/1 + t \tag{14}$$

According to the probability calculation, it can be obtained:

$$\min\{1, \exp(-\Delta f / T(k))\} > random[0, 1] \tag{15}$$

The algorithm then converges to the optimal solution.

Given English teaching scenarios, this study makes three aspects of educational adaptation to ESN:

1 time compression: control the delay of speech emotion analysis within 200 ms to meet the needs of real-time classroom interaction

2 noise filtering: develop an acoustic model exclusive to English classrooms to distinguish between student discussion sounds and environmental noise

3 cross-modal fusion: couple speech emotion data with behavioural data such as electronic textbook click heat maps and note input rates to build a multi-dimensional learning concentration prediction model.

This improvement makes ESN a key hub connecting ‘emotional perception’ and ‘teaching decision making’ in the intelligent education system.

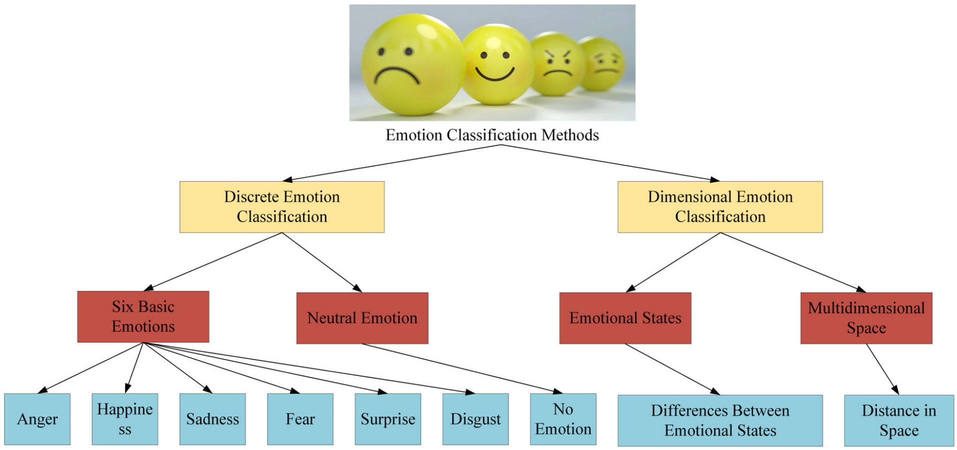
3.4 Emotion classification

Discrete emotion classification uses adjectives to describe emotions, such as anger, happiness, sadness, and so on. It is the most direct emotion classification method and the most widely used emotion classification in our daily lives. Currently, there are six basic emotions (also known as bugfixes) that are most widely used in the field of emotion research. Also, in real life, there is a neutral emotion, that is, no emotion at all. For familiar people, the real emotion of the other party can be identified based on the other party’s facial expressions, voice intonation, and obvious gestures or body language (Zafrullah et al., 2024). The memory function of emotion can be realised by recording the data, to predict the trend of emotion change of a specific object, and corresponding measures can be taken according to the prediction.

Dimensional emotion classification uses a multi-dimensional space to represent emotional states. This approach typically defines an emotional state as a point in a multidimensional space. In other words, any emotional state can find a corresponding point in space, and the difference between different emotional states can be represented by the distance in space (Guo et al., 2024). The structure of the emotion classification model is shown in Figure 4.

Constructivism has argued that teaching is not just about imparting learning content to students, but a process in which students actively acquire knowledge through interaction with the external environment (for example, autonomy, inquiry, etc.). Mastering new characteristics can provide reference and help to promote the construction of a smart learning environment. Particular attention should be paid to the design of the learning environment, which requires the learning environment to be as consistent as possible with the real world and to reflect its complexity. The learning environment should also provide different tools and information resources to help students learn and achieve established educational goals. The first type is a one-way interaction; the second type is a partial two-way interaction; the third type is a three-way interaction.

Figure 4 Emotion classification model (see online version for colours)



In intelligent education applications, traditional discrete emotion classification needs to be reconstructed into teaching effectiveness indicators: 'sadness' is converted into 'learning frustration index', 'anger' is converted into 'cognitive conflict intensity', and 'happiness' is converted into 'knowledge internalisation confirmation signal'. For example, when the system detects a frustration index peak of more than 3 times/minute, the stratified teaching mechanism is automatically activated: students with weak foundations are pushed animated micro-classes, and advanced students trigger peer mutual assistance matching (Tang, 2024). This classification system converts emotion recognition into an executable teaching instruction set.

4 Teaching effectiveness test

To evaluate the actual application effect of emotion recognition English teaching screen supported by artificial intelligence in smart learning, this paper takes fifth-grade students of a primary school as sample objects, randomly selects 60 people from parallel classes, 30 people in the experimental group (EG) and 30 people in the control group (CG). The experimental period is 8 weeks, with three English classes per week, each lasting 40 minutes. The course is based on the school's unified English textbook and is coordinated with different teaching methods to ensure an experimental environment consistent with the actual teaching scenario. A comparative test is conducted after each course.

In EG, the emotion recognition English teaching screen supported by artificial intelligence is used to assist in teaching. It is equipped with a multi-channel emotion recognition system that can monitor students' emotional status in real-time. The teaching screen will dynamically adjust the teaching strategy according to the student's emotional status. When it is found that the student's attention is reduced, it will automatically adjust the teaching speed or add interactive content to improve students' learning enthusiasm.

CG uses the traditional English teaching model, that is, teaching according to the pre-set teaching outline without the assistance of the teaching screen. The classroom interaction mode relies more on the teacher's experience and judgement and does not adopt an adjustment mechanism based on the student's emotional state.

During the experiment, EG and CG were taught by the same English teacher with more than 5 years of teaching experience, which ensured consistency in teaching content, teaching style, and teaching progress, and minimised interference from external factors.

Before the experiment was carried out, all students and parents who participated in the experiment had a comprehensive understanding of the experiment and strictly unified the experimental environment. To ensure the objectivity of the experiment, the basic information of the two groups of samples was compared, as shown in Table 1.

Table 1 Sample information

Indicators	*EG (n = 30)*	*CG (n = 30)*	*t*	*p*
Age	10.2 ± 0.5	10.3 ± 0.6	0.681	0.533
Gender (male: female)	16:14	15:15	—	—
English pre-test score (maximum score 100)	72.8 ± 5.3	73.2 ± 5.6	0.294	0.772
Vocabulary (maximum score 25)	18.2 ± 2.1	18.5 ± 2.0	0.566	0.589
Listening comprehension (maximum score 25)	17.8 ± 2.4	17.9 ± 2.6	0.164	0.878
Reading comprehension (maximum score 25)	19.3 ± 2.2	19.5 ± 2.3	0.343	0.742
Oral expression (maximum score 25)	17.5 ± 2.8	17.3 ± 2.9	0.274	0.797
Daily use of electronic devices (hours/day)	1.5 ± 0.6	1.4 ± 0.7	0.634	0.539

From Table 1, the gender ratio of the two groups of samples is roughly balanced. The students have similar foundations in English learning. The p value of the inter-group comparison of their English proficiency pre-test results is greater than 0.05, and there is no significant difference in English learning interest, classroom concentration, and electronic device use time.

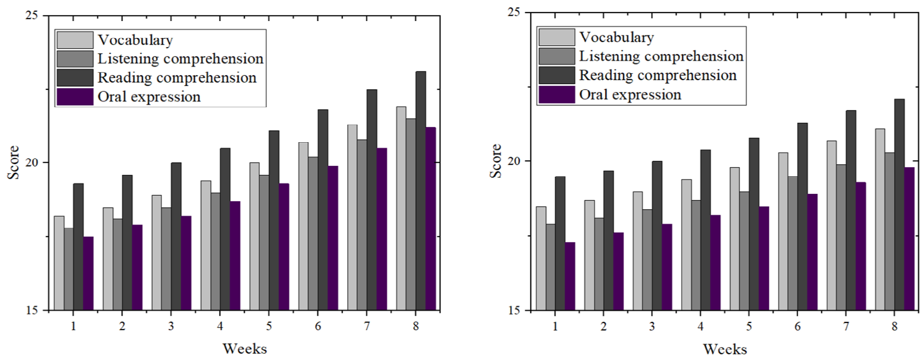
To comprehensively evaluate the effects of different teaching modes, specific evaluation indicators include:

1. English learning performance: based on the English proficiency post-test, the average scores of the two groups of students are statistically analysed
2. student learning interest: the students after the experiment were scored for interest tests
3. class participation: using the smart teaching screen (EG) and the teacher's (CG) classroom observation, the average concentration time of students is calculated, and the number of speeches in class is counted to compare the classroom participation of the two groups of students.

4.1 English learning performance

By comparing the English learning performance of the two groups before and after the experiment, the effects of the two different teaching modes on students' performance were analysed. The results are shown in Figure 5.

Figure 5 Comparison of English learning results, (a) the EG results (b) the CG results



(a) (b)

As can be seen in Figure 5, compared with CG students, EG, who use smart teaching screens for English teaching, has better English learning results. In Figure 5(a), the average test scores of EG in the four dimensions of vocabulary, vocabulary, listening comprehension, reading comprehension, and oral expression in the 8th week are 21.9 points, 21.5 points, 23.1 points, and 21.2 points; in Figure 5(b), the average test scores of CG in the four dimensions in the 8th week are 21.1 points, 20.3 points, 22.1 points, and 19.8 points. Compared with them, the average English learning scores of EG have increased by 3.8%, 5.9%, 4.5% and 7.1% respectively.

Based on ROI theory, this paper constructs an emotion-recognition English teaching screen supported by artificial intelligence, uses ESN to establish a student emotion recognition model, accurately identifies students' learning emotions, and can predict students' learning tendencies, and then dynamically adjust the teaching content, teaching rhythm, and interaction mode to improve students' understanding and acceptance. In sharp contrast, the CG group adopted the traditional teacher feedback method, which made students learn more passively, and their learning effect was limited.

4.2 Changes in learning interest

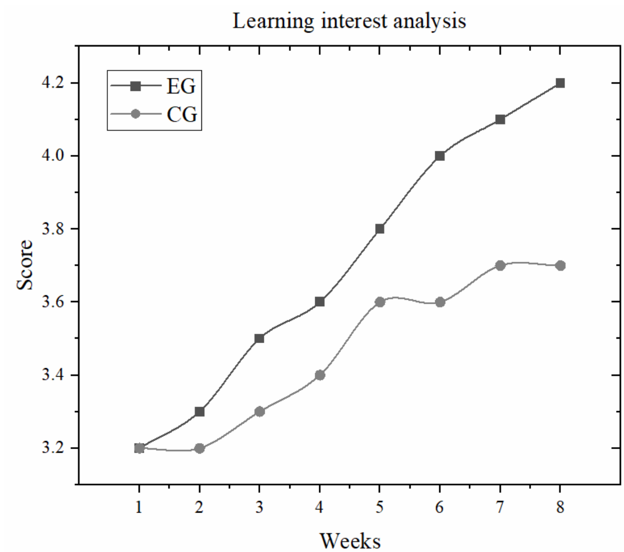
Learning interest is an important factor affecting students' learning outcomes and long-term academic achievements. The changes in the learning interest scores of the two groups of students during the 8-week experimental period are compared, and the results are shown in Figure 6.

In Figure 6, as the experimental period lengthens, the learning interest scores of students in each group continue to rise, but the EG score has a more significant increase. In the first week, the average learning interest score of both groups of students was 3.2 points; in the eighth week, the average learning interest score of EG students reached 4.2

points, while the average learning interest score of CG students was only 3.7 points; compared with CG, the average score of EG students increased by 13.5%.

In general, students using smart teaching screens show a higher interest in learning English, mainly because it can monitor students' learning status in real-time. When their attention is not focused, the fun of the class can be improved by adjusting the teaching speed or adding interactive content, thereby stimulating students' long-term learning motivation. The interactive mechanism of the CG group mainly relies on the teacher, so the students' learning interest improves relatively slowly.

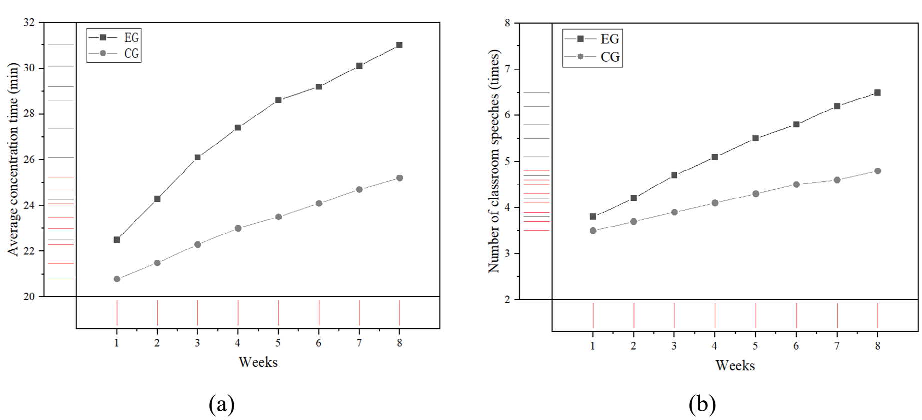
Figure 6 Comparison of learning interests



4.3 Classroom participation

The classroom participation of different groups of students is compared, as shown in Figure 7.

Figure 7 Class participation comparison, (a) the average concentration (b) the number of times students speak in class (see online version for colours)



From the comparison of classroom participation in Figure 7, the difference between the two groups in the eighth week is more significant. In Figure 7(a), the average concentration time of EG students in the eighth week reached 31 minutes, while that of CG students was only 25.2 minutes; in Figure 7(b), the average number of classroom speeches of EG students in the eighth week reached 6.5 times, while that of CG students was only 4.8 times. In the specific comparison, the average concentration time and the average number of classroom speeches of EG students increased by 23.0% and 42.5% respectively compared with CG students.

Through the adjustment mechanism combining emotion recognition with real-time feedback, teachers can better mobilise students' enthusiasm in class, encourage them to speak, and actively participate in class, thereby improving students' classroom participation. In sharp contrast, CG cannot make corresponding adjustments according to students' immediate emotional changes, resulting in students' inattention and low speaking frequency in class. From the overall results, the method in this paper can effectively improve students' classroom participation and optimise teaching effects.

5 Practical application of emotion recognition technology in English teaching

As an important branch of artificial intelligence, emotion recognition technology has gradually attracted attention in the field of education. By capturing students' emotional states in real-time, emotion recognition technology can provide more accurate and efficient teaching support for English teaching.

Emotion recognition technology captures students' emotional changes through voice, expression, and body movements, and provides real-time feedback to teachers. Voice analysis technology can capture students' intonation, speaking speed, and pauses when answering questions, determine whether students are confused or nervous, and adjust the teaching rhythm in time. When the system detects that students frequently hesitate or pause when answering questions, teachers can slow down the explanation or increase interactive links to help students better understand the knowledge points. Expression recognition technology captures students' facial expressions through cameras and identifies whether students are interested, confused, or tired, to help teachers adjust teaching content or increase interactive links in time. Body movement analysis technology determines whether students are focused or less engaged by analysing students' sitting posture, gestures, and other movements, and reminds teachers to take measures to improve students' classroom participation.

Emotion recognition technology can provide personalised learning support for each student. By analysing students' emotional data, the system can identify students' emotional fluctuations in learning and adjust teaching strategies accordingly. When the system detects that students frequently express confusion on a certain knowledge point, it can automatically push relevant auxiliary learning resources. In addition, emotion recognition technology can also divide students into different levels according to their emotional state and learning performance, and provide targeted teaching support for students at different levels. The peer mutual assistance mechanism can also be realised through emotion recognition technology. When the system detects that students have large emotional fluctuations, it automatically matches learning partners to help students relieve stress and improve learning effects through peer mutual assistance.

Emotion recognition technology can also help teachers optimise the design of teaching content. By analysing students' emotional reactions to different teaching links, teachers can understand which content can better stimulate students' interest and which content may cause students to feel confused or tired. According to students' emotional feedback, teachers can adjust the presentation of teaching materials to make them more vivid and interesting. In addition, through real-time emotion analysis, teachers can dynamically adjust the teaching rhythm to prevent students from feeling tired due to long-term concentration. The precise setting of teaching goals can also be achieved through emotion recognition technology. Teachers can ensure that the teaching content matches the students' emotional state and learning needs based on students' emotional data.

Emotion recognition technology provides more accurate data support for the evaluation of teaching effectiveness. Through the analysis of long-term emotional data, teachers can have a more comprehensive understanding of students' learning status and learning effects. Emotional trajectory analysis can evaluate the effectiveness of teaching strategies by analysing students' emotional changes in class. Dynamic monitoring of learning interest helps teachers understand which teaching methods can better stimulate students' learning interest by analysing students' emotional reactions to different teaching links. The basis for teaching improvement can also be achieved through the analysis of emotional data. Teachers can find out the deficiencies in teaching and optimise teaching strategies accordingly.

The application of emotion recognition technology in English teaching has achieved some practical results. Speech recognition-assisted oral practice can help students get real-time feedback on oral expression and improve pronunciation and intonation. Expression recognition to optimise classroom interaction has been applied in some smart classrooms. The camera captures students' expressions in real-time, and teachers adjust teaching content according to emotional feedback to improve the quality of classroom interaction. Personalised learning platforms have also introduced emotion recognition technology to provide students with personalised learning paths and resource recommendations.

6 Conclusions

The emotion recognition of artificial intelligence is a technology in which the computer extracts the emotional characteristic parameters of the collected signals and models the relationship between the emotional characteristic parameters and the emotion categories. It is very helpful for constructing the English teaching picture in smart learning. This paper has built a fuzzy comprehensive evaluation model for emotion recognition, evaluated and screened the intercepted emotional speech database from five aspects: emotion accuracy, background noise impact, clarity, naturalness, and picture sense, and finally established an emotional speech database. The basic structure of the Elman neural network and the schematic diagram of the structure of multiple intelligences have been drawn. This paper uses fifth-grade students from a primary school as sample objects and evaluates them from three dimensions: English learning performance, students' learning interests, and classroom participation. The experimental results show that compared with the traditional teaching model, the EG students' vocabulary, vocabulary, listening comprehension, reading comprehension, and oral expression test scores based on the teaching method of this paper have increased by 3.8%, 5.9%, 4.5%, and 7.1% respectively. The application of artificial intelligence-supported emotion recognition English teaching screen in smart learning can effectively improve the effect of English teaching, and enhance students' learning interest and classroom participation. However, this paper also has limitations. This paper fails to evaluate the continuous impact of long-term application on students' learning effects and lacks the verification of applicability for students of different school ages. Future research can explore the application effectiveness of smart teaching screens in long-term learning, and verify the adaptability of different school ages and different learning styles, to more comprehensively support the personalised development of smart teaching environments.

Declarations

The author has stated explicitly that there are no conflicts of interest in connection with this article.

The author was aware of the publication of the paper and agreed to its publication.

References

Al-Nakhal, M.A. and Naser, S.S.A. (2017) *Adaptive Intelligent Tutoring System for Learning Computer Theory*.

Berta, R. and Moreno-Ger, P. (2018) 'Introduction: intelligent learning assessment in serious games', *International Journal of Serious Games*, Vol. 5, No. 1, pp.3–4.

Coskun, K. (2019) 'Development of facial emotion recognition and empathy test (FERET) for primary school children', *Children Australia*, Vol. 44, No. 1, pp.23–31.

Guo, S., Zheng, Y. and Zhai, X. (2024) 'Artificial intelligence in education research during 2013–2023: a review based on bibliometric analysis', *Education and Information Technologies*, Vol. 29, No. 13, pp.16387–16409.

Khan, M.A., Zhang, Y.D., Sharif, M. and Akram, T. (2021) 'Pixels to classes: intelligent learning framework for multiclass skin lesion localization and classification', *Computers & Electrical Engineering*, Vol. 90, p.106956.

Li, W. (2020) 'Ontological approach to automatic generation of questions in intelligent learning systems', *Doklady BGUIR*, Vol. 18, No. 5, pp.44–52.

Lin, J., Wang, W. and Chai, J. (2021) 'Discussion and practice of blended teaching in basic theory and application of GSM-R', *Journal of Contemporary Educational Research*, Vol. 5, No. 11, pp.75–82.

Ma, X.L., Liang, J., Li, S. and Liu, J.J. (2018) 'The design and application of an intelligent learning support system based on knowledge structure', *US-China Education Review*, Vol. 8, No. 8, pp.313–331.

Mehta, D., Siddiqui, M.F.H. and Javaid, A.Y. (2018) 'Facial emotion recognition: a survey and real-world user experiences in mixed reality', *Sensors*, Vol. 18, No. 2, p.416.

Razek, M.A. and Frasson, C. (2017) 'Text-based intelligent learning emotion system', *J. Intell. Learn Syst. Appl.*, Vol. 9, No. 1, pp.17–20.

Reis, H.M., Alvares, D., Maillard, P.A.J. and Isotani, S. (2021) 'A proposal of a model of emotional regulation in intelligent learning environments', *Informatics in Education*, Vol. 20, No. 2, pp.1–16.

Sharma, A. and Shekhar, H. (2020) 'Intelligent learning-based opinion mining model for governmental decision making', *Procedia Computer Science*, Vol. 173, pp.216–224.

Tang, H. (2020) 'Research on teaching quality evaluation method of network course based on intelligent learning', *International Journal of Continuing Engineering Education and Life Long Learning*, Vol. 30, No. 4, pp.415–427.

Tang, K.H.D. (2024) 'Implications of artificial intelligence for teaching and learning', *Acta Pedagogia Asiana*, Vol. 3, No. 2, pp.65–79.

Troussas, C., Krouska, A. and Sgouropoulou, C. (2021) 'Improving learner-computer interaction through intelligent learning material delivery using instructional design modeling', *Entropy*, Vol. 23, No. 6, p.668.

Won, R. (2018) 'Intelligent learning with light', *Nature Photonics*, Vol. 12, No. 10, pp.571–573.

Xu, B. and Sun, F. (2017) 'Composite intelligent learning control of strict-feedback systems with disturbance', *IEEE Transactions on Cybernetics*, Vol. 48, No. 2, pp.730–741.

Zafrullah, Z., Meisya, A. and Ayuni, R.T. (2024) 'Artificial intelligence as a learning media in English education: bibliometric using Biblioshiny analysis (2009–2023)', *ELTR Journal*, Vol. 8, No. 1, pp.71–81.