Journal of Contemporary Educational Research, 2024, Volume 8, Issue 11
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

Implementation of Digital Classroom State System for Teachers and Students Based on Large Models

Wenbo Lyu*, Guangmin Zhu, Ziyi Qin, Mengting Yan, Jie Zhang

Saxo Business School of Fintech, University of Sanya, Sanya 572000, Hainan Province, China

***Corresponding author:** Wenbo Lyu, wenbolv@sanyau.edu.cn

Abstract: Deep learning has become a hot field of artificial intelligence, and the deep learning large model framework has become a bridgehead for the active layout of Chinese and foreign technology companies. Large models play a significant role in the application field, greatly improving the efficiency of training and optimization, and contributing to the landing of many innovative artificial intelligence tools. Based on the Chinese PaddlePaddle large model framework, an application system is designed in combination with the intelligent classroom teaching scenario, which uses machine vision algorithms to distinguish and present teachers' and students' behaviors, that is, the digitization and multi-classification scheme of class character states. After having digital data, data analysis can be carried out to evaluate the class status of teachers and students, and the traditional subjective judgment such as peacetime grades and teaching ability can be upgraded to the objective judgment of artificial intelligence.

Keywords: Large model; Machine vision; Digitalization; Status; Technical realization

Online publication: November 27, 2024

1. Introduction

As deep learning techniques continue to evolve, large models excel in the fields of image recognition, large-scale image classification, and object detection tasks. The realization of these technologies cannot be separated from the rapid development of deep learning models and the support of massive data. Large models have powerful learning and inference capabilities. It can be trained to learn a variety of tasks and data and make accurate judgments and predictions when reasoning.

The traditional smart classroom emphasizes more on the upgrading of teaching equipment and the digital presentation of teaching results, but the evaluation and statistics of these dimensions are difficult to some extent. Firstly, there is no extra manpower to make special statistics. Secondly, teachers tend to be more subjective and one-sided in recording students' status in class. However, it is necessary to conduct digital and statistical evaluations of each teacher and student's status, which is part of their performance. The

use of machine vision and artificial intelligence to achieve classroom teaching effect evaluation has strong practical value and significance.

In 2019, Liu *et al.* [1] built a classroom teaching behavior model of collection, modeling, and analysis, which mainly evaluates teacher behavior. In 2018, Duan [2] modeled the indicators of students' concentration in class, but the algorithm was only applied to the environment of small classrooms. In 2019, Jia *et al.* [3] realized the deep learning algorithm to detect and classify more than 60 key variables, but it is difficult to apply in a large classroom environment. In 2019, Qin [4] used a transfer learning algorithm to research classroom behavior recognition methods, but the training set was only pictures, not teaching activity videos, and the expansion was low. In 2020, Xu *et al.* [5] took real teaching videos captured by surveillance as the training set, and innovated the use of human skeleton information to automatically identify students' large-scale behaviors. In the same year, Wang [6] designed the application of artificial intelligence in educational videos. In the summer of 2021, Xia *et al.* [7] used the Father-RCNN model to divide student behavior into eight categories, which put forward high requirements on the efficiency of real-time recognition of hardware computing power. In 2021, Xiong [8] enhanced the application value of machine vision in smart classrooms, and Zhang [9] expounded on the enhancement role of 5G in artificial intelligence smart classrooms.

Therefore, artificial intelligence teaching reform has been the trend of the times, but it is still mainly based on theoretical and experimental exploration. In the real intelligent classroom environment, there exist some problems, such as difficulty in applying to large classrooms with many people, limited classification of students' behaviors, low recognition efficiency and high power consumption, and inability to extract and identify student concentration indicators.

To solve a wide range of classroom effect evaluation problems more objectively and accurately, this paper designs a digital identification and presentation system for teachers' and students' status in smart classrooms based on the PaddlePaddle large model framework. It can not only obtain the judgment results of teachers' and students' status in real-time but also prompt teachers and students to pay attention to classroom discipline and concentration and optimize the allocation of teaching resources. The students in good condition are encouraged and affirmed, and the students in poor condition are reminded, which improves the overall teaching efficiency of the class. At the same time, the work of non-teaching core figures such as attendance, question and answer records, and performance management is completed by machine vision and artificial intelligence. After the event, the system's machine algorithm can automatically realize the evaluation and prediction of the digital results, and present the digital evidence of the students' long-term learning ability and attitude, to facilitate the guidance of human means such as counselors.

2. System overview

PaddlePaddle large model framework, as a deep learning platform independently developed by Baidu, is the only deep learning platform in China that provides systematic deep learning technical service support. Compared with similar platforms abroad, PaddlePaddle has higher cost performance and technical feasibility. It has a convenient development framework and supports both dynamic chart and static chart programming, which can take into account ease of use and efficiency. It has rich model options and a large number of models precipitated in industrial practice, which is greatly convenient for the preparation of this paper. It also has large-scale distributed scenario training capability. The most important thing is that its end-to-end deployment capability is very complete, which can make the reasoning and prediction process in this

paper smooth, convenient deployment of deep learning applications to the computer terminal platform, and compatible with the mainstream camera equipment and mainstream video software in the market.

Based on the above advantages, this paper chooses the PaddlePaddle large model framework as the main channel of technical realization. Aiming at such a scenario of character status recognition in a smart classroom, this paper makes full use of the PaddleDetection (ppdet) object detection model and is committed to improving the efficiency of education management. Through object detection technology, the system accurately identifies all faces in the classroom, and on this basis carries out intelligent classification of students' class status, including three states: Normal (marked with a green box and Normal), Abnormal (marked with a red box and Abnormal), and Unknown (marked with yellow box and Unknown). The normal state is defined as the student looking up and paying attention to the lecture. The system accurately captures the student's attention state by analyzing the student's facial expression and head posture and provides the teacher with real-time engagement feedback. The abnormal state includes the behavior of the lower head and distraction, etc. The system accurately determines whether the students' attention deviates from the class through detailed analysis of the students' movements and recognition of the head direction. The setting of unknown states further takes into account some possible limitations of facial recognition, such as the face being blocked or lack of light, and the system determines these situations as unknown states, alerting teachers to possible problems. The specific classification criteria of normal and abnormal are divided into four categories, namely facial expression classification, student concentration direction classification, face relative position classification, and face relative distance classification. The logical basis is that the expression of the normal state of listening to the class should conform to the public expression, rather than a specific serious expression or smile. When teachers teach humorous content, they should generally smile, and for serious content, they should be generally serious. Therefore, by classifying a few characteristic abnormal expressions to distinguish, the direction of student focus should be the teacher's direction. In this training, the relative positions of the teacher and the camera are close and unchanged, so the default is the camera direction. Although the situation that students sometimes lower their heads to read books is considered, it is marked as abnormal state in the marking and classification process, but in the later data evaluation and prediction, the threshold classification will be made for the proportion of the total time of abnormal state, which is equivalent to increasing the tolerance of the state of lowering their heads to read books. The distance between adjacent faces is assessed, with a too-close distance indicating an abnormal state of people gathering to converse. Additionally, the relative positions of the adjacent environment—such as the faces and shoulders in relation to the table and chairs—are evaluated. A situation in which a person is sleeping on their stomach is also considered too close and marked as abnormal. Conversely, situations that do not exhibit these characteristics are marked as normal. At the same time, the confidence degree of the marked state is given. Considering the reliability and efficiency of machine learning, the confidence degree is used to achieve error tolerance of the real-time state and the training state. The data with less than 50% confidence (0.5) can be discarded and saved on the contrary. The system is set to record the data digitally every five seconds, and automatically generate and save the data in the Excel document named after the time, forming a time series of the class state digitalization, for further data analysis and processing.

The system is equipped with a data generation function, through which users can save multi-frame forecast data as a local CSV file. Data saving is displayed in terms of time, rectangular box coordinates, forecast status, and forecast confidence to provide users with a comprehensive perspective of data analysis. The timestamp of each frame records the time when the data was collected, and the rectangular frame

coordinates clarify the position information of each student in the class. The predicted state details the student's state in that frame, such as normal, abnormal, or unknown. The predicted confidence level reflects the system's level of confidence in the predicted outcome. This function enables users to store, view, and analyze multi-frame data conveniently, which provides strong support for teaching management. Through the form of a CSV file, users can look back at the historical data at any time, an in-depth understanding of students' learning status changes, to better adjust and improve teaching. This flexible and detailed data generation function provides strong support for the real-time monitoring technology of the system. **Figure 1** shows the digital data of teachers' and students' status. Considering the security and confidentiality of digital data, the system has designed a power-on password link, so that the system has a safe and reliable password login function, using the default fixed password, only the authorized person can access the program. This ensures the privacy of the user's identity and the security of access rights. Considering that in most cases there is video first, and then the program system is used for digitization and data analysis, two different paths of calling camera or retrieving video recording are set up when the system is started to achieve these functions.

Figure 1. System real-time discriminant effect rendering expectation

3. Technical implementation

To achieve all the functions of the above system, the following steps are followed for technical implementation (some non-core code refer to the official website of the PaddlePaddle public tutorial).

Technical realization is divided into two major plates, the first plate is the training, optimization, and export of the model. The second part is the visualization of the program on the computer side and the operation of the Windows system.

After the environment is installed, data labeling is carried out to generate training data. The wizard assistant software is used to classify and label a large number of video screenshots of various classroom environment backgrounds, and at the same time, the green screen background of typical abnormal state categories is made and labeled. The number of annotations for each scene is ensured to be more than 1000, and finally the annotated data are compressed and uploaded to the AI for training based on the ppdet model.

```
# Annotate and train datasets
! Unzip the -o data/test data. The zip - d PaddleDetection/dataset/test_det
# Load split file package
! python split_data.py
```

```
# configure datasets to load YML files
%cd /home/studio/PaddleDetection/
! pip install -r requirements.txt
! pip install paddle-serving-client
%cd /home/studio/PaddleDetection/
! python tools/train.py -c configs/ppyolo/ppyolo_r18vd_coco.yml\
--eval\
--use_vdl True
-r output/ppyolo_r18vd_coco/21
# set reads the configuration parameters and model validation of a data set
 % CD/home/studio/PaddleDetection /
! python tools/eval.py -c configs/ppyolo/ppyolo_r18vd_coco.yml\
-o weights='output/ppyolo_r18vd_coco/25'
# Make predictions of the model after validation
%cd /home/aistudio/PaddleDetection/
! python tools/infer.py -c configs/ppyolo/ppyolo_r18vd_coco.yml\
-o weights='output/ppyolo_r18vd_coco/21'\
--infer_dir 'dataset/test_det/JPEGImages'\
--output_dir 'output'\
--draw_threshold 0.1\
--save_txt True
# good forecast result, can export model
 % CD/home/aistudio/PaddleDetection /
! python tools/export_model.py -c configs/ppyolo/ppyolo_r18vd_coco.yml\
-o weights='output/ppyolo_r18vd_coco/21'\
--output_dir '/home/aistudio/export_model'\
--export_serving_model True
```

At this point, the model training based on ppdet object detection suite has been completed, and exported to the terminal for deployment.

4. Conclusion and outlook

The era of the big model has quietly arrived, and the big model of education has great application value and potential. Many researchers make unremitting efforts to explore how to integrate the big model of education into the teaching activities of smart classrooms. This topic is based on the digital presentation of the status data of teachers and students and studies the plan to realize the generation and processing of digital data, which can not only make an objective and true evaluation of students' listening status but also make a new dimension evaluation of teachers' teaching quality. This result also has some limitations. For example, due to the limitation of the training data set, the recognition accuracy rate is higher in the training set classroom, conventional light, and training background environment. How to realize character status recognition and digitization in general gathering occasions is the direction of further research in the future.

An attractive, harmonious, and friendly classroom state is an important feature of the high-quality

development of Chinese modern education. The results of this project can better evaluate the classroom quality according to this requirement, greatly simplify the traditional evaluation process of experts, and improve efficiency. Good evaluation has a positive role in promoting the improvement of teachers' and students' literacy. It is believed that with the continuous development of China's digital economic model, the connotation and extension of smart classrooms will be more abundant, and science and technology education and educational technology will usher in a new era of mutual integration and mutual promotion.

Funding

(1) Education Department of Hainan Provincial (Hnky2024-43)
(2) Sanya University's Industry-Education Integration Project (USY-CJRH2313)
(3) Financial Innovation and Multi- Asset Intelligent Trading Laboratory of the Key Laboratory of Philosophy and Social Sciences in Hainan Province of University of Sanya.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Liu QT, He HY, Wu LJ, et al., 2019, Classroom Teaching Behavior Analysis Method Based on Artificial Intelligence and Application. China Audio-Visual Education, (09): 13–21.

[2] Duan JL, 2018, Analysis and Evaluation System of Students' Concentration in Class Based on Machine Vision, dissertation, Zhejiang Gongshang University, 3–15.

[3] Jia PY, Zhang CH, Zhao XY, et al., 2019, Analysis of Classroom Student Status Based on Artificial Intelligence Video Processing. Modern Educational Technology, 29(12): 82–88.

[4] Qin DY, 2019, Student Classroom Behavior Recognition Based on Deep Learning, dissertation, Central China Normal University, 9–12.

[5] Xu JZ, Deng W, Wei YT, et al., 2020, Automatic Recognition of Student Classroom Behavior Based on Human Skeleton Information Extraction. Modern Educational Technology, (05): 108–113.

[6] Wang P, 2020, Application Analysis and Design of Artificial Intelligence in Educational Videos. Journal of Audio-Visual Education Research, 41(03): 93–100 + 121.

[7] Xia DX, Tian XY, Tang SN, et al., 2021, Analysis of Students' Classroom Behavior Based on Visual Attention. Journal of Guizhou Normal University (Natural Science Edition), 39(04): 83–89 + 120.

[8] Xiong JR, 2021, A Machine Vision Facial Expression Recognition and Detection System. Electronic Production, (04): 39, 40 + 38.

[9] Zhang K, 2021, Application Prospects of Artificial Intelligence Smart Classroom in the 5G Era. Post and Telecommunications Design Technology, (06): 84–87.