

Research Article

Convolutional Neural Network-Assisted Strategies for Improving Teaching Quality of College English Flipped Class

Tiankun Liu 

School of Foreign Languages, Hulunbuir University, Inner Mongolia Hulunbuir, 021008, China

Correspondence should be addressed to Tiankun Liu; liutiankun@hlbec.edu.cn

Received 13 July 2021; Revised 3 August 2021; Accepted 6 August 2021; Published 15 August 2021

Academic Editor: Yuanpeng Zhang

Copyright © 2021 Tiankun Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The “flipped classroom” teaching paradigm not only follows the cognitive rules of the learners, but it also subverts and reverses the standard classroom teaching process. Problem-oriented, teacher-led, student-centered, and mixed teaching approaches are the key teaching methods in the flipped classroom teaching model, which focuses on students’ procedural knowledge acquisition and critical thinking training. There are a lot of studies on the specific practice path of the “flipped classroom” teaching style right now, but there are not many on the learning involvement of college English students in this approach. According to studies, the level of student participation in classroom learning is the most important factor limiting the efficiency of teaching. The lack of research in this subject greatly limits the “flipped classroom” teaching model’s ability to improve college English classroom teaching quality. The degree of engagement between teachers and students, the enthusiasm of students in class, and the competence of teachers to educate are all reflected in student conduct in the classroom. Understanding and evaluating the behaviors and activities of students in the classroom are helpful in determining the state of students in the classroom, as well as improving the flipped classroom teaching technique and quality. As a result, the convolutional neural network is used to recognize student behavior in the classroom. The loss function of VGG-16 has been enhanced, the distance inside the class has been lowered, the distance between classes has been increased, and the recognition accuracy has improved. Accurate recognition of classroom behavior is beneficial in developing methods to improve teaching quality.

1. Introduction

The flipped classroom teaching model [1–3] has grown into a magnificent landscape of education and teaching reform [4] as a new teaching model [5–7]. Classroom teaching content in traditional classrooms relies on one-way teaching and transmission of book knowledge. The time for students to acquire internalized knowledge in the classroom is very restricted due to classroom teaching time constraints, and the influence of classroom teaching on the development of students’ critical thinking capacity is not ideal. The flipped classroom is learner-centered and problem-oriented, and the classroom flipping of declarative and procedural knowledge allows teachers to devote more time in the classroom to students answering questions and deepening their knowledge understanding. Simultaneously, students bring problems from preclass research into the classroom [8–10], resulting in more focused and successful classroom discussions. In

traditional classrooms, students’ learning is also easy to change from passively listening to lectures or transcribing teachers’ classroom teaching notes. Constructiveness (based on learner experience), structurality (knowledge structure development), criticality (critical assessment of knowledge and viewpoints), comprehension (memory learning that leads to understanding), and reflectiveness are all factors that contribute to efficiency (continuously in the learning process). The phenomena of collective quiet among students in classroom teaching have been overcome by in-depth learning defined by reflection and monitoring.

English acquisition [11, 12], in the final analysis, is to achieve better communication between people. Silence in the process of classroom English acquisition will inevitably lead to obstacles to English learning. At present, university classroom teaching can be roughly divided into “five levels,” namely, silence, answer, dialogue, critical, and debate. Teaching practice shows that the use of the flipped classroom

teaching model in college English classroom teaching can well promote the improvement of foreign language learners' various language skills, greatly improve the quality of college English classroom teaching in colleges and universities, and help deepen the reform of college English education and teaching. The flipped classroom teaching model takes students as the center of classroom teaching, allows students to participate in the process of teaching activities, and greatly improves students' participation and enthusiasm for course learning [13]. In order to explore the impact of flipped learning on English learners' second-language speaking, second-language listening, and participation and participation in curriculum materials and activities outside the classroom, some scholars divided 67 English university freshmen into three groups: structured flipped learning group, semistructured flipped learning group, and the traditional learning group; the research results show that flipped learning helps to improve the oral and listening skills of English learners and enables them to participate more in extracurricular activities. It can be seen that the flipped classroom teaching model can indeed successfully bring the traditional Chinese college English classroom out of the "silent" quagmire and make the classroom teaching model move towards the realm of "dialogue," "questioning," and "debating."

However, the current foreign scholars' research on the flipped classroom teaching model mainly focuses on theoretical research such as teaching model construction, teaching practice exploration and application [14], teaching implementation methods [15, 16], comparison studies with traditional classroom teaching models, and empirical research on the effects of teaching practice after the model is applied. However, most of the academic research on the application of the flipped classroom teaching model of college English focuses on the teacher teaching mode, student learning mode, teacher teaching ability (teaching design, team building, and information literacy), and course assessment methods of the flipped classroom teaching model. However, there are relatively few studies on the degree of student participation in classroom learning behind the phenomenon of classroom activity in the literature. The monitoring and promotion of students' participation [17] in classroom learning belong to the category of teaching quality monitoring. A high degree of student participation in classroom learning is an indispensable prerequisite for any teaching model to ensure its teaching quality. If there is only an increase in the "temperature" of the atmosphere in the classroom, but no deep participation of students in the true sense of learning, then the implementation of the flipped classroom teaching model will greatly lose its true meaning of education.

Since student conduct in the classroom reflects the level of interaction between teachers and students [18, 19], students' interest in class, and teachers' competence to educate, understanding and evaluating the actions and activities of students in the classroom are beneficial to understanding the status of students in class, as well as improving flipped classroom teaching methods and improving the quality of flipped classroom teaching. As a result, convolutional neural networks [20–23] can be used to recognize student behavior in classes [24, 25].

The main contributions of this paper are as follows:

- (1) This paper proposes a teaching quality promotion model for college English flipped classroom based on the assistance of convolutional neural network, which can improve the teaching method of flipped classroom and improve the quality of flipped classroom teaching
- (2) Aiming at the characteristics of similar faces, the loss function of VGG-16 is improved to reduce the intra-class distance and increase the interclass distance. The improved VGG-16 network improves the accuracy of emotion recognition in classroom students

2. Background

At present, the research on class behavior recognition is mainly reflected in the behavior state of students in class, such as raising hands, standing up, and sleeping. There are also some studies that reflect students' abnormal behavior in the examination room. These studies can also be classified as classroom behavior recognition because they are behavior recognition in a classroom setting. The standing recognition algorithm based on the region of interest uses the characteristic that the standing behavior occurs in the upper half of the image, crops the upper half of the image into the region of interest, and realizes standing recognition under different backgrounds through different threshold segmentation algorithms [26, 27]. The student behavior recognition algorithm based on the ResNet network [28] first trains on the ImageNet data set, then uses transfer learning to apply the ResNet network to student behavior recognition, and realizes the recognition of student behaviors such as looking left and right, raising hands, standing, and sleeping. Based on Faster R-CNN to identify the behavior state of students in the classroom, the YOLOV3 algorithm is used to extract the behavior sequence of the students, and finally, the behavior of students is classified through the ResNet network to realize the recognition of the behavior state of the students in the classroom. Based on the gradient histogram and the equivalent local binary mode histogram, the students' head-and-shoulders characteristics are merged, and the support vector machine is used to train the classifier to achieve target detection on the experimental data set. An algorithm based on sparse reconstruction is also proposed. Based on the 3D convolutional neural network, the time dimension is added to the two-dimensional convolutional neural network, which can better learn the time domain features. Simonyan and Zisserman [29] proposed a dual-stream CNN algorithm for human behavior recognition. This algorithm trains two CNN classifiers, one CNN mainly extracts optical flow features, and the other CNN extracts RGB image information, and finally the two classifiers fusion of features. Feichtenhofer et al. [30] proposed a new spatiotemporal structure based on the dual-stream architecture, which has a new convolutional fusion layer and a spatial fusion layer, which can better extract human behavior characteristics.

The detection and analysis of student behavior in the classroom scene can rapidly and efficiently identify the student's learning status and the teacher's teaching quality,

allowing for focused teaching technique changes and increased student learning efficiency. In general, students' behaviors in class include the following: paying attention in class, raising hands, standing up, napping, and using cell phones. Neural networks can be used to monitor and analyze student behavior [31, 32].

3. Methodology

3.1. Flipped Classroom Theory. The essence of the flipped classroom is to return the dominance of learning to the students, lead the students' subjective teaching methods, and have the energy to create education for the future. And "flipped classroom as a teaching concept and teaching model is affecting and changing traditional classroom teaching." It uses Internet technology and information technology to break through the boundaries of traditional classrooms, expand the time and space of classroom teaching, and optimize. The learning process of students enhances students' learning ability, realizes the deep integration of artificial intelligence and curriculum teaching, and promotes students' deep learning. At the same time, the flipped classroom model, as a part of the education reform movement, will completely subvert the traditional printing-based classroom teaching structure and teaching process and trigger a series of changes in the role of teachers, curriculum models, and management models.

Autonomous learning theory, cooperative learning theory, and mastering learning theory are three common theories in flipped classrooms.

Unlike the typical accepted learning approach, autonomous learning places a greater emphasis on pupils' ability to study independently. Students achieve their learning goals through independent analysis, investigation, practice, and invention as the main body of learning. Autonomous learning theory is founded on the principle of inquiry learning, which involves presenting students with a situation in which they must do their own research, solve problems, and gain expertise in the topic. Students in the classroom are given additional opportunity to "experience and interact" with knowledge as part of discovery learning. Traditional education and teaching approaches emphasize passive acceptance as a learning mode. Autonomous learning needs instructors to prioritize school instruction and complement it with required, scientific, and reasonable family and social education, so that children can learn to seek information, live, and survive through autonomous learning and have the ability to adapt to modern society. Encourage pupils to have a more and deeper understanding of knowledge in order to build the required abilities and basic attributes to continuously support their own development.

Cooperative learning is aimed at organizing classroom activities to promote academic and social learning experiences. Cooperative learning is not just about dividing students into small groups, but has been described as "building positive interdependence." In cooperative learning, students must work together in small groups and work together to achieve learning goals. Unlike individual learning, which can be competitive in nature, cooperative learning

students can leverage each other's resources and skills, seek information from each other, evaluate each other's ideas, monitor each other's work, and so on. In addition, the role of the teacher has also changed, from teaching knowledge to facilitating students' learning and from teaching to guiding. Compared with students in individual or competitive learning environments, students in cooperative learning environments achieved more and better reasoning, higher self-esteem and, for example, more social support when a group of classmates completed many learning tasks together.

According to the mastery of learning theory, education should be centered on the amount of time it takes for different students to acquire the same information and achieve the same level of mastery. To put it another way, all pupils can attain the same degree of knowledge as long as they have adequate time. The difference in a student's learning ability is only connected to the amount of time it takes him to master the knowledge, not whether or not he can master the topic. Traditional teaching focuses on the disparities in students' skills, and students' learning time and teaching techniques are essentially the same. Learning theory mastery contrasts sharply with typical teaching methods. Education is no longer just to allow a small group of students to fully learn what the school teaches, but to care about the development of each student and provide all students with the necessary knowledge and skills. In mastering learning, the responsibility of learning has changed. The failure of the students is more due to the factors of guidance, not necessarily the lack of ability of the learners. Therefore, in a mastered learning environment, the challenge becomes to provide students with enough time and appropriate teaching strategies so that all students can reach the same level of learning.

3.2. Improved VGG-16 Network. According to the size of the convolution kernel and the depth of the network, the VGG network model is divided into six network structures, including A, A-LRN, B, C, D, and E. VGG-16 is a class D network structure. VGG-16 consists of 13 convolution layers and 3 full connection layers. Each convolution layer contains the pooling layer and activation function, and the size of the convolution kernel is 3×3 . The network diagram of VGG-16 is shown in Figure 1. The size of the fixed input image of each VGG-16 model is $224 \times 224 \times 3$. The first and second layers are $64 \times 3 \times 3$ convolution kernels with a step size of 1, and the maximum pooling operation is adopted. The third layer and the fourth layer are $128 \times 3 \times 3$ convolution kernels, which also adopt the maximum pooling operation. The size of the convolution kernel in the fifth, sixth, and seventh layers is 33, and the number of convolution kernels is 256. The remaining six convolutional layers have 512 convolution cores, and the maximum pooling operation is performed every three layers. Finally, a comprehensive connection layer with three levels is introduced. The first two layers of the entire connection layer include 4096 neurons, whereas the last layer contains 1000 neurons. Finally, the predicted value is determined using the softmax function's categorization.

VGG-16 was a simple and easy-to-train deep network structure at the time, and it produced good results in image recognition. VGG-16 is separated into various blocks, as

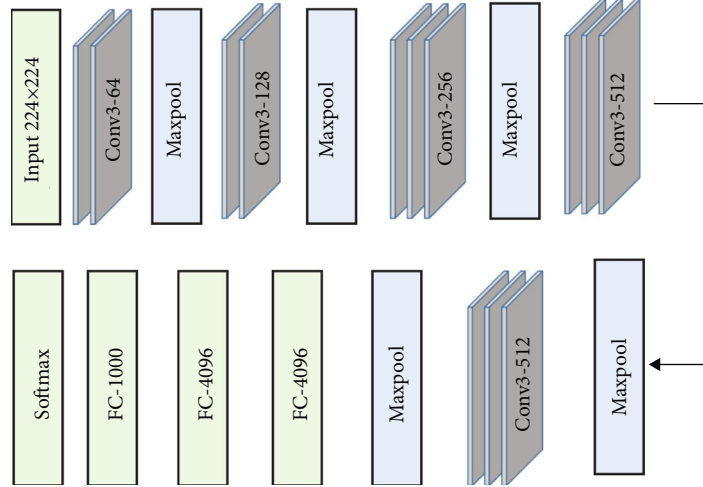


FIGURE 1: VGG-16 network structure diagram.

indicated in Figure 1. There are numerous convolutional layers and a pooling layer in each block. The convolutional layer's number of channels remains constant in the same block. After reaching 512, the number of convolution kernels doubles, doubling the number of channels, and the total number of channels remains unaltered. Despite the fact that many other network topologies have since been presented, VGG-16 remains a popular convolutional neural network model.

The classic VGG-16 network contains three fully connected layers, which leads to too many VGG-16 network parameters and slower training speed. This paper uses a face recognition algorithm based on the improved VGG-16 network to achieve the purpose of student face recognition in the classroom scene. By improving the classic VGG-16 network structure, it is more suitable for face recognition scenes; for face recognition scenes, the loss function is improved. In the task of student face recognition, the student's face is usually centered in the center of the image, and I hope that traits like eyes are learnt in various areas across the image. As a result, the last layer of VGG-16, as well as the third-to-last convolutional layer, is changed to a local convolutional layer in this article, allowing various parts of the image to learn completely separate features. At the same time, the VGG-16 network's fully connected layer is lowered by two layers, and the last layer's pooling layer is enhanced to an average pooling layer. 13 convolutional layers, 5 pooling layers, and 1 fully connected layer make up the upgraded VGG-16 network.

3.3. Improved Loss Function. I propose using center loss to build a mixed loss function that will increase the model's discriminative ability as well as its generalization capacity. Softmax loss can categorize images, and the central loss function can increase the distance between classes while decreasing the distance inside the class, as well as improve the model's accuracy. The center loss function's calculation equation is as follows:

$$L_C = \frac{1}{2} \sum_{i=1}^m \|x_i - C_{y_i}\|_2^2. \quad (1)$$

And the calculation equation of the softmax loss function is as follows:

$$L_S = - \sum_{i=1}^m \log \frac{\exp(w_{y_i}^T x_i + b_{y_i})}{\sum_{j=1}^n \exp(w_j^T x_i + b_j)}. \quad (2)$$

The calculation equation of the improved mixed loss function is as follows:

$$L = L_S + \lambda L_C, \quad (3)$$

where λ is the weight of the central loss function.

3.4. Classroom Student Behavior Recognition. The teaching quality promotion strategy of the English flipped classroom proposed by us is divided into two parts. First, I use the improved VGG-16 for student face recognition and behavior recognition. Second, I evaluate the recognition results and give each student developed teaching strategies and plans.

3.4.1. Classroom Student Behavior Recognition. The improved VGG-16 network structure for student face and behavior recognition is shown in Figure 2.

As shown in Figure 2, what I propose is a convolutional neural network architecture with twin parallel branches. The upper branch is mainly used for student facial emotion recognition, and the lower branch is mainly used for subject behavior recognition. And through the optimized softmax loss and center loss, respectively, the proposed algorithm has a higher accuracy rate.

3.4.2. Teaching Quality Promotion Strategy. Figure 3 shows the teaching quality promotion strategy I proposed in flipped classroom. First of all, I evaluate the students' face recognition and behavior recognition, respectively. The evaluation is based on the score of the control group.

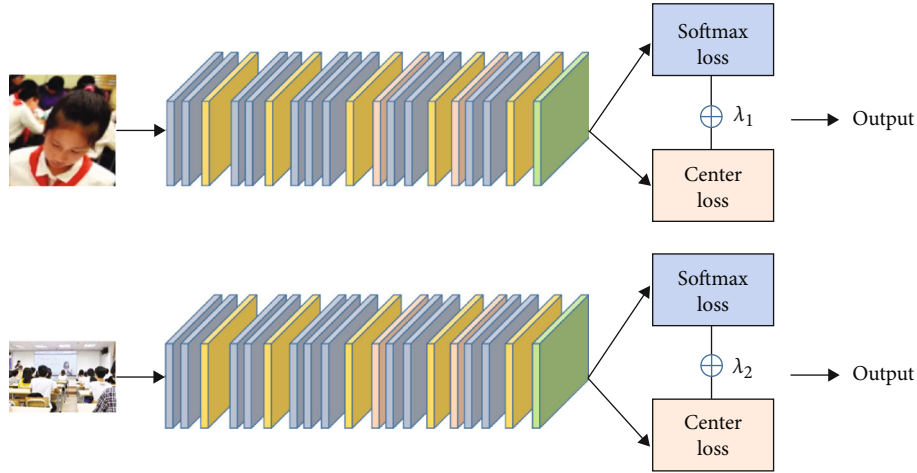


FIGURE 2: Schematic diagram of student face and behavior recognition architecture.

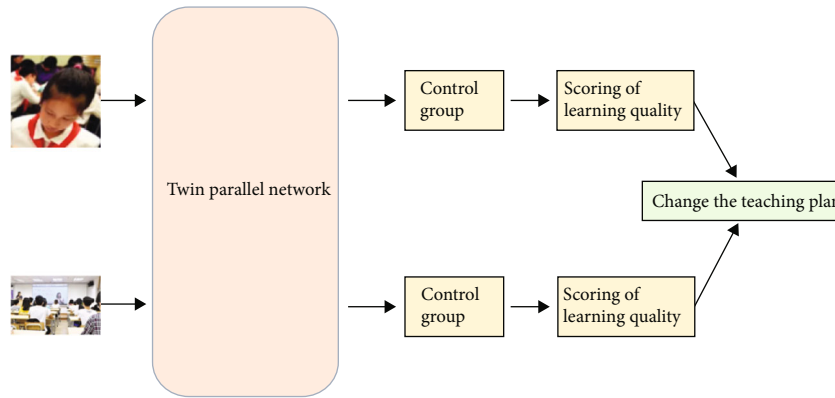


FIGURE 3: Schematic diagram of teaching quality promotion strategy framework.

4. Experiments

4.1. Environment Configuration. The experimental hardware environment in this paper is dual-channel Xeon E5 2678 V3 processor, the main CPU is 2.5 GHz, a total of $12 * 2$ cores, the memory is 16 G, and the graphics card is NVIDIA RTX2060 Super (8 G video memory). The software environment of this experiment is Ubuntu 18.04 LTS operating system, the programming language is Python, the deep learning framework is TensorFlow and Caffe, and the GPU acceleration is carried out using CUDA10.0+ CUDNN7.6.4. The hyperparameter settings are shown in Table 1, and we divide the data set into a training set and a test set, 70% and 30%, respectively, and the batch size is 100.

4.2. Data Sets. The data set used to train face recognition models in this study was Labeled Faces in the Wild (LFW). Face identification in an uncontrolled context is studied using the LFW data set, which is a face photo database. More than 13,000 facial photographs were collected from the Internet, with each face annotated with the name of the person in the shot, and 1,680 people in the sample had at least two different photos. The LFW data set is utilized to train a convolutional neural network in this paper, with some students from self-recorded films added to the data set for training.

TABLE 1: Hyperparameter settings.

Type	Hyperparameter
OptimizerA	Adam
Learning rate	0.0001
bata_1	0.99
bata_2	0.99
Epsilon	$3e-8$
Decay	$3e-8$

In addition, the data set of students' classroom behavior comes from our collection in 3 months.

4.3. Evaluation Methods. The quality of the teaching quality promotion strategy of college English flipped classroom depends on the accuracy of face and behavior recognition. Therefore, I use the evaluation method of classification problem to evaluate the algorithm of this article. In the classification problem, the commonly used indicators include accuracy rate, recall rate, and accuracy rate. Accuracy is the ratio of the number of samples that are predicted to be positive to all samples that are predicted to be positive. The calculation equation is as follows:



FIGURE 4: The effect of face recognition of students in classroom.

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (4)$$

$$\text{Re call} = \frac{TP}{TP + FN}, \quad (5)$$

$$\text{ACC} = \frac{TP + TN}{TP + FN + FP + TN}. \quad (6)$$

4.4. Experimental Results. The classroom face recognition effect based on the improved VGG-16 network is shown in Figure 4. It can be seen from the figure that, except for the occlusion phenomenon of individual students, the faces of other students can be detected. In addition, it can be found that the improved VGG-16 network can also identify the information of the students in the back row, but the matching degree of the students in the back row is still flawed. The faces of individual students can be detected, but they cannot be matched. In the figure, there is a student whose face is occluded, and the algorithm has not recognized it. This is related to the lack of features provided by the student's face.

Table 2 shows the test results of the improved VGG-16 network face recognition method on the LFW data set. Compared with the classical VGG-16 network model, the accuracy of the improved VGG-16 network model in classroom face recognition is improved by 2.6%. In the study of classroom behavior recognition, student face recognition is mainly used to verify student information. For the verification of standing up students' information, the accuracy rate of face recognition reaches 96.81, fully meeting the expected requirements.

In summary, the face recognition algorithm based on the improved VGG-16 has achieved the expected effect. According to the experimental results, it can be concluded that facial features can be used as a method of confirming student identity information when performing classroom behavior recognition.

4.5. Ablation Experiments. In order to verify the effectiveness of softmax loss and center loss in the proposed method, an ablation experiment is set up in this section; SL stands for softmax loss, CL stands for center loss, T stands for upper

TABLE 2: Comparison result of recognition accuracy.

Method	Precision	Recall	ACC
VGG-16	93.71	84.25	84.23
VGG-19	94.22	85.26	83.25
Ours	96.81	88.01	89.65

TABLE 3: The results of the ablation experiments.

Method	Precision	Recall	ACC
T-SL-SL,B-SL-CL	92.15	81.63	83.65
T-SL-SL,B-CL-CL	91.69	82.65	84.12
T-SL-CL,B-CL-CL	90.25	84.25	85.25
T-CL-CL,B-CL-CL	92.65	84.96	81.56
T-SL-CL,B-SL-CL	96.81	88.01	89.65

TABLE 4: The results of the ablation experiments for VGG.

Method	Precision	Recall	ACC
T-VGG16,B-VGG19	91.11	82.23	84.12
T-VGG19,B-VGG16	92.25	83.69	85.12
T-VGG19,B-VGG19	91.10	83.36	84.14
T-VGG16,B-VGG16	96.81	88.01	89.65

branch, B stands for lower branch, and the results of the ablation experiment are shown in Table 3.

It can be seen from Table 3 that when the upper branch and the lower branch use softmax loss and center loss at the same time, the best performance is achieved. Therefore, this proves that the proposed method is scientific and effective.

4.6. Ablation Study for VGG. In this paper, a VGG ablation experiment was carried out. Considering the complexity and cost of the model, we chose VGG16 and VGG19 for the ablation experiment. The experimental results are shown in Table 4.

It can be clearly seen from Table 4 that the upper and lower branches use VGG16 to obtain the best performance, which proves that the proposed method is effective. In addition, we found that the upper branch uses VGG19 and the lower branch uses VGG16 to obtain suboptimal performance.

5. Conclusion

The degree of student participation in classroom learning is the main factor that restricts the effectiveness of teaching. The lack of research in this subject greatly limits the “flipped classroom” teaching model’s ability to improve college English classroom teaching quality. The degree of engagement between teachers and students, the enthusiasm of students in class, and the competence of teachers to educate are all reflected in student conduct in the classroom. Understanding and evaluating the behaviors and activities of students in the classroom are helpful in determining the state of students in the classroom, as well as improving the flipped classroom teaching technique and quality. As a result, the convolutional neural network is used to recognize student behavior in the classroom. The loss function of VGG-19 has been enhanced, the distance within the class has been lowered, the distance between the classes has been increased, and the recognition accuracy has improved. Recognizing classroom behavior aids in the development of teaching quality improvement initiatives. In addition, the experimental results show that the proposed method achieves an accuracy of 96.81% and achieves a competitive performance.

In the following research, we will focus on the processing and identification of real-time data.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author does not have any possible conflicts of interest.

Acknowledgments

The study was supported by the “science research project of Inner Mongolia Autonomous Region Higher Learning Institutions (Grant No. NJSY20311).”

References

- [1] Z. Turan and B. Akdag-Cimen, “Flipped classroom in English language teaching: a systematic review,” *Computer Assisted Language Learning*, vol. 33, no. 5-6, pp. 590–606, 2020.
- [2] T. Tang, A. M. Abuhmaid, M. Olaimat, D. M. Oudat, M. Aldhaeabi, and E. Bamanger, “Efficiency of flipped classroom with online-based teaching under COVID-19,” *Interactive Learning Environments*, pp. 1–12, 2020.
- [3] R. Kay, T. Mac Donald, and M. DiGiuseppe, “A comparison of lecture-based, active, and flipped classroom teaching approaches in higher education,” *Journal of Computing in Higher Education*, vol. 31, no. 3, pp. 449–471, 2019.
- [4] P. Zheng, X. Wang, and J. Li, “Exploration and practice of curriculum ideological and political construction reform——take” information security” course as an example,” *ASP Transactions on Computers*, vol. 1, no. 1, pp. 1–5, 2021.
- [5] L. Liang, Q. Yin, and C. Shi, “Exploring proper names online and its application in English teaching in university,” *ASP Transactions on Computers*, vol. 1, no. 1, pp. 24–29, 2021.
- [6] S. S. Ismail and S. A. Abdulla, “Virtual flipped classroom: new teaching model to grant the learners knowledge and motivation,” *Journal of Technology and Science Education*, vol. 9, no. 2, pp. 168–183, 2019.
- [7] Ö. Canaran and İ. H. Mirici, “A new model of team teaching for teacher professional development: a case study of in-service English teachers,” *Egitim ve Bilim*, vol. 45, no. 201, 2020.
- [8] J. L. Jensen, E. A. Holt, J. B. Sowards, T. H. Ogden, and R. E. West, “Investigating strategies for pre-class content learning in a flipped classroom,” *Journal of Science Education and Technology*, vol. 27, no. 6, pp. 523–535, 2018.
- [9] J. Jovanovic, N. Mirriahi, D. Gašević, S. Dawson, and A. Pardo, “Predictive power of regularity of pre-class activities in a flipped classroom,” *Computers & Education*, vol. 134, pp. 156–168, 2019.
- [10] R. M. Yilmaz and O. Baydas, “An examination of undergraduates’ metacognitive strategies in pre-class asynchronous activity in a flipped classroom,” *Educational Technology Research and Development*, vol. 65, no. 6, pp. 1547–1567, 2017.
- [11] J. K. Klingner and S. Vaughn, “Promoting reading comprehension, content learning, and English acquisition through Collaborative Strategic Reading (CSR),” *The Reading Teacher*, vol. 52, no. 7, pp. 738–747, 1999.
- [12] C. McKee, “A comparison of pronouns and anaphors in Italian and English acquisition,” *Language Acquisition*, vol. 2, no. 1, pp. 21–54, 1992.
- [13] J. J. Steel, “Microbial murders crime scene investigation: an active team-based learning project that enhances student enthusiasm and comprehension of clinical microbial pathogens,” *Journal of Microbiology & Biology Education*, vol. 18, no. 2, p. 18, 2017.
- [14] K. Van Lehn, S. Ohlsson, and R. Nason, “Applications of simulated students: an exploration,” *Journal of Artificial Intelligence in Education*, vol. 5, pp. 135–135, 1994.
- [15] T. E. Dorgu, “Different teaching methods: a panacea for effective curriculum implementation in the classroom,” *International Journal of Secondary Education*, vol. 3, no. 6, p. 77, 2016.
- [16] V. A. Dinh, J. Frederick, R. Bartos, T. M. Shankel, and L. Werner, “Effects of ultrasound implementation on physical examination learning and teaching during the first year of medical education,” *Journal of Ultrasound in Medicine*, vol. 34, no. 1, pp. 43–50, 2015.
- [17] J. Chen, “Research on the promotion of Chinese vocational college students’ participation in online and offline blended teaching,” *International Journal of New Developments in Education*, vol. 2, no. 3, 2020.
- [18] P. Cooper and D. McIntyre, “Patterns of interaction between teachers’ and students’ classroom thinking, and their implications for the provision of learning opportunities,” *Teaching and Teacher Education*, vol. 10, no. 6, pp. 633–646, 1994.

- [19] N. Mercer and L. Dawes, "The study of talk between teachers and students, from the 1970s until the 2010s," *Oxford Review of Education*, vol. 40, no. 4, pp. 430–445, 2014.
- [20] J. Zhang, X. Jin, J. Sun, J. Wang, and A. K. Sangaiah, "Spatial and semantic convolutional features for robust visual object tracking," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 15095–15115, 2020.
- [21] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.
- [22] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.
- [23] S. Qi, X. Ning, G. Yang et al., "Review of multi-view 3D object recognition methods based on deep learning," *Displays*, vol. 69, article 102053, 2021.
- [24] P. Chonggao, "Simulation of student classroom behavior recognition based on cluster analysis and random forest algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 2, pp. 2421–2431, 2021.
- [25] A. Vilorio, J. P. Lis-Gutiérrez, M. Gaitán-Angulo, A. R. M. Godoy, G. C. Moreno, and S. J. Kamatkar, "Methodology for the design of a student pattern recognition tool to facilitate the teaching-learning process through knowledge data discovery (big data)," in *International conference on data mining and big data*, pp. 670–679, Springer, Cham, 2018.
- [26] W. Cai, Z. Wei, Y. Song, M. Li, and X. Yang, "Residual-capsule networks with threshold convolution for segmentation of wheat plantation rows in UAV images," *Multimedia Tools and Applications*, pp. 1–17, 2021.
- [27] D. P. Campos, P. J. Abatti, F. L. Bertotti, J. A. G. Hill, and A. L. F. da Silveira, "Surface electromyography segmentation and feature extraction for ingestive behavior recognition in ruminants," *Computers and Electronics in Agriculture*, vol. 153, pp. 325–333, 2018.
- [28] J. Zhang, J. Sun, J. Wang, and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 8427–8440, 2021.
- [29] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," 2014, <https://arxiv.org/abs/1406.2199>.
- [30] C. Feichtenhofer, A. Pinz, and A. Zisserman, "Convolutional two-stream network fusion for video action recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1933–1941, Las Vegas, NV, USA, 2016.
- [31] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.
- [32] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.