Open Learning Environment for Multimodal Learning Based on Knowledge Base Technology

https://doi.org/10.3991/ijet.v18i11.39397

Yanli Huang(✉)
Guilin University, Guilin, China
13978304405@163.com

Abstract—With the development of Internet technology, multimodal data have become the main data resource in the information age. Multimodal learning mode, as an education and teaching mode developed on multimodal data technology, provides more convenience for multimedia teaching. However, many challenges persist in its actual development and application. Currently, the multimodal learning model is susceptible to classroom noise, lack of teaching information, and other factors, which adversely affect multimodal data collection, teaching application, and achievement output. Thus, this paper optimizes the multimodal learning model in the open learning environment and takes 120 engineering students from a university in Guangxi Province as the research object. First, a sequential modal extraction method is proposed by constructing a multimodal probability generation model and then the data are modeled. Semi-supervised learning is then achieved by analyzing and combining the supervised and unsupervised learning processes. Finally, the knowledge base technology with information fusion characteristics is applied to the multimodal teaching mode. This teaching mode has been proven to improve students' learning ability and learning achievement and teachers' teaching effectiveness.

Keywords—knowledge base, multimodal learning, semi-supervised learning, open learning environment

1 Introduction

Multimodal learning refers mainly to digitizing the information contained in multimodal data into feature vectors, which is the basis of learning activities. The above concepts clearly show that multimodal learning can fully integrate and efficiently utilize knowledge [1]. At the same time, this education and teaching mode can establish the relationship between different information, obtain more information through the analysis of information, help open up learners' minds, inspire learners, and help them find effective learning methods that work for them, thereby helping students achieve continuous growth, which has an irreplaceable and important value for learners [2].

The subject knowledge of engineering majors is relatively complex, and the program focuses on basic knowledge and technical experience to cultivate the advanced technology talents in engineering. Therefore, focusing on methods and clear goals is important in the education and teaching of engineering majors. However, certain shortcomings and difficulties have been encountered in multimedia teaching in engineering teaching in colleges and universities, specifically in the following aspects. First, in the aspect of insufficient innovation and serious homogeneity [3]. The teaching materials are mainly in the form of teachers' homemade, network materials, school experts' materials, and purchased materials. Creating the homemade multimedia teaching materials is time-consuming and labor-intensive, and relatively few teachers choose to use homemade materials. Moreover, the homemade materials are mostly borrowed from the existing courseware materials, which do not pay enough attention to students' personalized needs and the characteristics of engineering majors. The multimedia teaching materials are also completed with serious homogenization. Therefore, stimulating the students' interest can be difficult. Second, updating teaching materials is slow, and keeping pace with the development of teaching demand is also a challenge. The studies of engineering majors involve considerable basic knowledge of professional theory, which can provide a necessary guarantee for students' learning and innovation. With the continuous development of the times, the data indicators of different majors in engineering are also constantly updated. For example, in recent years, China has updated its standards for steel members several times. However, the speed of updating related multimedia courseware is far from keeping pace with the updating of teaching national standards. In addition, students studying for a long time could easily cause derailment between students and enterprises, which is not conducive to students' development and innovation. Third, insufficient attention is given to the blackboard writing and teaching plan design. In multimedia teaching, many teachers add the teaching plan contents in multimedia courseware, believing that the content of the multimedia courseware is the instructional design, and thus, focus less attention on blackboard writing. If the multimedia equipment does not work properly, it will have many negative effects on the development of teaching activities.

In conclusion, multimodal technology, as an important form of the information age, can analyze and control different data and provide more convenience for people's lives and work. However, the teaching mode of multimodal technology is still in the stage of continuous exploration in university teaching. As mentioned earlier, inconsistencies can be found in the teaching materials, which do not match the constantly updated technology, resulting in a high workload for teachers to prepare lessons. Problems such as interference by outside parties and loss of student data during teaching also occur.

2 State of the art

Crescenzi et al. [4] (2020) pointed out that one of the greatest advantages of the multimodal learning mode is its characteristic of automatic data saving, which is a simple way to process lost data. That is, the modality has a natural processing method. Bayoudh et al. (2022) [5] proposed that in deep multimodal learning, one can focus on visual image modalities, such as images from a set of discrete signals from various image sensors. Our environment usually includes many forms in which we can see objects, hear tones, feel textures, smell scents, and so on. Tiong et al. [6] verified many possibilities for implementing a multimodal model through experimental analysis. Partial order will appear when an asymmetric and dissimilar model is executed. Cheng [7] proposed applying multivariate space-time models with alternate spatiotemporal interactions in multimodal models. In terms of teaching applications, Paul [8] conducted an in-depth study on the synergy of multimodality in second language teaching, and posited that language teaching will increasingly be assisted by multimedia technology. Gordani [9] studied the application principles and practical effects of multimodal PPT demonstration and pointed out that the use of multimodal PPTs to support teaching is more attractive to students' attention and allows them to have a more visual understanding of the information. To further explore the relationship between multimodal teaching and modern media technology, Cadamuro [10] studied from the perspective of teachers using interactive whiteboards in the classroom. He pointed out that that with the rapid development of modern media technology, classroom teaching is no longer dominated by language modality but is completed through the interaction and cooperation of various modalities. By exploring the expression of images and texts, Soleymani [11] promoted various modalities to assist teaching with the help of network information and multimedia technology. Their study achievements affirmed the contribution of the development of multimedia technology to multimodal teaching.

Relevant studies on China's multimodal teaching mode are still in the initial stage of development. It faces many difficulties and problems in practical study and application. Specifically, heterogeneity gaps in multimodal data, lack of accuracy in data modeling, difficulties in conducting regular evaluation, and difficulties in securing data security and ethics, still exist and require more in-depth research and exploration in the future. Chao et al. (2022) [12] proposed that based on ImageNet, HowNet, and CCD, a new multimodal knowledge base can be constructed by manual annotation, and the mapping of 21,455 noun and verb concepts in ImageNet was completed to map the concepts in HowNet and CCD into ImageNet effectively. This data set can be applied to natural language processing and computer vision tasks [13]. Liu et al. [14] (2021) proposed that multimodal learning analysis (MMLA) as a key technology for intelligently exploring the mechanisms by which effective learning occurs. Ma et al. [15] (2022) proposed that when adopting multimodal data to evaluate the occurrence and extent of deep learning, issues, such as heterogeneity gaps in multimodal data, lack of accuracy in data modeling, difficulties in conducting regular evaluation, and difficulties in securing data security and ethics still exist, and require more in-depth research and exploration in the future.

Hence, this study attempts to apply the multimodal learning mode in the teaching process of engineering disciplines while optimizing the previous multimodal learning mode because of the inadequate application of the multimodal learning mode in the open environment. The knowledge base corpus technology is innovatively integrated into lesson preparation. Then, by aiming at the problem that data can easily be missed, a semi-supervised classification and clustering framework is proposed. This new teaching method is applied to practical teaching to provide a new perspective for teaching reform in various disciplines in an open environment.

3 Methodology

The professional knowledge in engineering colleges and universities is relatively complex, making it more difficult to learn and acquire. The effect of traditional education and teaching mode is not ideal. Multimodal teaching mode needs to be practiced and applied continuously in the curriculum of engineering colleges and universities to gain more recognition and attention. Based on the research on the basic information of multimedia teaching and multimodal learning mode, this paper makes full use of the advantages of information collection and analysis of multimodal learning mode to provide students with more specific and fine learning methods and more references for engineering multimedia teaching mode.

3.1 Design of multimodal learning models

The multimodal learning mode seeks to gather different modal knowledge for learning. Through the establishment and analysis of the multimodal model, the problems of correlation between knowledge and information as well as those in learning can be solved. Multimodal learning mode can be expressed by the following equation:

$$f(x)=\sum_{k=1}^{M}\sum_{i=1}^{N} w_k a_{ik} k_k (x_{it}, x_k) + b.$$

The above equation indicates that each modality is classified as a marked sample. The variations are produced to carry out repetitive verification and training to improve the complementarity between modalities and the scientific and efficient application of multimodal learning modes in multimedia teaching.

The different modalities assumed to generated from the shared subspace, and the unedited prediction rates of different modalities are consistent, indicating the correlation between them. The correlations of different modalities are analyzed by typical correlations, and the optimization of the correlations of the two modalities is carried out continuously by decreasing the latitude of the two modalities to the d-latitude space with the following model:

$$p = tr\left(\frac{w_1^T X_1 X_2^T w_2}{\sqrt{(w_1^T X_1 X_1^T w_1)}\,(w_2^T X_2 X_2^T w_2)}\right).$$

In the above model, two modalities of different latitudes are represented as follows:

$$X_1 \notin R^{d_1 \times N}, X_2 \notin R^{d_2 \times N}.$$

Predictive consistency has a certain binding force for different modes when no labeled sample prediction consistency can be found. This method is usually applied in the semi-supervised learning process and can be described by the following models:

$$\left\| f_1(x_{i^1}) - f_2(x_{i^2}) \right\|_F^2 .$$

Multimodal semi-supervised learning methods must fully consider the prediction of different modal consistencies. Only by ensuring the consistency of different modalities can the effectiveness and scientificity of learning resources be improved. A large amount of unlabeled data can be represented as follows:

By fully considering the consistency of these unmarked data and potential predictions, the goal of collaborative regularization under such context can be expressed as follows:

$$\min_{f_1 \notin H_1, f_2 \notin H_2} \sum_{i=1}^{N_1} \ell(f_1(x_{i^1}), f_2(x_{i^2}), y_i) + r_1 \left\| f_1 \right\|_{H_1}^2 + r_2 \left\| f_2 \right\|_{H_2}^2 + \sum_{i=N_I+1}^{N_1+N_M} r_1 \left\| f_1(x_{i^1}) - f_2(x_{i^2}) \right\|_F^2 .$$

$\left\| f_1(x_{i^1}) - f_2(x_{i^2}) \right\|_F^2$ can also indicate that different modes can have some consistency in prediction while $r_1 \left\| f_1 \right\|_{H_1}^2$ and $r_2 \left\| f_2 \right\|_{H_2}^2$ are processed using the complexity of multiple models of reproducing kernel Hilbert space norms, which ultimately guarantees consistency of constrained countless data.

3.2 Construction of multimodal learning mode

In constructing a multimodal learning mode, it is necessary to ensure that the classifier accurately classifies and the similarity matrix is decomposed to ensure prediction consistency. Therefore, the overall framework of constructing a multimodal learning mode is defined as follows:

$$\min_{f_w, \gamma} \sum_{v=1}^{v} (\widehat{L}_v(f_v(X_v), \widehat{Y}) + \frac{\lambda_2}{2} \widehat{L}_v(X_v, \widehat{Y})).$$

In the above framework definition, constructing a multimodal learning mode requires establishing single modal classifier further. First, without losing generality, the loss function of the v and modality can be expressed by the following equation:

$$\min_{f_v} \ell(f_v(x_v), \widehat{Y}) + \frac{\lambda_1}{2} r(f_v).$$

The third step is to calculate the SLIM function using the linear function as follows:

$$F_v = X_v W_v + 1 b_V^T \Theta P_V.$$

The fourth step adopts the square loss method and redefines the loss function based on the equation of the third step, as follows:

$$\min_{w_v, b_v} \frac{1}{2\eta_v} \left\| F_V - \widehat{Y}\Theta P_V \right\|_F^2 + \frac{\lambda_1}{2} \left\| W_V \right\|_F^2.$$

After the single modal classifier is formulated through the above four steps, a clustering learner of unlabeled samples should be established. The topic information of unlabeled samples can alleviate the problem of modal loss. It is a modal classifier that can obtain a matrix to compensate for the similarity of modal loss, which can be defined as follows:

$$\tilde{L}_V = \left\| R_\Omega(M_V) - R_\Omega(\widehat{Y}\widehat{Y}_T) \right\|_F^2.$$

In an open environment, real data usually have certain interference factors, making clustering errors easy. Under semi-supervised monitoring, there is a need to reduce the influence of interference factors and increase robustness needs to be reduced. In such a context, the weights of modalities can be expressed as follows:

$$\frac{1}{\eta_V \left\| R_\Omega(M_V) - R_\Omega(\widehat{Y}\widehat{Y}^T) \right\|_F}$$

The second subparagraph is redefined as follows:

$$\min_{\widehat{Y}} \frac{1}{\eta_V} \left\| R_\Omega(M_V) - R_\Omega(\widehat{Y}\widehat{Y}^T) \right\|_F$$
$$s.t. \widehat{Y}^{\Theta 1} = Y, 0 \le \widehat{Y} \le 1.$$

By combining the above calculation equations and ideas, it is finally concluded that SLIM can be expressed as follows:

$$\min_{W_V, b_v, \widehat{Y}} \sum_{V=1}^{V} \frac{1}{2\eta_v} \left\| F_V - \widehat{Y}\Theta P_V \right\|_F^2 + \frac{\lambda_1}{2} \left\| F_V \right\|_F^2 + \frac{\lambda_2}{\eta_V^2} \left\| R_\Omega(M_V) - R_\Omega(\widehat{Y}\widehat{Y}^T) \right\|_F$$
$$s.t. 0 \le \widehat{Y} \le 1, \widehat{Y}^{\Theta 1} = Y$$

3.3 Application of multimodal learning mode in multimedia teaching

Multimodal learning mode can provide more ideas and resources for multimedia teaching, which can effectively assist the multimedia teaching of engineering majors in colleges and universities. Determining how the integration of multimodal learning mode and multimedia teaching and giving full play to the value of multimedia teaching is a topic worthy of in-depth study. The application of multimodal learning mode in multimedia teaching is shown in Figure 1:

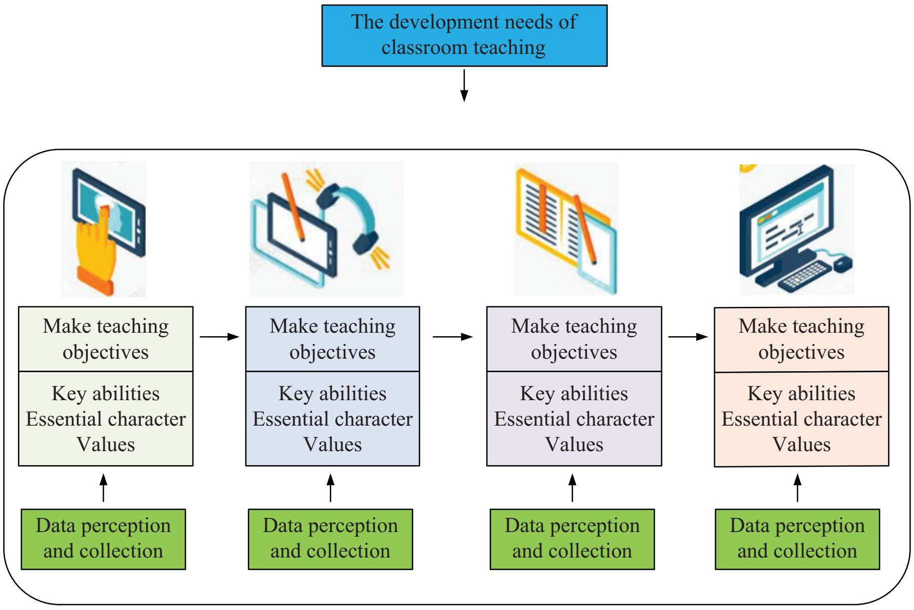

Fig. 1. Teaching process to improve the quality of multimedia classrooms based on multimodality

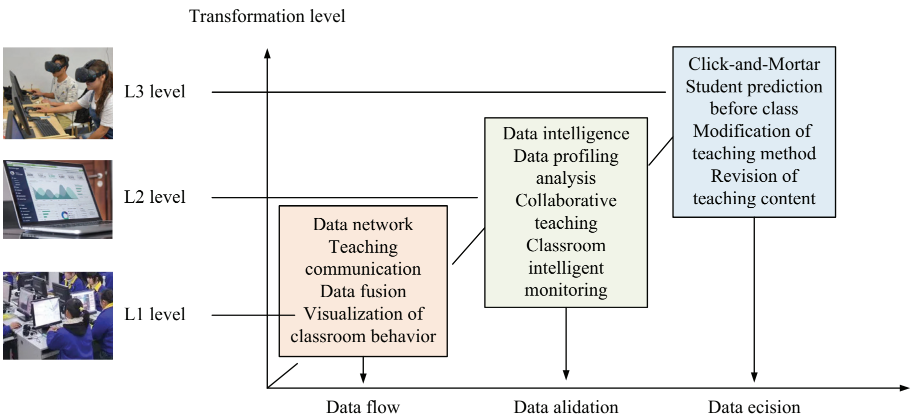

Fig. 2. Multimodal learning model transformation hierarchy chart

As shown in Figure 2, a multimodal knowledge base is built in this paper, with the main purpose of mapping the concepts in ImageNet to HowNet and CCD. Because HowNet and CCD contain English concept descriptions, they are mapped separately in English. First, the multimodal data program is used to screen and determine the multimedia teaching resources in establishing the above multimodal knowledge base. Second, the screened resource information is refined further through manual screening. Finally, the data obtained after the screening is applied to multimedia teaching activities. Through the above process, the multimodal information database is found to be capable of identifying and classifying the information learned by engineering students and finally storing it in the knowledge base. In multimedia teaching, it is possible to directly select information from the multimodal database to integrate multimodal learning mode and multimedia teaching. The multimodal knowledge base construction process is shown in Figure 1, and the multimodal learning mode hierarchy is shown in Figure 2.

4 Results and analysis

4.1 Teaching example

In this paper, the application of multimodal learning mode based on knowledge base technology was studied in depth in the application effect in the College English course of engineering majors with the following teaching flowchart:

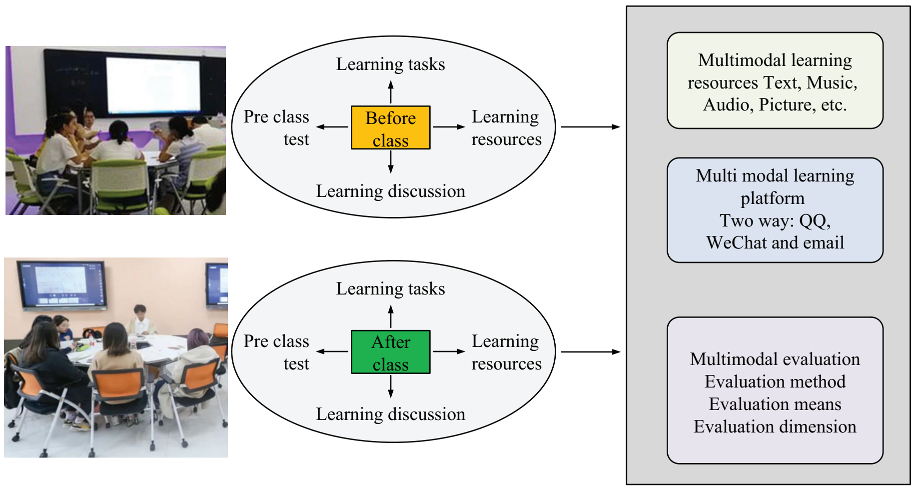

Fig. 3. Data-based flowchart of multimedia teaching in classrooms of engineering majors in colleagues and universities

As shown in Figure 3, the educational model constructed in this paper integrates the multimodal learning mode with multimedia teaching technology through an education and teaching mode that combines online and offline learning. A new teaching mode framework is built by combining the language knowledge base technology to meet the characteristics and needs of engineering English courses. The above teaching model reveals that the multimodal learning mode is integrated into all aspects of resource collection, learning interaction, and teaching evaluation, thereby promoting efficient multimedia teaching activities. Specifically, there are three main aspects:

First, as shown in Figures 4 and 5, in collecting multimodal learning resources, the resources mainly come from authoritative knowledge bases, such as corpora. The correlation between different teaching resources is established through an in-depth analysis of related teaching resources. Finally, the most suitable material for education and teaching is selected to form a multimodal teaching resource database.

Fig. 4. Information presentation of the multimodal knowledge base (I)

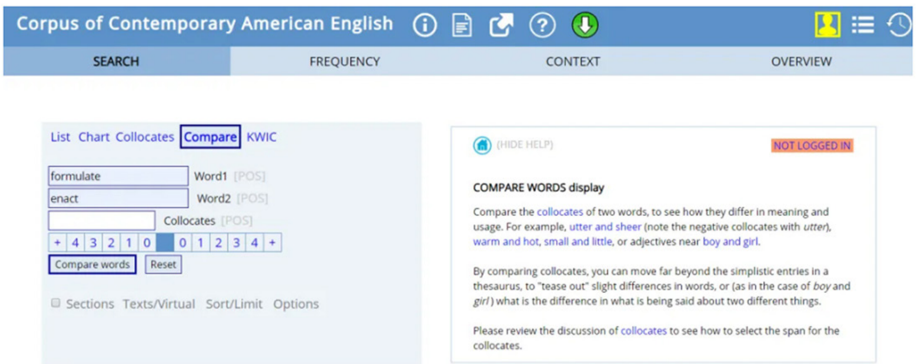

Fig. 5. Information presentation of the multimodal knowledge base (II)

Second, in the pre-course learning session, by screening the teaching resources and the design of multimedia teaching courseware, the multimedia teaching activities are carried out in a targeted manner by combining the learning characteristics and needs of students, forming a virtuous learning mode of active learning, active interaction, in-depth investigation of college students (as shown in Figure 6). In the process of education and teaching, modern information means can be fully utilized to establish a teaching mode that combines online and offline teaching, thereby helping students to choose an online learning mode suitable for them according to their habits and needs. In addition, multimedia information input and output channels such as audio, video, pictures, PPT, and text best suit students' individual and multi-dimensional needs.

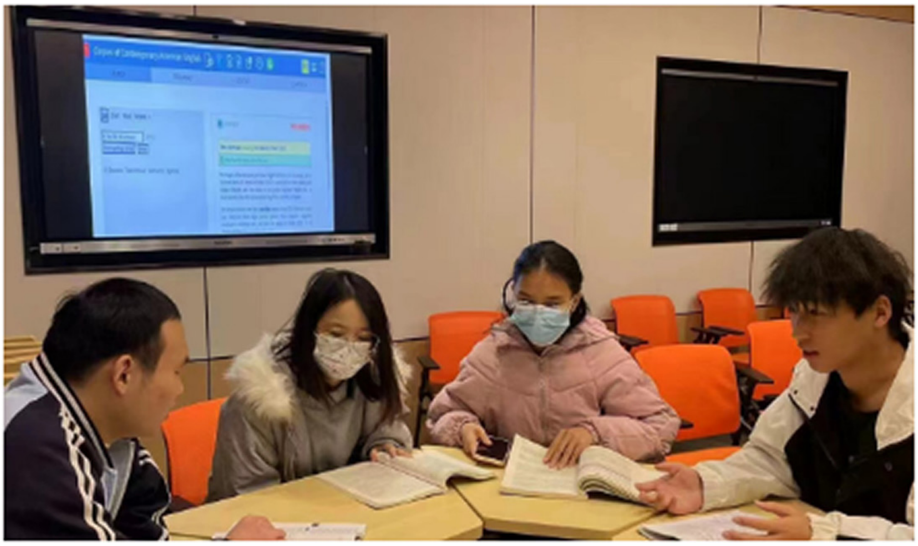

Fig. 6. Multimodal learning model for pre-course learning sessions

Third, teachers collect and analyze the contents of education and teaching according to the requirements of multimedia teaching in classroom teaching and apply these to the contents to promote the application of teaching practice. They also select targeted evaluation methods and means and dimensions based on a multimodal learning mode to achieve mutual growth of teachers and students.

At the same time, we have to consider the theoretical framework of multimodal instructional design, which contains three important factors: the combination of modality types, the relationship between modalities, and the principles of selecting modalities. This instructional design divides modalities into language: auditory, visual, body, and spatial modalities from the perspective of feeling. (see Figure 7).

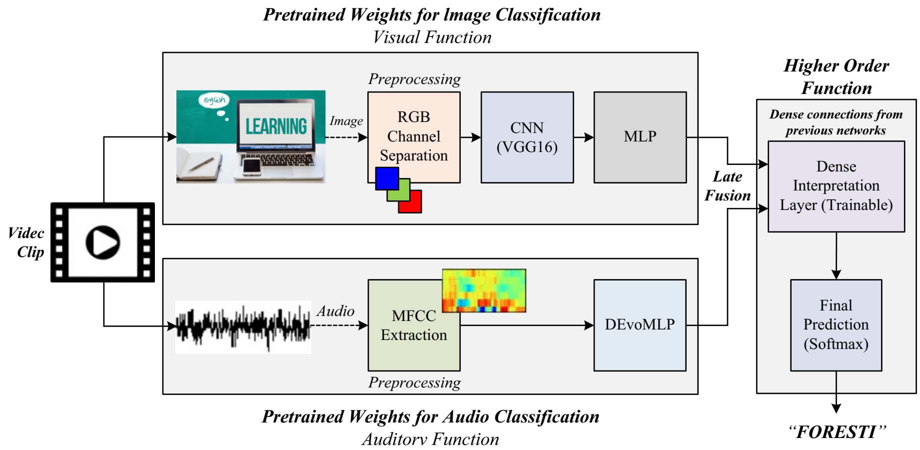

Fig. 7. Modal analysis presentation of multimodal learning mode

4.2 Teaching effect

From March to June 2022, 120 students, including 73 males and 47 females, were selected from an undergraduate class of 2022 engineering majors at a university in Guangxi Province. All students were divided equally into the experimental and control groups according to their school numbers, and the differences in general information between the two groups were not statistically significant ($P > 0.05$). Selection criteria: Test scores of basic theoretical knowledge are above "good" at the time of selection; students can correct their learning attitude and actively participate in the course.

Table 1. Comparison of learning achievements between two groups in multimodal learning mode

Group	Definition (10 Points)	Choice Questions (35 Points)	Short Answer Questions: (25 Points)	Image Analysis Questions (30 Points)	Total Score (100 Points)
Experimental group (n=60)	8.58±1.97	27.98±6.23	22.67±3.51	26.78±5.88	85.16±11.52
Control group (n=60)	7.31±2.16	23.54±5.98	21.09±5.08	21.09±5.27	71.09±15.03
t value	–0.959	2.925	3.409	3.551	4.058
p value	0.341	0.005	0.001	0.001	<0.001

Table 2. Comparison of learning abilities between two groups in multimodal learning mode

Group	Learning Attitude	Knowledge Acquisition Skills	Analytical Thinking Skills	Teamwork Skills	Communication and Expression Skills
Experimental group (n=60)	7.60±0.51	7.80±0.37	7.95±0.36	7.98±0.35	8.10 (6.75, 10.00)
Control group (n=60)	6.20±0.36	6.30±0.38	6.36±0.39	6.48±0.33	8.10 (5.00, 8.00)
t value	–2.322	–2.896	–2.896	–3.122	–3.018
p value	0.023	0.005	0.005	0.003	0.003

The statistical results in Tables 1 and 2 show that applying a multimodal learning mode based on knowledge base technology positively affects students' learning achievement and learning ability. At the same time, the teaching ability of teachers can be improved. Specifically, the multimodal learning mode of multimedia teaching can analyze many well-known engineering or design cases so that the instructional design process can fully integrate students' characteristics and needs and select teaching resources. Through the display of teaching cases, students' motivation and interest along with their sense of honor and interest will be stimulated through excellent local cases and designs. In turn, it can enhance the effect of education and teaching. Second, the teaching mode helps teachers make intelligent choices in terms of advanced scientific research results when preparing lessons, building a database of professional results, helping students have a full and comprehensive understanding of professional knowledge and results, and reducing teachers' workload, improving their effect and raising their teaching level. At the same time, this innovative practical learning is more targeted in learning content and more active in classroom atmosphere. Third, it can help teachers to use their own advantages to teach. The multimodal learning mode can make teaching more diverse, provide teachers with more opportunities to show their personal teaching ability, and guide students to participate more in teaching activities. Students can share and reflect with teachers, stimulate interest in the process, cultivate professional interest, and help them master learning methods suitable for them. Multimedia teaching is more dependent on multimedia resources, and it can give full play to the advantages of multimedia resources. However, teachers' ability to analyze teaching resources, including how to choose education and teaching resources with relevance, scientificity, and high value is limited The multimodal learning mode in this study can help teachers to analyze teaching resources and choose more valuable multimedia teaching resources. In particular, multimedia teaching of engineering, the multimodal teaching mode can extract the HOG features of original pictures, MFCC features in voice, sensor signals and other picture, text, voice and device information for multimedia teaching, which can be found out through the multimodality to finally form a multimodal teaching library model, thus improving the comprehensiveness of teaching content and helping to improve students' learning ability and learning achievement.

5 Conclusions

In view of the characteristics of engineering majors, this paper proposes a new multimodal learning mode in an open learning environment that optimizes the traditional multimodal learning mode by applying knowledge base technology and semi-supervised learning technology, and carries out a series of exploratory reforms and practices. The new teaching mode was found to help improve students' learning ability and learning achievement, reduce the workload of teachers, and improve the efficiency of teaching work. Such teaching mode has contributed greatly to the advancement of teaching effectiveness in engineering majors, which is analyzed as follows:

(1) The information library of the multimodal learning mode can identify and classify the information learned by engineering students, and automatically save the data. Such kind of intelligent selection of multimodal database can realize the integration of multimodal learning mode and multimedia teaching.
(2) Semi-supervised learning makes full use of unlabeled samples in student information, allowing multimedia learners to automatically use unlabeled samples to improve learning performance without relying on external interaction. Thus, it could effectively help teachers extract teaching information and make classroom content more relevant.
(3) The integration of knowledge base technologies in the multimodal learning mode is more conducive to improving teaching effect. Knowledge base technology can greatly enhance the integration of teaching resources and provide teachers with more comprehensive and advanced lesson preparation resources. Enhancing the effectiveness of the application of multimodal learning mode can better achieve the sorting of complex knowledge and advance the development of multimedia teaching.

6 References

[1] Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. From signals to knowledge: A conceptual model for multimodal learning analytics. Journal of Computer Assisted Learning, 2018, vol. 34(4), pp. 338–349. https://doi.org/10.1111/jcal.12288
[2] Cukurova, M., Giannakos, M., Martinez-Maldonado, R. The promise and challenges of multimodal learning analytics. British Journal of Educational Technology, 2020, vol. 51(5), pp. 1441–1449. https://doi.org/10.1111/bjet.13015
[3] Du, E.D, Zheng, L., Peng, M.G., et al. Countermeasures and prospect of online teaching of engineering experiment courses. Jiangsu Science & Technology Information, 2020, vol. 37(31), pp. 59–62.
[4] Crescenzi-Lanna, L. Multimodal learning analytics research with young children: A systematic review. British Journal of Educational Technology, 2020, vol. 51(5), pp. 1485–1504. https://doi.org/10.1111/bjet.12959
[5] Bayoudh, K., Knani, R., Hamdaoui, F., et al. A survey on deep multimodal learning for computer vision: Advances, trends, applications, and datasets. The Visual Computer, 2022, vol. 38(8), pp. 2939–2970. https://doi.org/10.1007/s00371-021-02166-7

[6] Tiong, L.C.O., Kim, S.T., Ro, Y.M. Implementation of multimodal biometric recognition via multi-feature deep learning networks and feature fusion. Multimedia Tools and Applications, 2019, vol. 78(16), pp. 22743–22772. https://doi.org/10.1007/s11042-019-7618-0
[7] Cheng, W., Gill, G.S., Ensch, J.L., et al. Multimodal crash frequency modeling: Multivariate space-time models with alternate spatiotemporal interactions. Accident Analysis & Prevention, 2018, vol. 113, pp. 159–170. https://doi.org/10.1016/j.aap.2018.01.034
[8] Paul, A., Miguel, F. Reading with eyes wide open: Reflections on the impact of multimodal texts on second language reading. Íkala, revista de lenguaje y cultura, 2017, vol. 22(1), pp. 57–69. https://doi.org/10.17533/udea.ikala.v22n01a04
[9] Gordani, Y., Khajavi, Y. The impacts of multi-modal PowerPoint presentation on the EFL students' content knowledge attainment and retention over time. Education and Information Technologies, 2020, vol. 25(1), pp. 403–417. https://doi.org/10.1007/s10639-019-09979-z
[10] Cadamuro, A., Bisagno, E., Di Bernardo, G.A., et al. Making the school smart: The interactive whiteboard against disparities in children stemming from low metacognitive skills. Journal of e-Learning and Knowledge Society, 2020, vol. 16(1), pp. 33–43.
[11] Soleymani, M., Garcia, D., Jou, B., et al. A survey of multimodal sentiment analysis. Image and Vision Computing, 2017, vol. 65, pp. 3–14. https://doi.org/10.1016/j.imavis.2017.08.003
[12] Chao, R., Zhang, K.L., Wang, J.J., et al. Construction of Chinese multimodal knowledge base. Journal of Guangxi Normal University (Natural Science Edition), 2022, vol. 40(3), pp. 31–39
[13] Chen, Y., Zhou, G., Lu, J.C. Survey on construction and application research for multi-modal knowledge graphs. Application Research of Computers, 2021, vol. 38(12), pp. 3535–3543.
[14] Liu, Q.T., Li, X.J., Xie, K., et al. Developments and prospects of empirical research on multimodal learning analysis. E-education Research, 2022, vol. 43(1), pp. 71–78, 85.
[15] Ma, Y.F., Zheng, X.D., Zhao, R., et al. Research on the concurrence mechanism and multi-modal evaluation of deep learning. Journal of Distance Education, 2022, vol. 40(1), pp. 50–60.

7 Author

Yanli Huang, Master, she is a lecturer in Guilin University, Guilin, China. Her research interests focus on Multimodal Learning Mode and English Teaching (13978304405@163.com).

Article submitted 2023-03-06. Resubmitted 2023-05-04. Final acceptance 2023-05-05. Final version published as submitted by the authors.