EAI Endorsed Transactions
on Scalable Information Systems

Research Article EAI.EU

Improved Convolutional Neural Network Algorithm for Student Behavior Detection in the Classroom

Yihua Liu[1] and Weirong Wang[1,*]

[1] Xi'an Siyuan University, Xi'an,710038, Shaanxi, China

Abstract

The performance of the existing student classroom behavior detection model is affected by various aspects such as dataset, algorithm and height as well as the differences between different classrooms, and there are problems such as a single dataset, low accuracy and low efficiency. In order to improve the accuracy of student classroom behavior detection algorithm, this paper proposes a student classroom behavior detection method based on improved convolutional neural network algorithm. Firstly, the student behavior detection dataset is constructed, and the student classroom behavior detection technology scheme is designed; secondly, in order to improve the detection accuracy, the features are extracted by using the new jumping bi-directional paths, and the attention mechanism module is added at different positions to improve the path aggregation network; weekly, the embedding positions of the attention mechanism strategy are determined by analyzing multiple sets of experiments, and the proposed student classroom behavior detection algorithm's effectiveness and superiority.

Keywords: student classroom behavior detection, convolutional neural network, path aggregation network, attention mechanism strategy

Received on 21 January 2024, accepted on 23 April 2024, published on 2 May 2024

doi: 10.4108/eetsis.5872

*Corresponding author. Email: wwrchjkk@163.com

1. Introduction

Classroom teaching activities as the focus of research in the field of education, with students as the main body of classroom teaching activities, the use of students' classroom behavioral status to respond to classroom efficiency, an important reference indicator for evaluating classroom quality [1]. With the rapid development of artificial intelligence technology, school education and teaching intelligence has also gradually developed, and has received the attention and research of experts and scholars in the field of education [2]. Students' classroom behavior status is usually recorded by classroom video equipment or sensors, combined with artificial intelligence, deep learning, intelligent optimization algorithms and other related technologies to analyze students' behavior status data, design and development of intelligent detection and analysis of students' classroom behavior technology is not only conducive to the classroom education reform, but also from the side of the evaluation of the quality of teachers' classroom [3]. Currently, the classroom evaluation method based on expert experience is not only time-consuming and laborious, but also lacks objectivity, and the development of artificial intelligence, image processing, deep learning and other technologies bring new ideas to solve the problem of teaching evaluation based on the detection of students' classroom behaviors, which not only improves the accuracy of evaluation, but also realizes the intelligentization of teaching evaluation, which is of great significance to improve the efficiency of teaching evaluation and promote the development of education [4].

Student classroom behavior detection is to obtain students' classroom behavior status through classroom monitoring equipment, and identify and analyze them, and use the results of identification and analysis to improve teaching methods and improve teaching quality [5]. Currently, there is a lack of research on student classroom behavior detection and identification [6], and the research methods mainly include machine learning and deep learning algorithms [7], and the research content mainly includes research on students' classroom attendance, classroom concentration, and abnormal behavior detection [8]. Literature [9] collects monitoring data, constructs a student behavior database, and uses VGG network design to implement an abnormal behavior identification method for student behavior; Literature [10] proposes a classroom behavior identification method for students based on three-dimensional convolutional neural network; Literature [11] adopts Bayesian network to select the target features, and constructs a student behavior detection model; and Literature [12] proposes a student behavior-based intelligent teaching assessment method from the spatial and temporal dimensions to construct a student behavior recognition model based on deep spatio-temporal residual convolutional neural network, and achieved a high accuracy rate; Literature [13] used wearable smart bracelets to analyze learners' handwriting state and heart rate changes to analyze learners' current state; Literature [14] used support vector machine students' head gestures, hand movements, and facial information to analyze the relationship between the degree of student classroom participation degree and the relationship between each feature; Literature [15] for the current detection algorithms on student behavior leakage detection problem, the use of K-means and region division strategy to improve the accuracy of the student classroom behavior detection model; Literature [16] proposed based on the Rs-YOLOv3 detection algorithm for the detection of abnormal behavior of the student method. For the above literature analysis, the performance of the existing student classroom behavior detection model is affected by various aspects such as dataset, algorithm and height as well as the differences between different classrooms, and there are problems such as a single dataset, low accuracy and low efficiency [17].

Aiming at the problems of current student classroom behavior detection methods, this paper proposes a student classroom behavior detection method based on improved convolutional neural network. In this paper, we constructed a student behavior detection dataset, designed a technical scheme for student classroom behavior detection, optimized the convolutional neural network by using jumping bidirectional path and attention mechanism module, and at the same time, lightweighted the detection network. Finally, the performance of the improved algorithm is analyzed through multiple sets of experiments to verify the effectiveness and robustness of the algorithm improvement.

2. Building a database for student behavior detection

2.1. Definition of Student Behavioral Classification

Students' classroom behaviors are mainly analyzed and defined from two perspectives: the independent learning stage and the guided independent learning stage [18]. From the viewpoint of students' independent learning stage, students' learning process is more autonomous, and common classroom behaviors include studying, playing cell phones, and sleeping; from the viewpoint of guided independent learning stage, students' classroom behaviors mainly include studying, sleeping, and drinking water. Therefore, in order to construct the students' classroom behavior monitoring model, the analysis and definition of the type of students' classroom behaviors include four types of behaviors such as studying, playing cell phone (phone), sleeping, and drinking [19], and the description of the specific classroom behavior standards is shown in **Figure 1**.

No.	Behavior	Description standard
1	Study	Reading, writing and looking at blackboard
2	Phone	Touching phone with hand
3	Sleeping	Laying on the table
4	Drinking	Holding a water bottle in hand

Figure 1 Description of student behavioral judgments

2.2. Data Acquisition and Labeling

In order to fully obtain sufficient sample data, this paper uses a sample dataset in which large and small classes, different population densities, different classroom lighting environments, different shooting angles and other complex influences are all considered [20]. In this paper, the first surveillance video and the second cell phone recording video are used to capture the university classroom video in two ways. According to the demand of video number collection, single person learning video, multi-person learning video, small classroom video, large classroom video; according to the demand of teaching classroom scene selection, ordinary classroom classroom scene, multimedia classroom classroom scene; according to the angle of video collection, the front camera and rear camera in surveillance video are used to shoot the video. According to the above collection requirements, the student behavior detection data set is obtained, and the schematic diagram is shown in **Figure 2**.

Figure 2 Data acquisition effect diagram

In this paper, we use the dataset in VOC format with Labeling annotation tool [21], and the specific interface of the annotation tool is shown in **Figure 3**. The specific process of dataset annotation includes: 1) download Labeling software and open it; 2) set the path to save the annotation file; 3) name the annotation according to the classification, which is mainly divided into four categories: studying, playing cell phone, sleeping and drinking, and the annotation is set to study, phone, sleeping and drinking; 4) generate the XML file; 5) name the XML file with the corresponding image in a unified format, and store it in the Annotation Tool [22]. The XML files and corresponding images are named in a uniform format and stored in Annotations and JPEGImages respectively.

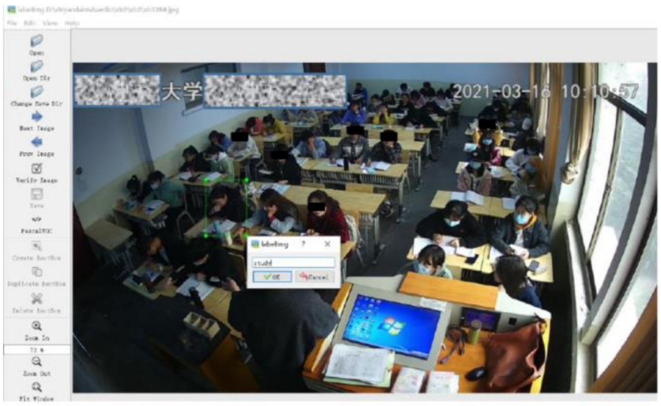


Figure 3 Data acquisition effect diagram

According to the above operation, finally 10,288 college student behavior detection datasets and 10,288 corresponding XML files are obtained.

3. Design of a technology solution for student behavior detection in the classroom

3.1. Detection algorithm selection

Current student classroom behavior detection algorithms include video-based semantic information class methods and image-based semantic information class methods [22]. Video-based semantic information class methods include LSTM neural network algorithms, spatio-temporal dual-stream network algorithms, TSN structure algorithms, Slowfast algorithms, etc. These algorithms are time-consuming to train and have poor real-time performance, and are mainly used in behavior detection problems where the target has fewer changes in movements that are greatly affected by time. Based on the image semantic information class methods include SSD, Faster-RNN, YOLO, etc., which are mainly used in the detection of multi-targets, small behavioral changes, and dense scene detection problems. Most of the student classroom behavior detection scenarios studied in this paper are large number of people and dense scenarios, therefore, image-based semantic information class methods are used.

3.2. Detection method flow

In order to obtain better detection results, this paper adopts YOLOv4 detection algorithm [23] as the basic algorithm, which is mainly used for students' classroom behavior detection research. Firstly, the dataset is constructed, and the detection model is trained to generate weights, the input image is feature extracted using cross-stage local network, the sense field is increased by feature pyramid network, then the path aggregation network is used for bidirectional feature extraction, fusing features of different scales, and finally the classification and regression are carried out, and the output is classified and predicted box, so as to realize the detection of student's classroom behavior.

Based on the above analysis, the specific process of detecting student behavior in the classroom is shown in **Figure 4**.

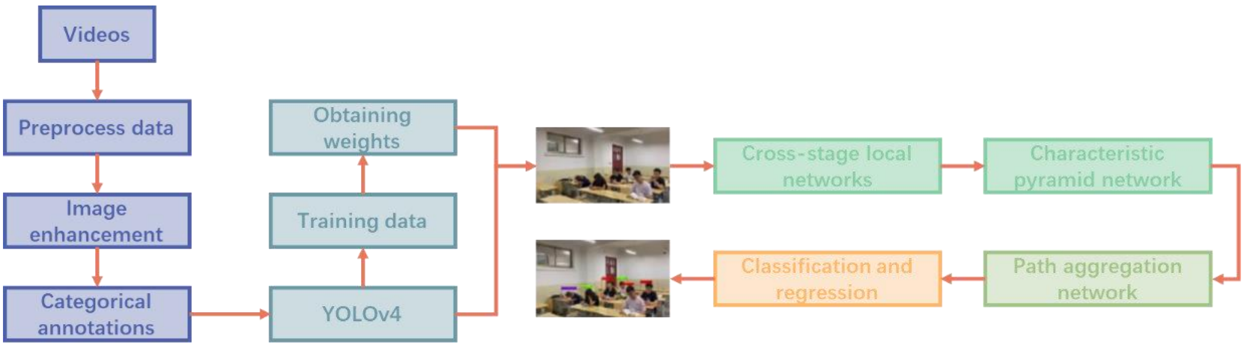


Figure 4 Flowchart of students' classroom behavior detection

4. Improved Algorithm for Detecting Student Behavior in the Classroom

In this paper, we consider datasets under different conditions, and the original YOLOv4 detection algorithm is difficult to achieve the desired effect. For this reason, this paper adopts the jumping bidirectional cross-scale feature fusion strategy, attention mechanism to improve the YOLOv4 detection algorithm.

4.1. Path Aggregation Network

Path aggregation network (PAN) [24] is based on the FPN network to add the bottom-up path, which can further get the accurate positioning in the lower layer to strengthen the information, and enhance the feature extraction ability of the network, and at the same time, adaptive feature pooling is proposed to aggregate all the candidate features, and enhance the prediction ability. The specific network structure of the path aggregation network is shown in **Figure 5**.

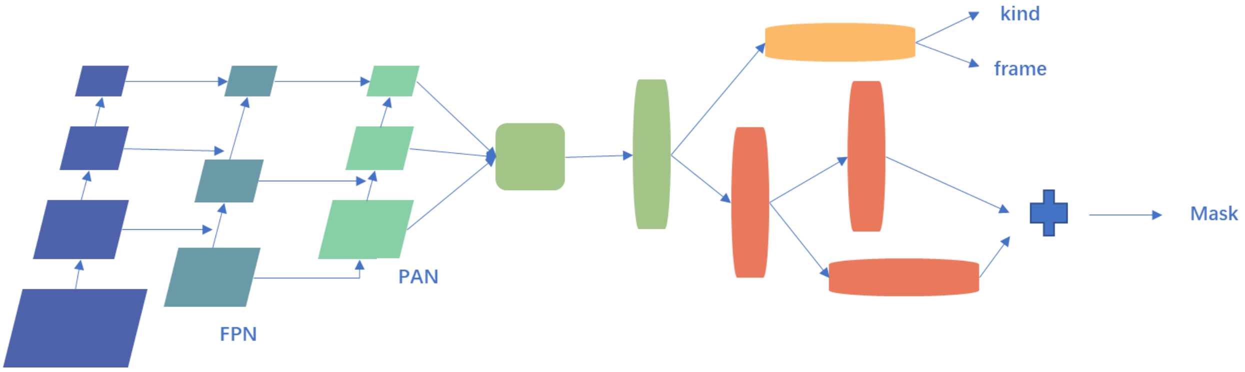


Figure 5 Path aggregation network structure

4.2. Jump bi-directional cross-scale feature fusion strategy

YOLOv4 adopts PANet as the network to strengthen the feature extraction part, which consists of FPN and PAN.YOLOv4 uses concat operation in PAN network and shortcut operation in PANet network. PANet fails to completely solve the problem of difficult detection caused by dim light and dense number of people with occlusion. In order to accelerate the network detection speed, this paper adopts the jump bidirectional cross-scale feature fusion strategy to improve YOLOv4 [25], the specific structure is shown in **Figure 6**.

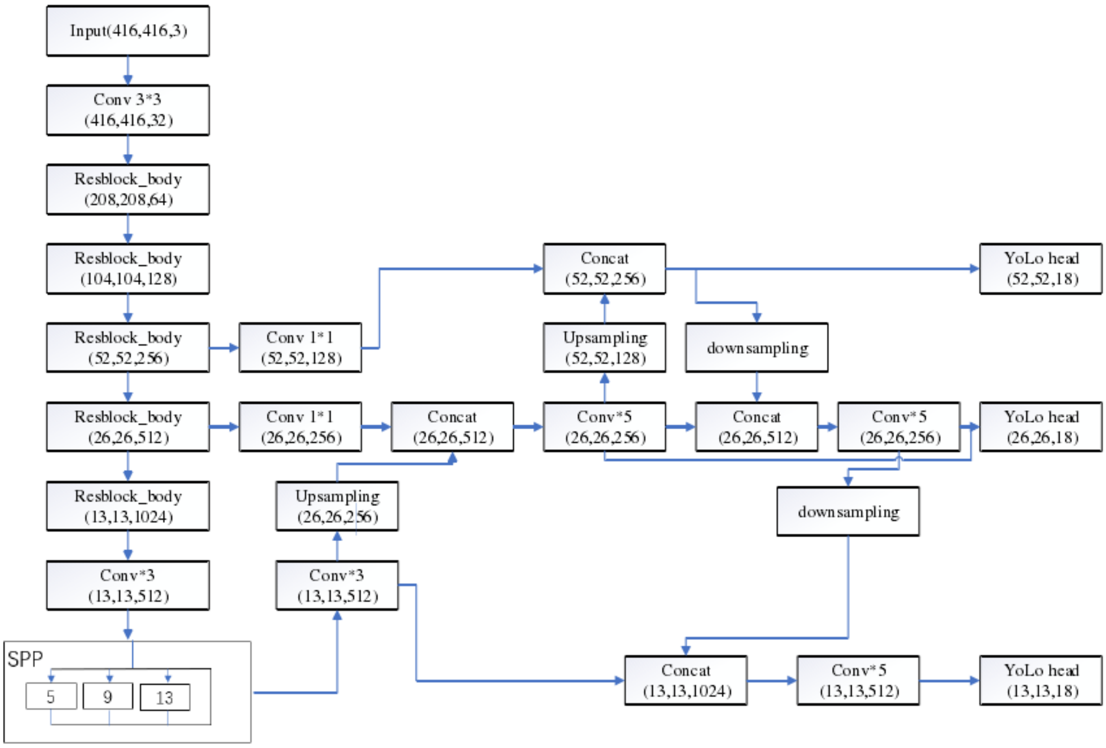
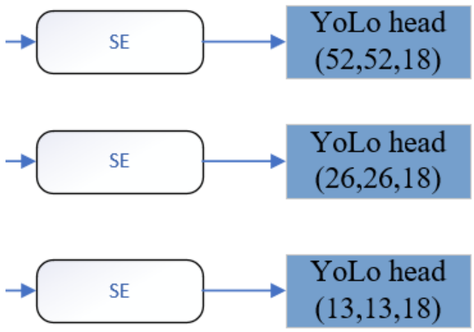


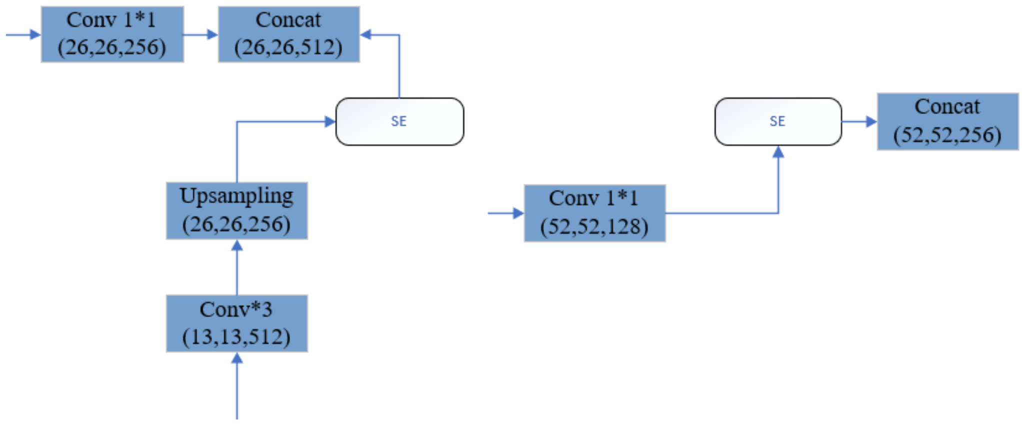
Figure 6 Jump bi-directional cross-scale feature fusion strategy

4.3. Attention Mechanism Strategy

In student behavior detection, students sitting in the back row belong to small targets in detection, which are easily confused with the background, and easily cause false detection and missed detection in the network [26]. In order to overcome the above problems, this paper adopts the attention mechanism module to improve the YOLOv4 detection algorithm. First, the SE module is added to emphasize the correlation between channels, which makes the network automatically learn the importance features between different channels, and the specific design method is shown in **Figure 7**. As can be seen from Figure 7, this paper adds the SE module from the YOLOhead former and enhanced feature extraction networks, and the networks are named YOLOv4-se-1 and YOLOv4-se-2, respectively.



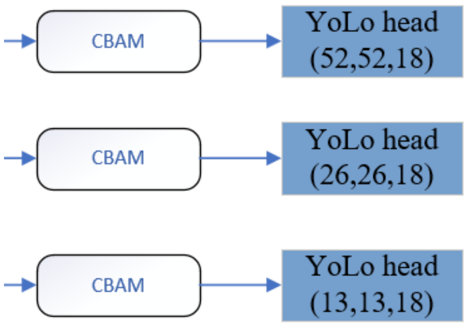
(a) YOLOv4-se-1



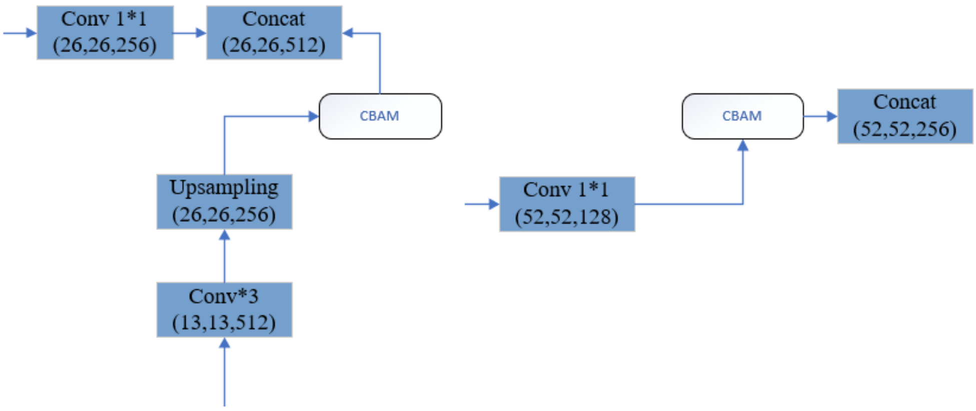
(b) YOLOv4-se-2

Figure 7 SE module embedding structure

Based on the previous experience of verifying that the CBAM module can lead to improved network performance, this paper adopts the CBAM module to improve YOLOv4, embedded in the same position as the SE module, to obtain YOLOv4-cb-1 and YOLOv4-cb-2, and the specific structure is shown in **Figure 8**.



(a) YOLOv4-cb-1



(b) YOLOv4-cb-2

Figure 8 CBAM module embedding structure

5. Experiments and analysis of results

5.1. Experimental setup

The computational node settings for this experiment are shown in **Figure 9**. Meanwhile, the ratio of training set and test set is 8:2, the input image size is 416*416, the number of batches is set to 4, the maximum learning rate is 0.001, and Epoch is set to 100.

Settings
8GB Tesla P4
AMD R7 4800H 2900MHz
6G GeForce RTX 2060
Python3.7，PyTorch1.2

Figure 9 Environment configuration

5.2. Algorithm evaluation metrics

In order to analyze and compare the performance of the models, this paper evaluates the models using Precision, Recall, Average precision (AP) and mean average precision (mAP) [26].

(1) Precision

$$P_{precision} = \frac{TP}{TP+FP} \tag{1}$$

Where TP denotes a true example, i.e., an example that is judged to be a positive sample and the fact that it is also a positive sample; and FP denotes a false positive example, i.e., an example that is judged to be a positive sample but the judgment is wrong.

(2) Recall

$$P_{recall} = \frac{TP}{TP+FN} \tag{2}$$

where FN denotes a false negative example, i.e., an example that is judged to be a negative sample but is in fact a positive sample.

(3) Average accuracy AP

$$AP = \int_0^1 p(r)\,dr \tag{3}$$

Where p denotes the check accuracy, r denotes the check completeness, and AP is the average accuracy of a class.

(4) Mean average accuracy mean mAP

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{4}$$

where mAP denotes the average of all class AP values.

5.3. Analysis of experimental results

(1) Attention mechanism strategy analysis

In order to evaluate the performance of attention mechanism strategies, the methods proposed in this paper are compared. Different student behavior detection algorithms including YOLOv4, YOLOv4-se-1, YOLOv4-se-2, YOLOv4-cb-1, and YOLOv4-cb-2 are analyzed by using college students' behavior detection dataset, and the mAP evaluation indexes with a threshold value of 0.5 are statistically evaluated, and the specific results are shown in **Figure 10**.

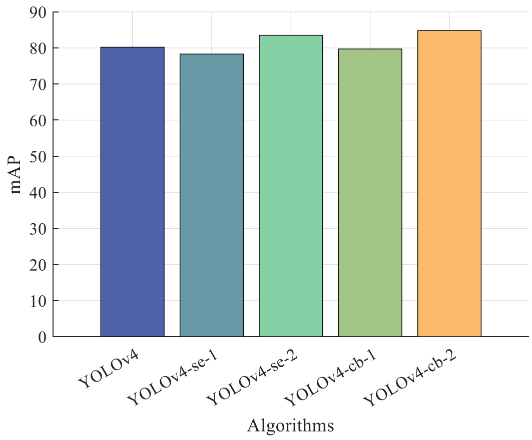


Figure 10 Comparison of performance evaluation results of student classroom behavior detection models with different attention mechanism strategies

As can be seen from **Figure 10**, the original YOLOv4 detection algorithm achieves a mAP of 80.2; YOLOv4-se-1 and YOLOv4-cb-1 embed the attention module into the position before the YoLo head, with mAP values of 78.3 and 79.7, respectively, which instead decreases when compared with the original YOLOv4 detection algorithm; YOLOv4-se-2 and YOLOv4-cb-2 embed the attention module into the enhanced feature extraction network, and their mAP values are 83.5 and 84.8, respectively, which improve the mAP values compared with the original YOLOv4 detection algorithm; comparing the mAP values of YOLOv4-se-2 and YOLOv4-cb-2, the CBAM module has a better enhancement effect than the SE module. Therefore, in this paper, we choose the attention mechanism based on CBAM module and it is embedded into the enhanced feature extraction network.

(2) Analysis of the results of the classroom behavior detection algorithm

In order to verify the effectiveness and superiority of the student classroom behavior detection method based on the YOLOv4-cb-2 algorithm, YOLOv4-cb-2 is compared with YOLOv4, YOLOv3, SSD, and Faster R-CNN models, and the evaluation results of each model are shown in Figure 11, Figure 12, and Figure 13.

Figure 11 gives the PR curve based on the YOLOv4-cb-2 algorithm under the students' classroom behavior detection dataset.PR curve is a kind of curve graph based on the precision rate and recall rate, where P is used to denote the precision and R is used to denote the recall rate.In the graph, the vertical axis is the precision, and the horizontal axis is the recall rate, and each PR curve should correspond to Each PR curve should correspond to a threshold value, and in this paper, the threshold value is selected as 0.5. It can be seen from **Figure 11** that the recall rate decreases with the increase of the check accuracy rate.

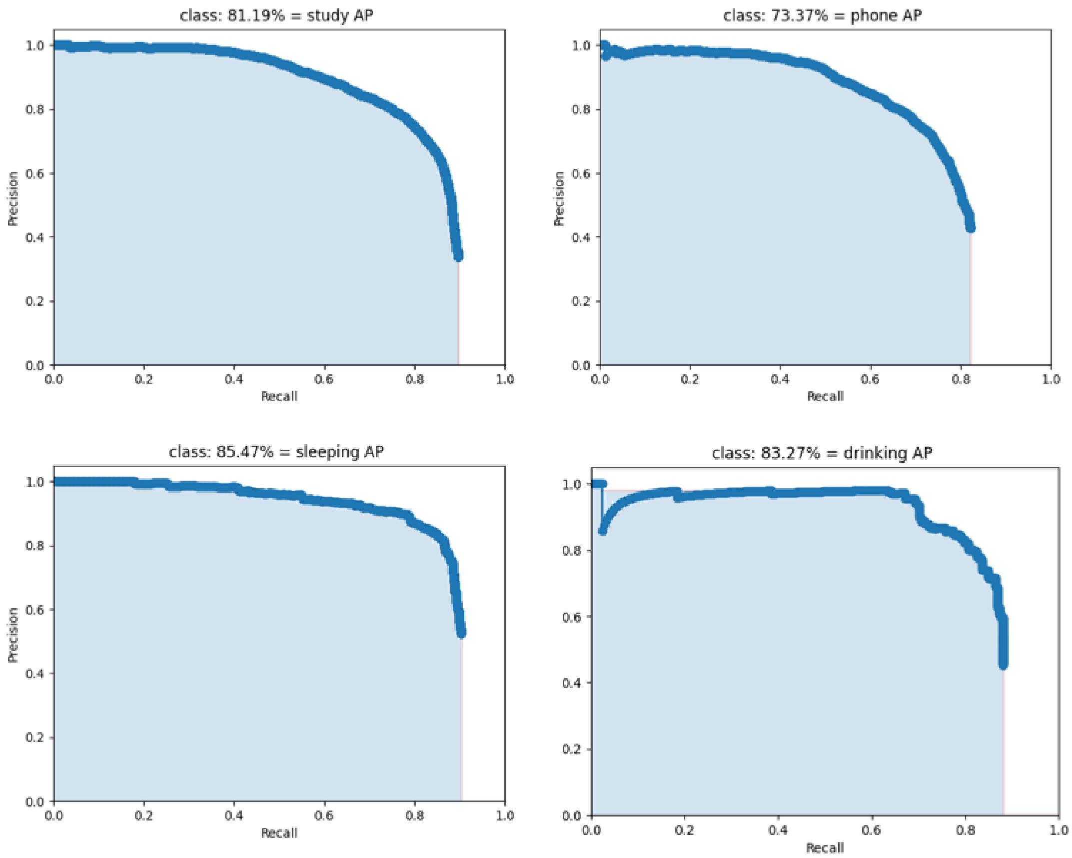


Figure 11 PR curves under the student classroom behavior detection dataset

Figure 12 gives a comparison of AP results based on different detection algorithms under the student classroom behavior detection dataset. From Figure 12, it can be seen that the student classroom behavior detection method based on YOLOv4-cb-2 algorithm demonstrates the best detection accuracy, followed by YOLOv4, Faster R-CNN, YOLOv3, and SSD, respectively.Meanwhile, the results demonstrate the feasibility of the improved detection algorithm of this paper for student classroom behavior detection.

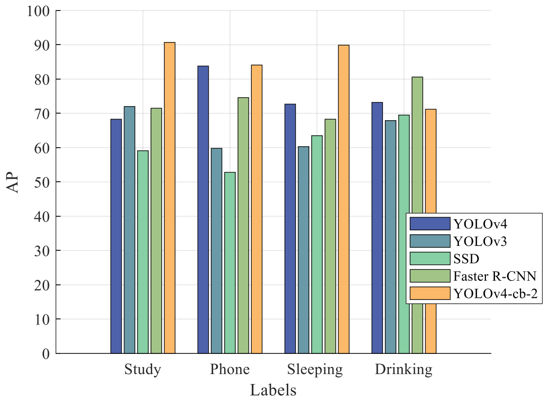


Figure 12 Comparison of AP results based on different detection algorithms under student classroom behavior detection dataset

The detection results based on different detection algorithms for four working conditions such as sparsely populated student classroom in a small class, brightly lit and densely populated student classroom in a medium class, dimly lit and heavily shaded student classroom in a large class, and student classroom in a rear-side shot are given in **Figure 13**, **Figure 14**, **Figure 15**, and **Figure 16.**

As can be seen from Figure 13, under the classroom behavior detection condition of sparse number of students in a small class, the classroom behavior detection method based on the YOLOv4-cb-2 algorithm has improved the precision of the rest of the students' behaviors, except for the first row of the cell phone playing girls whose detection precision has decreased, and the detection precision of the YOLOv4-cb-2 algorithm has a small difference between the YOLOv4 and the YOLOv4 detection precision.

(a) Original map

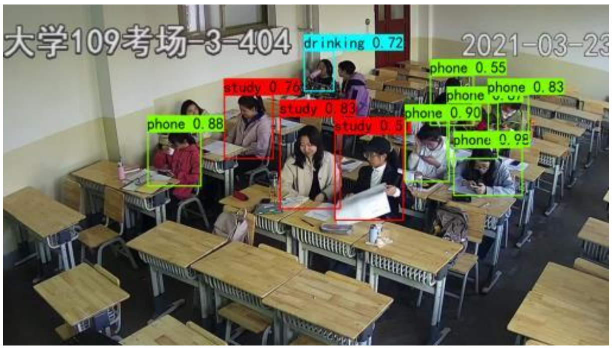
(b) YOLOv4

(c) YOLOv4-cb-2

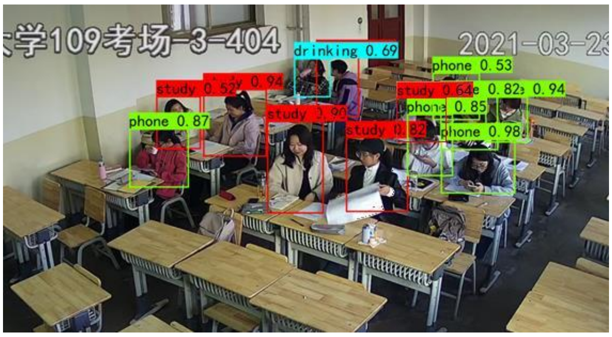
Figure 13 Comparison of the effectiveness of classroom behavior detection for students with small sparse class sizes

As can be seen from Figure 14, under the brightly lit and number-intensive student classroom behavior detection condition in the middle class, YOLOv4-cb-2 detection is more comprehensive and the detection accuracy is improved compared to the YOLOv4 detection algorithm.

(a) Original map



(b) YOLOv4



(c) YOLOv4-cb-2

Figure 14 Comparison of the effect of classroom behavior detection for brightly lit and densely populated students in the middle classroom

As can be seen in **Figure 15**, under the classroom behavior detection condition of a large class of dimly lit and heavily occluded type students, the YOLOv4-cb-2 algorithm detects significantly more than the YOLOv4 detection algorithm, and the detection accuracy is improved, with virtually no false detections.

(a) Original map

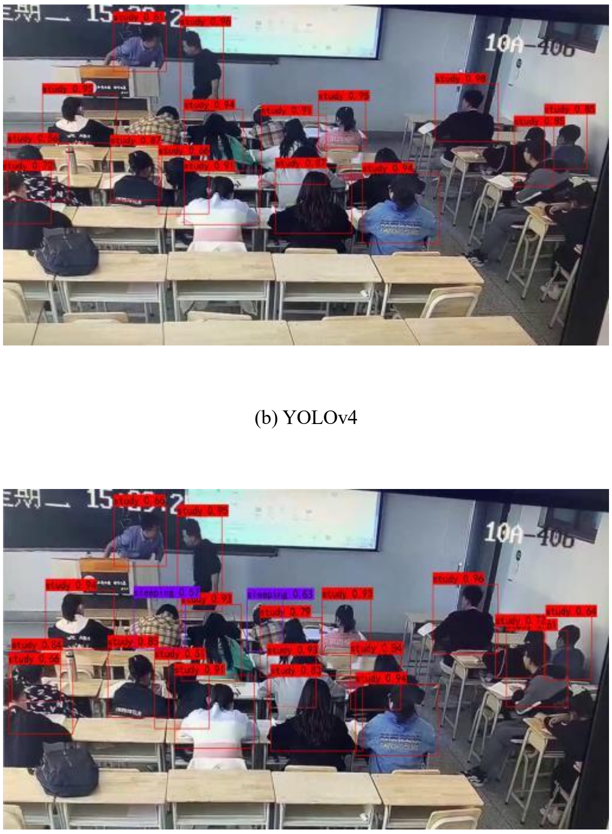
(b) YOLOv4

(c) YOLOv4-cb-2

Figure 15 Comparison of the effect of classroom behavior detection for students with dim lighting and severe shading type in large classes

As can be seen in **Figure 16**, the YOLOv4 detection algorithm fails to detect students sleeping on their stomachs in the front row under the classroom behavioral detection effect condition of the rear side shot students, and the improved YOLOv4-cb-2 algorithm detects all students and has a higher behavioral detection accuracy.

(a) Original map



(b) YOLOv4

(c) YOLOv4-cb-2

Figure 16 Comparison of the effect of back-side shots on detecting student behavior in the classroom

6. Conclusion

Aiming at the problem that the performance of existing student classroom behavior detection models is affected by various aspects such as datasets, algorithms and heights, as well as the differences between different classrooms, this paper proposes a method for student classroom behavior detection that improves the YOLOv4 algorithm. The method constructs student behavior detection data, designs a behavior detection scheme based on the characteristics of student classroom behavior, and improves the YOLOv4 algorithm by combining the jump bidirectional cross-scale feature fusion strategy with the attention mechanism strategy. Comparative experimental analysis of the proposed method using the collected classroom behavior detection datasets of college students under different working conditions shows that the improved YOLOv4-cb-2 algorithm has the highest detection accuracy.

The student classroom behavior detection algorithm implements some of the functions of the student classroom behavior detection management system, and has not been considered to incorporate face recognition detection algorithms in the detection system, which will be the next step.

References

[1] Gong D , Wang T , Su X , ZhangY .Equivalent Mutants Detection Based on Weighted Software Behavior Graph[J].International journal of software engineering and knowledge engineering, 2022.

[2] Lina S U , Fan L I , Xiu P , Zili T, Jiandong Z. Detection of vehicle lane-changing behavior based on radar data[J]. Journal of Beijing Jiaotong University , 2022, 46(4):60-67.

[3] Wang C , Zhu H .Representing Fine-Grained Co-Occurrences for Behavior-Based Fraud Detection in Online Payment Services[J].IEEE transactions on dependable and secure computing, 2022(1):19.

[4] Sun C , Guo X , Zhao T , Han Y. Real-time detection method of window opening behavior using deep learning-based image recognition in severe cold regions[J].Energy and buildings, 2022.

[5] Guo M , Zhao X , Yao Y , Bi C, Su Y. Application of risky driving behavior in crash detection and analysis[J].Physica A: Statistical Mechanics and its Applications, 2022, 591:126808-.

[6] Yang D , Zhang Y , He H .AI-Based Detection of Power Consumption Behavior of People in a Smart City[J].Journal of Testing and Evaluation: a Multidisciplinary Forum for Applied Sciences and Engineering, 2023.

[7] Wang H , Zhang S , Zhao S , Wang Q, Li D, Zhao R. Real-time detection and tracking of fish abnormal behavior based on improved YOLOV5 and SiamRPN++[J]. Computers and Electronics in Agriculture, 2022.

[8] Zhang G , Wang L , Chen Z .Hand-raising gesture detection in classroom with spatial context augmentation and dilated convolution[J].Computers & graphics, 2023.

[9] Ding Y , Bao K , Zhang J .An Intelligent System for Detecting Abnormal Behavior in Students Based on the Human Skeleton and Deep Learning[J]. Computational intelligence and neuroscience, 2022.

[10] Chen C , Arakawa Y , Watanabe K , Ishimaru S. Quantitative Evaluation System for Online Meetings Based on Multimodal Microbehavior Analysis[J]. Sensors and materials: an International Journal on Sensor Technology, 2022(8 Pt.1):34.

[11] Zhang Y , Tingbing X U , Wei Z .Pre-locate net for object detection in high-resolution images[J].Chinese Journal of Aeronautics, 2022(10):313-325.

[12] Annas G D , Freitas C .Warning behaviors and leaked intent: potential new avenues to prevent mass shootings[J]. 67(1):275-278.

[13] Mo J , Zhu R , Yuan H , Shou Z, Chen L. Student behavior recognition based on multitask learning[J].Multimedia tools and applications, 2023.

[14] Peltier W , Newell K L , Linton E , Holmes S C, Donaldson J M. Effects of and preference for student- and teacher-implemented good behavior game in early elementary classes[J].Journal of applied behavior analysis, 2023.

[15] Chen L , Wang L , Zhou Y .Research on Data Mining Combination Model Analysis and Performance Prediction Based on Students' Behavior Characteristics [J].Mathematical Problems in Engineering, 2022, 2022.

[16] Rosario G , Akila V , Govindasamy V .Detection of automated behavior on Twitter through approximate entropy and sample entropy[J]. ubiquitous computing, 2023.

[17] Zhao L , Zhang Y , Yu J , Yin G. Islandization of the University Town from the Perspective of Student Off-Campus Behavior: A Case Study of Changqing University Town, Jinan City, China[J].Journal of Urban Planning and

Development, 2023.
[18] Zhou J , Ran F , Li G , Peng J, Li K, Wang Z. Classroom Learning Status Assessment Based on Deep Learning[J].Mathematical Problems in Engineering, 2022,. 2022.
[19] Frick B , Lensing F , Beck-Werz L .When an exchange semester is no longer enough - Why and how the Bologna-reforms changed the behavior of high-ability students[J].Working Papers Dissertations, 2022.
[20] Obaidat I , Sridhar M , Pham K M ,Phung P H. Jadeite: a novel image-behavior-base approach for Java malware detection using deep learning[J]. Computers & Security, 2022(113-):113.
[21] Zheng Zhaoqian, Han Dongchen, Zhao Hui. Single-step fragment labeling for joint entity-relationship extraction model[J]. Computer Engineering and Applications, 2023, 59(9):10.
[22] Tan B,Yang Shujun. Research on student classroom behavior detection algorithm based on FasterR-CNN[J]. Modern Computer, 2018(22):3.
[23] Zhang X , Wang G .Stud Pose Detection Based on Photometric Stereo and Lightweight YOLOv4[J]. Journal of Artificial Intelligence Technology (English), 2022(001):002.
[24] Dang L , Huangfu P , Hou Y E , LiuY, Han H. A Path Aggregation Network Based on Residual Feature Enhancement for Object Detection in Remote Sensing Imagery[J].Remote sensing letters, 2023(4/6):14.
[25] ZHANG Qian,LIU Ziyan,CHEN Yunlei,et al. Fusion of Transformer and improved PANet for YOLOv5s traffic sign detection[J]. Journal of Sensing Technology, 2023, 36(2):232-241.
[26] Xu H .Prediction of Students' Performance Based on the Hybrid IDA-SVR Model[J].Complexity, 2022, 2022.
[27] Deng G , Lin L , Lin B , Jing W, Zhao X, Li Y. Design of Command and Dispatch System for Automatic Reading of Meter Images in Substations by UAV Inspection Photos[J].Computing Open, 2024, 02.