Immersive English Online Teaching Model Using Original Film as Teaching Resources

https://doi.org/10.3991/ijet.v18i02.35467

Rongrong Song[1(✉)], Mohammad Otair[2]
[1]School of Foreign Languages, Xiamen Institute of Technology, Xiamen, China
[2]Faculty of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan
songrongrong@xit.edu.cn

Abstract—The traditional English teaching mode generally consists of reciting words, phrases and texts with high intensity, mechanically memorizing grammar and doing a lot of exercises. This way not only causes the majority of students lack of practical ability to use English, but also difficult to stimulate students' interest in learning. In this study, an immersive on-line teaching mode of audio-visual film English is constructed. Compared with the general on-line teaching mode, the model has a learning state feedback module based on convolutional neural network and support vector machine. Teachers can pay attention to students' learning state in time and adjust teaching strategies according to their learning state. Three classes of the same school and grade were selected for comparative experiment. Offline teaching, traditional online teaching and improved online teaching were adopted in one semester respectively. The effectiveness of the improved method was proved by analyzing the class status and final grades of students in the whole semester. Through network training and practical tests, compared with the students who adopt traditional online teaching, the students who adopt this mode have greatly improved their learning state and final exam scores. At a time when the epidemic is still lingering, this mode can provide a positive reference for the development of online teaching.

Keyword—immersive English online teaching, action recognition, status classification, learning effect

1 Introduction

The development of multimedia technology and big data technology has caused great changes in the teaching environment and teaching mode. Especially after the prevalence of novel coronavirus in 2019, online teaching has been rapidly developed and widely promoted [1]. On the whole, there are no equipment and technical problems in most regions [2]. Teachers' literacy is different, but teachers usually have good learning ability and can improve relevant skills through self-learning [3]. The main problem is that that online teaching is difficult to grasp students' learning state timely and accurately [4]. Although current online teaching systems have functions such as check-in, IP address detection, and irregular questions to determine whether students are learning, it is not enough for teaching to just confirm whether students are learning. High-quality teaching requires a high degree of interaction between teachers and students, teachers need to timely grasp the learning status of students, adjust teaching strategies according to the learning status of students, and timely warn students with poor learning status to maintain concentration (Grivet C S et al 2021) [5]. In order to achieve this goal, this paper proposes an immersive online teaching model of original film English based on learning state feedback. This model uses the English original film as a tool to guide students into the specific English context, so as to have a deeper understanding of the use of phrases and grammar. In addition, bilinear interpolation and gradient sharpening Convolution Neural Network (CNN) model were added to the online teaching system to recognize the students' actions, and Support Vector Machine (SVM) was used to construct the learning state classification model. So that teachers can timely understand the learning status of students.

2 Related work

Compared with the traditional teaching methods, the development of multimedia technology and modern network information technology has brought many conveniences to education, which has led to major changes in the teaching mode. Students have the right to choose a variety of learning methods. Dorfsman et al. developed a typology based on conceptual model that can be applied to different crises. This model can accurately identify whether teachers can accept online teaching, which is helpful for schools to carry out training and publicity of online teaching [6]. Tayeh uses an adaptive survey community model to assess the satisfaction of Undergraduates of Jordan km university with online courses. Data analysis shows that there is a strong positive correlation between teaching and social existence and course satisfaction, while students' average level, major, academic year and other factors have little statistical impact on course satisfaction [7]. Karagul and other scholars believe that the main difficulty of online teaching lies not only in the lack of technology or infrastructure, but also in the difference between teachers' teaching methods and students' personal characteristics. Through research, they found that students' digital literacy has a significant statistical relationship with gender and education background, while age is not a significant factor [8]. Scholars such as jallad st al think that the transition from traditional teaching to online teaching is a transition of organizational type, which can be promoted or prevented by personal, community and social factors. Students need to understand personal awareness, responsibilities, changes to be faced, the start and expected end of the transition, key turning points and other factors [9]. Kank m investigated the students' cognition and participation in online education in a university in Northern Cyprus during the epidemic period. The survey showed that students generally had a negative view on the experience of online education. However, the research results showed that students' academic performance in this semester of online education was higher than that in the previous semester. He believes that as long as the students' attitude and engagement are sufficient, online teaching can be as effective as offline teaching [10].

The fundamental purpose of English education is to enable students to skillfully use English in their daily life. For language learning, interest is the best teacher. The original sound film resources are good language teaching tools with the characteristics of authentic language, wide content, diverse themes and audio-visual integration. Ryan and other scholars have studied the impact of input quantity, input quality and output on students participating in bilingual immersion teaching. The research results show that students who have more input have higher scores on the receptive vocabulary scale, while students who have more output have higher scores on the expressive vocabulary scale. Bilingual courses should also consider the input and output of the target language [11]. Ju and other scholars have investigated the contribution of two different Pinyin skills and spoken words to Korean immersion teaching, and believe that spoken words and tone recognition can better explain the common differences in Korean reading this time than phonetic awareness and spelling at the beginning and end of Pinyin, that is, meaning related skills are better than code related skills in immersion teaching [12]. Jafri N and other scholars believe that cultural immersion programs can play a positive role in the process of second language learning, and it is an effective means to understand the cultural core of the language through appropriate film and television works [13]. Putri g and other scholars have made in-depth research on many cultural immersion projects held in Indonesia. The purpose of these projects is to promote the interaction between local society and overseas students. The research results show that cultural immersion projects can not only effectively promote overseas students' understanding of Indonesia, but also promote the development of local tourism [14]. Sanchez and other scholars assessed the needs and development of immersion education in two schools in bangpanya city. Although immersion teaching did not bring significant impact on academic achievement, its improved students' learning attitude and interpersonal relationship, and analyzed the technical and management problems of the program [15].

It can be seen from the existing research that many countries and regions have some research on the development of online education. From the initial recording and broadcasting mode to the mainstream online live broadcasting mode, there is a certain foundation in terms of technology, equipment and teacher literacy. Relatively speaking, immersive teaching has higher requirements on all aspects. In the research on immersive teaching, most scholars tend to go on field trips and participate in cultural interaction activities, and some scholars believe that it is feasible to select appropriate film and television works as teaching tools. In view of the reality does not allow, the study USES the appropriate film and television works as immersion teaching tool, and will be online teaching emphasis on feedback to students' study condition, in the traditional online teaching system, add to timely monitor and feedback on students' learning status module, in order to improve the teaching quality of existing online teaching mode.

3 Construction of an immersive on-line teaching model of audiovisual English based on students' learning state feedback

3.1 Immersive English online teaching mode based on original sound film

The fundamental purpose of English education is to enable students to effectively listen, speak, read and write English when necessary, rather than simply doing questions [16]. There are many difficulties in second language teaching. Restricted by cultural awareness gap and language environment, it is difficult for students to understand language knowledge [17]. As a cultural carrier, movies can show the cultural customs of the countries using the target language from both visual and auditory aspects, and there are enough practical communication cases. They are excellent teaching materials for second language learning [18]. The original sound film is used as the teaching resource, and the film selection needs to be cautious. See Table 1 for the basic requirements.

Table 1. Movie selection

Film dialogue	Easy to understand and close to life	Correct example: <The Pursuit of Happiness> Error example: <Tenet>
Film theme	Inspirational or funny. Avoid blood, pornography, politics and crime	Correct example: <Forrest Gump>, <Madagascar>. Error example: <Schindler' List>, <House of Cards>, <Prison Break>
Film content	Educational and interesting	Correct example: <Three idiots> Error example: <Les Choristers>
Other requirements	Moderate duration, correct pronunciation, normal speaking speed and simple plot	

It can be seen from table 1 that when selecting films, we should not only consider the teaching quality, but also consider whether the film ideas will have a negative impact on students and whether the values expressed by the films are in line with the universal values. Choosing movies is only the first step. Using original movies as teaching resources is a great test of teachers' information literacy and teaching literacy. In the Internet age, information literacy refers to the ability to discover information, understand information, use information to innovate knowledge and learn. As early as the beginning of the 21st century, some universities in China have used original films as English teaching resources. The traditional teaching mode of using English original films is shown in Figure 1.

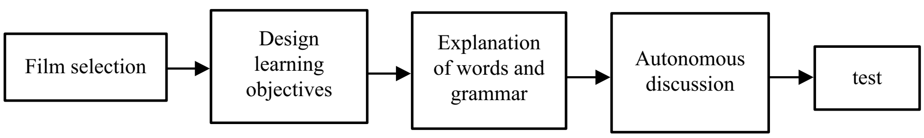

Fig. 1. Traditional English teaching using original sound film

As can be seen from Figure 1, this method is still a traditional exam-oriented teaching method in essence. It only changes the boring textbooks into English original movies. The teacher's thinking is still limited to the explanation of words and grammar, and does not really make use of the various functions of original movies. To some extent, this can stimulate students' interest in learning, but it cannot effectively improve students' practical English application ability. The immersive English teaching based on the original sound film proposed in this study focuses on the interaction and cooperation between teachers and students from the perspective of scene and culture, and fully mobilize students' subjective initiative. See Figure 2 for the specific teaching mode.

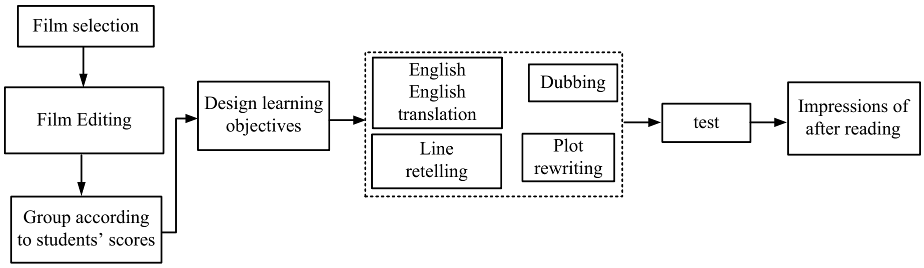

Fig. 2. Immersive audiovisual English teaching

As can be seen from Figure 2, the immersive education model no longer emphasizes the teacher's explanation of knowledge points, but emphasizes the participation of students. Students can use the English knowledge they have learned through various activities, consider the current language environment in their own position, and think about the connection between the language environment and language knowledge. This model does not emphasize the memory of vocabulary and grammar, but attaches more importance to the expansion of students' thinking and communication in oral and writing. In fact, grammar and vocabulary are no longer the most important points in the current college exams. Listening, reading comprehension and writing are the main difficulties. The immersive original film English teaching model can better improve students' relevant abilities. However, it is not enough to just change the online teaching strategy. The biggest difference between online teaching and traditional teaching is the separation of teacher and student space, which makes it difficult for teachers to grasp the learning status of students in time. Therefore, corresponding modules must be added to solve this problem.

3.2 A convolution neural network motion recognition model based on bilinear interpolation and gradient sharpening

As a classification algorithm, convolutional neural network has the advantages of high fault tolerance, adaptability to complex environments, and automatic feature extraction [19]. Its basic structure is divided into input layer, convolution layer, activation function layer, pooling layer, connection layer and output layer from the connection sequence [20]. Due to the different learning devices used by students, the pixels of the input image are different. However, the input dimension of the input layer is unchanged. In this study, bilinear interpolation method is used to zoom the image. The schematic diagram of bilinear interpolation is shown in Figure 3.

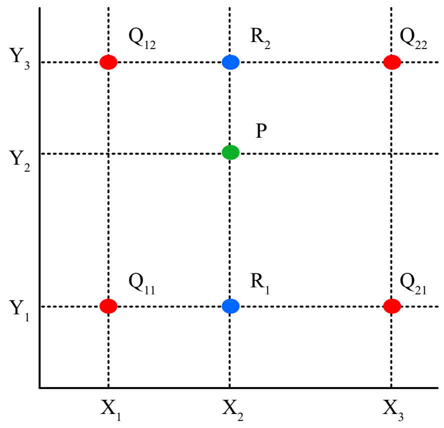

Fig. 3. Image scaling based on bilinear interpolation

As shown in Figure 3, first perform linear interpolation in the *x*-axis direction to obtain the sum of R_1 and R_2, and then perform linear interpolation in the *y*-axis direction to obtain *P*. see equations (1) and (2) for details. The interpolation direction does not affect the final result. It is also possible to interpolate in the *y*-axis direction first.

$$\begin{aligned} f(R_1) &= \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}) \\ f(R_2) &= \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}) \end{aligned} \tag{1}$$

In equation (1), *f* is an unknown function, *Q* is a known pixel point, and (x,y) is the coordinates of *Q* points.

$$f(P) = \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2) \tag{2}$$

In equation (2), *P* is a target pixel point. The interpolation order of bilinear interpolation method has no effect on the result, and it can also be interpolated on the axis first. Because the place where students study is not fixed, there may be complex background or even dynamic background. In this study, *Harr* feature face detection method is selected to recognize and eliminate background interference. After these two steps, the image may be blurred due to small pixels. In this study, the gradient differentiation method is used to sharpen the image. Let the gradient of image $f(x,y)$ at point (x,y) be a vector G. See formula (3) for details.

$$G[f(x,y)] = \left| f(i,j) - f(i+1,j) \right| + \left| f(i,j) - f(i,j+1) \right| \tag{3}$$

In equation (3), i is a constant increasing downward and j is a constant increasing right, simulating the computer display direction. The sharpened image will produce redundant noise and fringes, and the *Sobel* operator is selected to obtain the one step degree of the image. *Sobel* see equation (4) for the calculation method of operator, equation (5) for the calculation method of gradient size, and equation (6) for the calculation formula of gradient direction.

$$\begin{aligned} Sobel_x &= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * [1 \quad 0 \quad -1] = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \\ Sobel_y &= [1 \quad 2 \quad 1] * \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \end{aligned} \tag{4}$$

$$Sobel = \sqrt{Sobel_x^2 + Sobel_y^2} \tag{5}$$

$$\theta = \tan^{-1}\left(\frac{Sobel_y}{Sobel_x} \right) \tag{6}$$

When $\theta = 0$, it means that the image has a longitudinal edge at this point, and the left is darker than the right. The output of convolutional neural network is a three-dimensional feature vector. Let the first-dimension vector be the learning state, 0 means "bad", 1 means "good", and the state range is between [0, 1]. Next, we need to assign the learning state weight to the processed image, select multiple images that reflect the change of learning state to distinguish, "good" is 3 points, "normal" is 2 points, and "bad" is 1 point. See Table 2 for the action characteristics that reflect the learning state.

Table 2. Action characteristics reflecting learning state

Action Characteristics	Learning State
Duration of eye closure (More than 1s)	Good:(0, 10] s/5d Normal:(10, 30] s/5d Bad:(30, 300] s/5d
Mouth opening duration (More than 6s)	Good:(0, 10] s/5d Normal:(10, 30] s/5d Bad:(30, 300] s/5d
Duration of nod (More than 1s)	Good:(0, 10] s/5d Normal:(10, 30] s/5d Bad:(30, 300] s/5d
Line of sight deviation (More than 2s)	Good:(0, 10] s/5d Normal:(10, 30] s/5d Bad:(30, 300] s/5d
Length of body leaning forward or backward (More than 2s)	Good:(0, 10] s/5d Normal:(10, 30] s/5d Bad:(30, 300] s/5d

It is generally believed that the blinking time is 0.4s–0.6s, and the blinking frequency per minute is 15–20 times. Beyond this range, it can be preliminarily judged as poor mental health; If the mouth is opened for more than 6S, it can be preliminarily judged as yawning; However, if the students' heads are lowered and their eyes deviate for a long time, they will be distracted; The length of leaning forward or backward is used to judge whether students are sleeping or doing other things.

3.3 SVM based learning state classification model

After the pictures captured by the equipment camera are processed, the action features in the images are effectively recognized, and the data are converted into action feature vectors, which are input into support vector machine for classification. The workflow is shown in Figure 4.

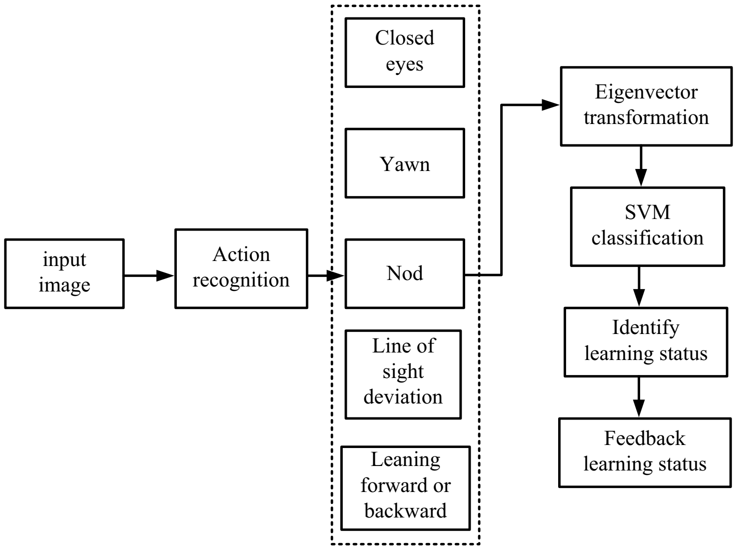

Fig. 4. State classification model based on CNN and SVM

In this study, the standard linear SVM is selected, and the activation function is linear activation function *Sigmoid*. See formula (7) for the specific formula.

$$f(x) = w \bullet x + b \tag{7}$$

In equation (7), w and b are the normal vector and intercept of the hyperplane respectively. Substitute equation (7) into function *Sigmoid* to obtain equation (8).

$$S\left(y = 1 \middle| f(x)\right) = \frac{1}{1 + \exp\left(Af(x) + B\right)} \tag{8}$$

In equation (8), exp is an exponential curve, and A and B are data sets. Use the maximum likelihood method to estimate A and B to obtain the target data set (a_n, b_n), where b_n is the value of the target probability. See formula (9) for details.

$$b_n = \frac{y_1 + 1}{2} (y_1 \in \{0, +1\}) \tag{9}$$

The final optimization form can be obtained according to equations (8) and (9). See equation (10) for details.

$$\min_{X=(A,B)} F(X) = \sum_{n=1}^{m} [b_n \log(p_n) + (1 - b_n) \log(1 - c_n)] \tag{10}$$

In equation (10), p_n is a posteriori probability, and the values of b_n and p_n are shown in equation (11).

$$b_n = \begin{Bmatrix} \dfrac{N_+ + 1}{N_+ + 2} \\ \dfrac{1}{N_- + 2} \end{Bmatrix}$$
$$p_n = \frac{1}{1 + \exp(Af(x_n) + B)} \tag{11}$$

In equation (11), 11 is the number of positive samples and 11 is the number of negative samples. Let any learning state be 11, and set an SVM classifier between each two learning states, with a probability of 11, and obtain equation (12).

$$P_{mn} = P(y = m | y = n, X) = \frac{1}{1 + \exp(Af + B)} \tag{12}$$

After calculating the posterior probability of which learning state X belongs to, the respective probabilities of the three learning states can be obtained. See Figure 5 for the online teaching mode constructed in this study.

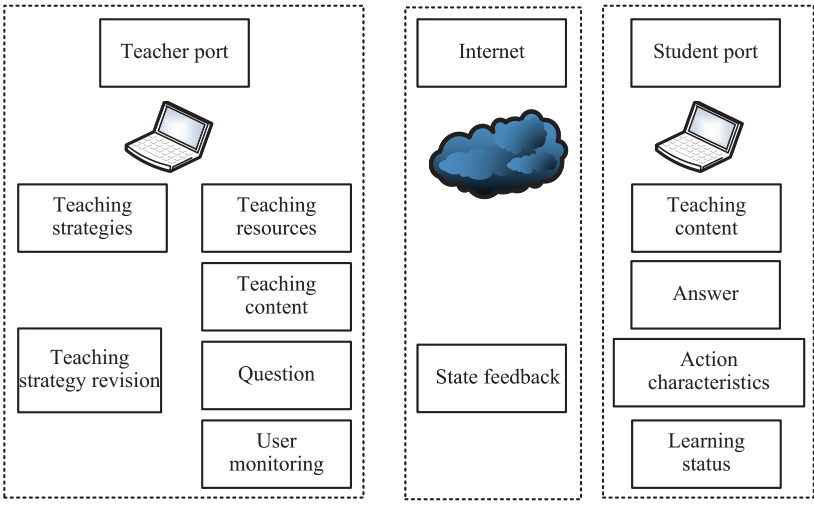

Fig. 5. Online teaching mode based on learning state feedback

In the current online learning environment, one to many class teaching modes is the most common teaching mode. Because teachers and students are in a state of space separation, it is difficult to detect the learning state of students. By adding motion capture and state classification modules to the teaching system, it can correctly reflect the learning state of most students, and teachers can adjust teaching strategies in time.

4 Comparative analysis of students' learning state based on time series

4.1 Training of state classification model based on CNN and SVM

For CNN networks, different parameter settings determine the performance in different problems. In this study, the most commonly used normal random distribution function is selected to initialize the parameters of CNN network, so as to enrich the diversity of the network and accelerate the convergence speed of the network. After debugging, the final parameters are set as follows: the number of convolution layers is 3, and the number of characteristic maps of each layer is 16, 32 and 64 respectively; The convolution kernel size is 3*3, and the pooled kernel size is 2*2; The step size of convolution layer is 1, and the step size of pooling layer is 2; The number of full connection layers is 2. Activation function select the activation function because it will cause gradient explosion or gradient disappearance. In order to facilitate the calculation and analysis of the experimental results, 100 volunteers were selected for the action feature recognition experiment. Including interference from complex background environment, different pixel devices, network fluctuations, insufficient light, etc. the accuracy of action feature recognition is shown in Table 3.

Table 3. Accuracy of action feature recognition

Action Characteristics	Accuracy
Closed eyes	83%
Yawn	92%
Nod	86%
Line of sight deviation	91%
Leaning forward or backward	94%

As shown in Table 3, under the interference of various factors, the CNN model based on bilinear interpolation and gradient sharpening has a high accuracy in identifying action features through student equipment cameras. The recognition accuracy of yawning, line of sight deviation and body tilt is higher than 90%, while the recognition accuracy of eyes closing and nodding is also higher than 80%, which can meet the demand. See Table 4 for the accuracy of linear SVM model in identifying students' learning state.

Table 4. Accuracy of learning state classification

Learning State	Accuracy	Average Accuracy
Good	83.4%	84.2%
Normal	81.6%	
Bad	87.7%	

It can be seen from table 4 that after the CNN model recognizes and processes the action data, the SVM model has a high accuracy in classifying the learning state. The recognition rate of "good" is 83.4%, that of "normal" is 81.6%, that of "bad" is 87.7%, and the average recognition rate is 84.2%. In this process, some "normal" states are identified as "bad", but in practical application, this error is acceptable, because the purpose is to enable teachers to timely adjust teaching strategies or warn students according to students' learning state. After being warned, the "normal" students who are wrongly identified as "bad" will also be more focused and achieve the expected results. In this study, WFLW data set was selected to train the model. The pictures in the data set include occlusion, posture, makeup, lighting, blur and other labels, and are marked with key points of face recognition, which can cover the situation faced by most students during learning. The training results are shown in Figure 6.

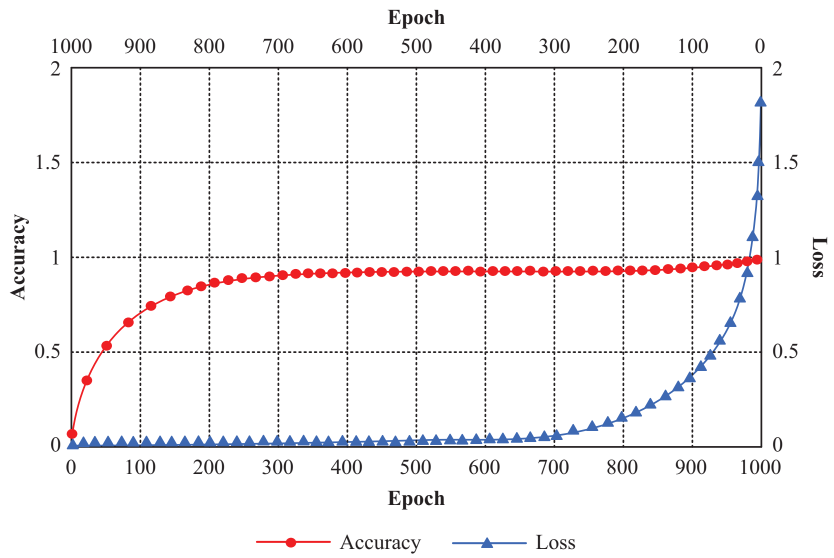

Fig. 6. Accuracy and loss

As shown in Figure 6, the accuracy rate of the learning state classification model combined by CNN and SVM tends to converge after about 100 iterations, and converges after about 300 iterations, with a fast convergence speed; After convergence, the accuracy is about 98%, which can meet the expectation. The loss value also tends to converge at about 80 iterations, and converges at about 300 iterations; The initial loss value is about 1.8, and the loss value after convergence is about 0.1, which can meet the expectation. No matter the accuracy or the loss value, there is no obvious fluctuation after convergence, and the model has good performance and stability.

4.2 Evaluation of learning effect of immersive soundtrack english online teaching mode based on learning state feedback

In this study, three classes of students of the same major in a university were selected as the test objects, which were marked as class 1, class 2 and class 3. The students in the three classes are all taught by the same English teacher. There is little difference in performance after excluding the particularly excellent and poor students. The three classes implement the traditional English online teaching mode, marked as mode a; Immersive audio-visual film English online teaching mode, marked as mode B; Based on the feedback of learning state, the immersive on-line teaching mode of audio-visual film English is marked as mode C. In all three modes, the camera of learning equipment will be used to capture students' action characteristics to analyze learning state, but only mode C will be fed back to teachers in time. The duration of each class in the school is 50 minutes. The system is set to feed back the learning state every 2 minutes. The whole semester is divided into initial stage, middle stage and final stage, and the average change of each period is obtained by integrating the learning state change of each class. See Figures 7, 8 and 9 for details.

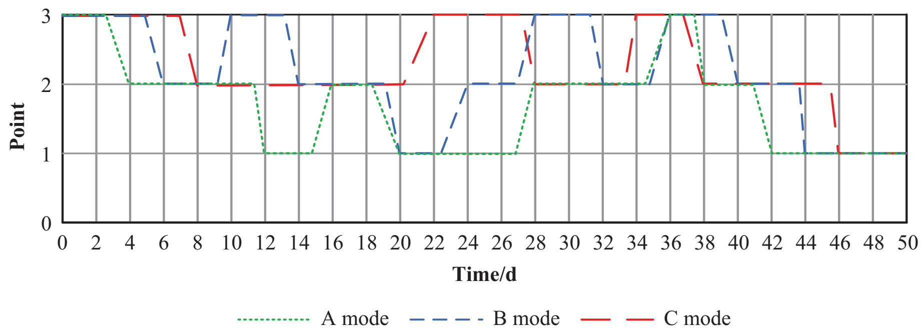

Fig. 7. Learning state at the beginning of the semester

As shown in Figure 7, at the beginning of the semester, students can ensure good learning state in 0–5 minutes, and after 5 minutes, their learning state drops to average; Using mode B and mode C can maintain this state for about 20 minutes, while using mode a will decline to the worse state again in about 10 minutes; In general, teachers will emphasize discipline in about 20 to 25 minutes, so the learning state of the three classes has rebounded during this period; After 40 minutes, students' learning state will decline again as the class is coming to an end.

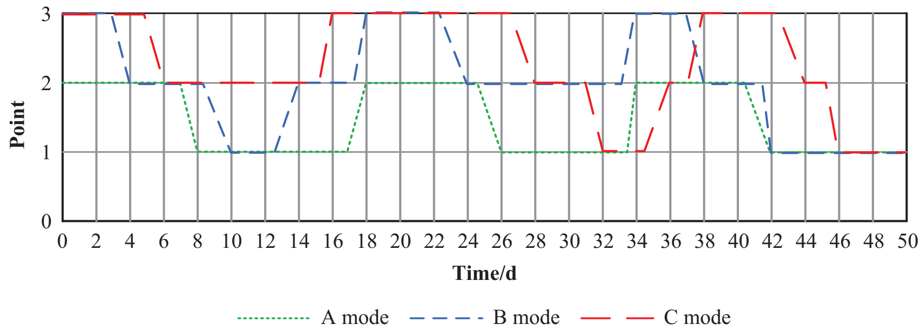

Fig. 8. Mid semester learning state

As shown in Figure 8, in the middle of the semester, the classes adopting mode B and mode C can still maintain a good learning state in 0–5 minutes. It can be considered that the immersive teaching of original sound and film can effectively maintain students' learning interest in the long semester; However, since mode B does not have a state feedback module, the teacher cannot find the students in poor state in time, and the state of mode B will decline again in about 10 minutes; The learning state of mode a fluctuates between average and poor in the middle of the semester, which is difficult to maintain a good state.

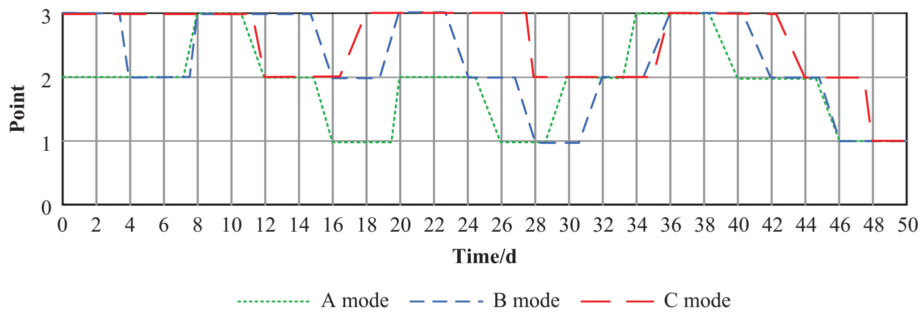

Fig. 9. Learning state at the end of the semester

As shown in Figure 9, at the end of the semester, students' learning state will pick up as they approach the exam. They can keep good or average in 0–10 minutes, and the period of decline will be postponed to 10 minutes later. As shown in Figures 7, 8 and 9, the learning state of mode a is the most unstable. It fluctuates between average and poor most of the time, and the fluctuation frequency is high. It fluctuates about every 4–6 minutes; However, mode B has a high attraction ability, students' learning state fluctuates less, and they can maintain a good learning state for a certain period of time, each time for about 5–8 minutes; C mode has high attractiveness and improves teachers' subjective initiative. It can remind students or change teaching strategies by observing changes in learning state in time, so as to improve students' learning state. In mode C, the holding time of each good state is about 10 minutes, and there are few bad states.

The duration of each bad state is 2–4 minutes, and the state fluctuation is less. The fluctuation of each class is 3–4 times, lower than 4–5 times in mode B and 5–7 times in mode C. Students' learning state is only a reference for learning effect. Generally speaking, the better the learning state, the better the learning effect. In most universities, the way to judge English level is still examination. The current mainstream examination papers mainly focus on students' listening, reading, translation and writing abilities. Because the results of the two semesters need to be compared in this study, the test paper difficulty of each semester is different, and the test state of students will fluctuate, so only the situation and average score of most students are considered, and extreme value and scattered distribution are not considered. See Figure 10 for the comparison of the scores of the three classes before and after the experiment.

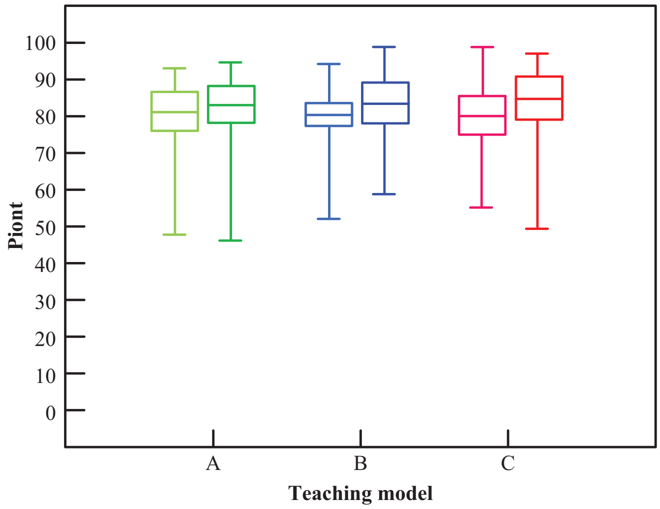

Fig. 10. Final exam

It can be seen from Figure 10 that after one semester of experiment, the scores of class 3 in mode C have improved significantly. Most students have increased from [75,86] to [79,91], with an increase rate of 5.33% and 5.81%, and the average score has increased from 80 points to 85 points, with an increase rate of 6.25%; Class 2 adopting mode B has increased from [77,84] to [78,89], the student rate is 1.29% and 3.71%, and the average score has increased from 81 points to 84 points, with an increase rate of 3.70%; Class 1 adopting mode a has changed from [76,88] to [78,87], and the average score has increased from 75 to 76, almost unchanged.

5 Conclusion

This study improves the existing online teaching system, adding a student status recognition module based on bilinear interpolation and gradient sharpening CNN algorithm and standard linear SVM. This module can timely identify the learning status of students and give feedback to teachers to help teachers improve the quality of teaching. In addition, immersive teaching is introduced into English online teaching, and English original video works are used as immersive teaching tools. This paper analyzes how to select appropriate film and television works in detail, and introduces how to use these works to construct the teaching framework. Through the comparative experiment, it is concluded that the students who adopt the new online teaching mode can maintain a better learning state in the whole semester, and the final average score is 2.25% higher than that of the reference group. The improvement effect is obvious. The research not only proves the effectiveness of the method, but also indicates that online teaching may become an important part of future teaching. The shortcoming of this study is that the selection of student samples does not consider the gender difference between men and women, nor does it compare the differences between different majors. The following research will focus on this direction.

6 Funding

The research is supported by: Educational research project for young and middle-aged teachers in Fujian Province (special research project on foreign language teaching reform in colleges and universities): The Smart Teaching of College English by the mode of "teaching goal oriented and student-centered" (NO: JSZW21036).

7 References

[1] K. Maha, E. Omar, E. Mohamed, "Design of Educational Scenarios of Activities in a Learning Situation for Online Teaching", vol. 1, pp. 49–64, 2021. https://doi.org/10.30574/gscaet.2021.1.1.0028

[2] N. Nartiningrum, A. Nugroho, "English Teachers' Perspectives on Challenges, Suggestions, and Materials of Online Teaching Amidst the Global Pandemic", *IJEE (Indonesian Journal of English Education)*, vol. 8, pp. 108–126, 2021.

[3] C. Z. Javid, N. S. Althobaiti, E. A. Al-Malki, "A Comparative Investigation of the Impact of Effective Online Teaching Strategies Practiced During Corona Pandemic in Ensuring Sustainable Pedagogy", *Universal Journal of Educational Research*, vol. 9, pp. 17–31, 2021. https://doi.org/10.13189/ujer.2021.090103

[4] G. Chinnasamy, "A Study on the Instant Switch to Online Teaching and Learning in Private Schools During the Covid-19 Pandemic", *The International journal of analytical and experimental modal analysis*, vol. 7, pp. 7–14, 2021.

[5] C. S. Grivet, M. Haj-Roussard, R. Broomé, "School administrators' perspectives of French immersion programs", *Foreign Language Annals*, vol. 54, pp. 1–25, 2021. https://doi.org/10.1111/flan.12512

[6] M. Dorfsman, G. Horenczyk, "The Coping of Academic Staff with an Extreme Situation: The Transition from Conventional Teaching to Online Teaching", *Education and Information Technologies*, vol. 27, pp. 267–289, 2021. https://doi.org/10.1007/s10639-021-10675-0

[7] Q. Tayeh, "The Role of Undergraduate Students' Satisfaction in the Transition to Online Education in Light of Teaching and Social Presence", *Multicultural Education*, vol. 7, pp. 104–113, 2021.

[8] B. I. Karagul, M. Seker, C. Aykut, "Investigating Students' Digital Literacy Levels During Online Education Due to COVID-19 Pandemic", *Sustainability*, vol. 13, pp. 1–11, 2021. https://doi.org/10.3390/su132111878
[9] S. T. Jallad, B. Isik, "Transitioning Nursing Students' Education from Traditional Classroom to Online Education During the COVID-19 Pandemic: A Case Study Applied to the Meleis Trial", *Florence Nightingale Journal of Nursing*, vol. 29, pp. 124–127, 2021. https://doi.org/10.5152/FNJN.2021.20090
[10] M. Kank, "Students' Perception of and Engagement in Reactive Online Education Provided During the Covid-19 Pandemic", *International Online Journal of Education and Teaching*, vol. 8, pp. 1063–1082, 2021.
[11] V. Ryan, "French Vocabulary Development of Early-Elementary Students in a Dual Language Immersion Program: The Role of Out-of-School Input and Output", *International Journal of Bilingualism*, vol. 25, pp. 1200–1213, 2021. https://doi.org/10.1177/13670069211000849
[12] Z. Ju, Y. Zhou, R. Delmas, "The Contributions of Separate Pinyin Skills and Oral Vocabulary to Chinese Word Reading of U.S", *Mandarin Immersion Third Graders*. *Reading and Writing*, vol. 34, pp. 1–21, 2021. https://doi.org/10.1007/s11145-021-10150-9
[13] N. Jafri, U. Manaf, M. M. Muhamad, "The Value of Cultural Immersions in Teaching Mandarin as a Second Language", *Universal Journal of Educational Research*, vol. 8, pp. 91–94, 2020. https://doi.org/10.13189/ujer.2020.081611
[14] G. Putri, A. Reswari, J. Kalimanzila, "Cultural Immersion Program as a Potential Live Laboratory for Applied English Students in Indonesia", *Journal of Vocational Education and Training*, vol. 2, pp. 11–15, 2020.
[15] R. Sanchez, P. J. Sarmiento, "Learning Together Hand-in-Hand: An Assessment of Students' Immersion Program in a School's Division", *International Journal of Research Studies in Education*, vol. 9, pp. 85–97, 2020. https://doi.org/10.5861/ijrse.2020.5809
[16] M. C. Franca, V. Boyer, M. I. Pegoraro-Krook, "Increasing Pluralistic Education through Cross-Cultural Interprogram Collaboration", *Perspectives of the ASHA Special Interest Groups*, vol. 5, pp. 1524–1529, 2020. https://doi.org/10.1044/2020_PERSP-20-00020
[17] C. P. Pagatpatan, M. Gurango, J. Madlansacay, "Medical Students' Perspectives on the Program Outcome Drivers of Community Immersion: A Realist-Informed Study", vol. 25, pp. 10–19, 2021.
[18] K. Sederberg, "Teaching Babylon Berlin: Language and Culture through a Hit TV Series", *Die Unterrichtspraxis/Teaching German*, vol. 54, pp. 200–216, 2021. https://doi.org/10.1111/tger.12171
[19] Y. Ding, Z. Ma, S. Wen, "AP-CNN: Weakly Supervised Attention Pyramid Convolutional Neural Network for Fine-Grained Visual Classification", *IEEE Transactions on Image Processing*, vol. 30, pp. 2826–2836, 2021. https://doi.org/10.1109/TIP.2021.3055617
[20] A. A. Aminu, N. N. Agwu, "General Purpose Image Tampering Detection Using Convolutional Neural Network and Local Optimal Oriented Pattern (LOOP)," vol. 12, pp. 13–32, 2021. https://doi.org/10.5121/sipij.2021.12202

8 Authors

Rongrong Song obtained her M.A. in British and American Literature (2010) from Shandong University, Jinan, China. Presently, she is working as a lecturer in the School of Foreign Language, Xiamen Institute of Technology, Xiamen, China. She has published articles in more than 5 national reputed peer reviewed journals and conferences proceedings. Her areas of interest include methodology of teaching English, British literature and American literature.

Mohammad Otair has a Ph.D. in Computer Information System/ Artificial Intelligence. His teaching experience includes information system courses over 21 years to University students. He has been working at Amman Arab University since 2010. He held different leading positions at the university: head for several departments (Computer Science, Computer Information System, and Software Engineering) and computer center manager in two universities, also he has an experience in management of projects. He is the author of 73 scientific publications including an organizer and a participant of more than 28 international scientific conferences and more than 45 international journals (most of them are about E-learning, Mobile Learning, and its Infrastructure). He supervised 32 master theses in Computer Science field. E-mail: otair@aau.edu.jo

Article submitted 2022-09-21. Resubmitted 2022-11-27. Final acceptance 2022-11-28. Final version published as submitted by the authors.