ISSN: 2664-9640

# Classroom Management System for College Students Based on Face Recognition and Object Detection

Shaoqi Yan [a], Yaoxin Yi

Beijing Wuzi University, Beijing 101100, China

[a]1622400870@qq.com

## Abstract

**Based on the research on face recognition and object detection, a set of systems suitable for classroom management is proposed, which includes two parts of functions, namely the function of recording student attendance information by scanning their faces before class and the function of catching abnormal behavior of students through object detection recognition algorithm in class. Based on an open source project named face recognition, built a highly efficient and realistic face check-in subsystem. By combining the one-stage object detection algorithm of Yolov5s with the Coordinate Attention mechanism which is more suitable for lightweight networks, the abnormal behavior detection function of the system is realized. After deployment and testing, it is found that the performance of the system can better meet the classroom management's needs of the university, and greatly improve the school's efficiency and level of classroom management.**

## Keywords

**Classroom management, face recognition, abnormal behavior detection, Coordinate Attention, Yolov5 algorithm.**

## 1. Introduction

In recent years, the rapid development of the field of computer vision makes its application in various fields close to mature and the classroom situation of college students has always been a key issue of the school. Attendance methods and students' performance in the classroom need to be improved, the original attendance methods such as paper check-in, designated App sign-in, teacher roll call sign-in. There are bad phenomena such as student signing, leaving the classroom after signing, resulting in the falseness of attendance results, and will indirectly hit the attendance enthusiasm of some students. However, in the classrooms of college students in some colleges and universities, there are generally non-concentration phenomena such as students sleeping, talking, and playing mobile phones in class, which greatly affects the learning efficiency of these students. In view of the above two problems, based on the two technologies of face recognition and object detection, a set of classroom management system has been developed, which can timely and accurately understand the attendance status and classroom performance of students, thereby improving the efficiency of school management students, and also improving students' enthusiasm and consciousness in learning. The face recognition part of this system is mainly used to obtain students' attendance, and can also be applied to other similar fields such as the check-in attendance of enterprise employees, and has strong promotion and generalization capabilities. This article mainly introduces the classroom management system and implementation method based on face recognition and object detection, which includes two parts of the function, namely the function of recording student attendance information by brushing face check-in before class and the function of catching students' abnormal behavior through object detection recognition algorithm in class. The

ISSN: 2664-9640

system can automatically capture the student's on-site face information and real-time comparison and verification with the information in the image database, so it can ensure the accuracy of the student attendance information, solve the drawbacks of the traditional check-in method, such as the phenomenon of signing on behalf of the traditional check-in method, and the system can detect the abnormal behavior of students in real time through the target recognition algorithm in the classroom, which plays a positive role in urging students to study seriously and increase the quality of the classroom.

## 2. Classroom management system structure design

The system structure mainly includes 4 parts, namely face entry, brush face check-in, generation of punch card results, and target detection and recognition. The role of the first three parts is to record students' check-in before class, and the last part is used to detect and identify abnormal behavior of students in the classroom. The system structure flow is shown in Figure 1

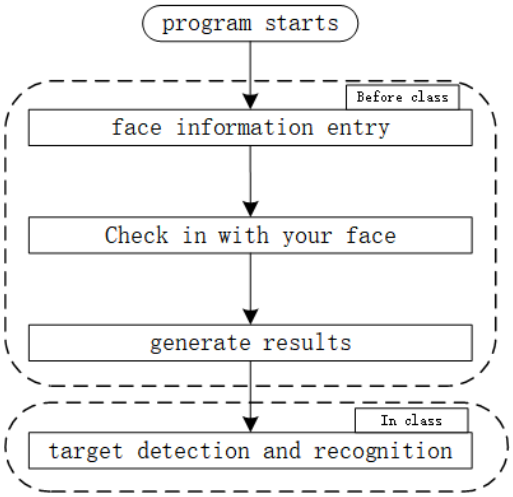


Figure 1 Classroom management system structure diagram

### 2.1. Face entry

Entering a student's face information into the database is a prerequisite for face recognition check-in. The school can organize students to enter faces in an orderly manner during the free time period, students need to carry their ID cards to the designated place to swipe the card, after the card swipe, the instrument will prompt the students to stand in the designated location, and then call the camera through the Open Cv library to collect the student's face. The specified location should be selected in a place with moderate light to ensure that the image acquisition process works well. The face library will automatically retrieve whether the students have repeatedly entered to ensure the uniqueness of the data, and after the collection is completed, the interactive interface will automatically prompt the input to complete.

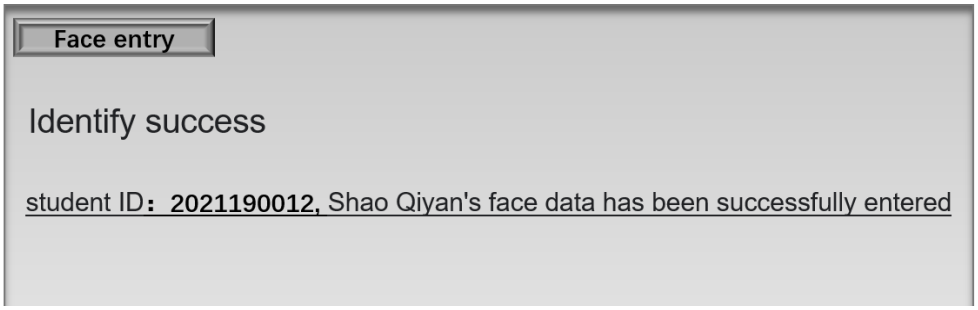


Figure 2 Prompt for completion of acquisition

ISSN: 2664-9640

## 2.2. Brush your face to check in

Brush face check-in is the most important link in the whole process of recording attendance information, by calling the camera to collect the student's face code information in real time, and then measure the distance with the face code information in the database to determine its similarity. Commonly used distance measurement methods include Euclidean distance, cosine similarity, Hamming distance, Manhattan distance, Chebyshev distance, etc., which have their own advantages and disadvantages and are suitable for different scenarios. Since cosine similarity is better at measuring high-dimensional data, and the module used by the face recognition program in this article will obtain 128 dimensions of data for the face, this paper uses cosine similarity to measure the similarity of the compared images. The core of the cosine similarity algorithm is: the image feature group formed by n elements of the image matrix can constitute an n-dimensional space, and the two image matrices that are compared can be regarded as a point in the n-dimensional space[1], so that x and y represent these two points, and the similarity can be calculated by the cosine similarity formula:

$$x = (x_1 + x_2 + x_3 + \cdots + x_n) \tag{1}$$

$$y = (y_1 + y_2 + y_3 + \cdots + y_n) \tag{2}$$

$$D(x,y) = \cos(\theta) = \frac{x \cdot y}{\|x\| \cdot \|y\|} \tag{3}$$

The system can recognize and record multiple faces at the same time, and the density of crowds during the peak period of class will be relatively large, and the camera is likely to exist in the case of capturing multiple faces at the same time. One-to-many face recognition improves the efficiency of student check-in and avoids crowded queues during peak periods, preventing students from entering the classroom in time. If the student successfully checks in, the interactive interface displays the information of the successful check-in, as shown in Figure 3.

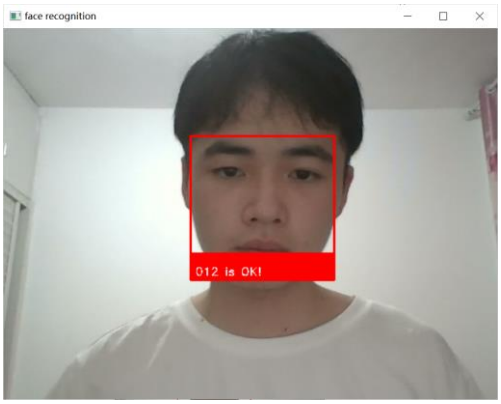


Figure 3 Successful display of check-in

## 2.3. Generate punch card results

Users can artificially modify the punch time according to the specific class time of different courses, which has greater flexibility. This article shows that the punch card critical point is set for the course start time at 8:00 a.m. Students who check in with their faces before 8 o'clock will be recorded as unsuccessful; If the student has already checked in but the check-in time has passed 8 o'clock, it is recorded as late; If the student does not clock in on that day, the attendance sheet will not generate any information, which means that the student is absent from class on that day. After the student checks in with the face, the punch card result will be generated at the same time and will be recorded in the attendance record sheet. Classroom teachers can check the attendance table before the start of the class to get the overall attendance, and later they can also easily query the attendance record sheet to warn students with high absenteeism rates or add or subtract from students' final assessments accordingly.

ISSN: 2664-9640

The attendance record sheet mainly includes name, student number, punch date, attendance status, and check-in time accurate to the second, which clearly and comprehensively reflects the daily check-in results of students in the classroom.

| Name | student ID | 2022-1-26 | Time | 2022-1-27 | Time | 2022-08-04 |
|---|---|---|---|---|---|---|
| Shao Qiyan | 2021190012 | Be Late | 2022-1-26 7:31 | OK | 2022-1-27 8:14 | Be Late |

Figure 4 Records of late arrivals and successful check-ins

## 2.4. Object detection and identification

This step is mainly used to detect abnormal behavior of students in the classroom, so as to supervise students to study seriously. Considering the difficulty of obtaining the dataset and the most common unattended behavior of students in the classroom, this paper divides the abnormal behavior label into three categories: sleeping, talking, and playing with the mobile phone, and the system captures and detects the students' pictures in the classroom in real time by calling the camera. When the system detects the above three behaviors of students, the students who are detected for abnormal behavior are automatically marked with a box line and a category and confidence level are attached to the upper right corner. This step does not record students who have abnormal behavior, college classes are different from middle and high schools, college students have a certain degree of freedom in the classroom, excessive supervision will be counterproductive, and there is a certain misjudgment of the target detection and recognition algorithm. The classroom teacher or the relevant management personnel can perform irregular sampling of the classroom video, which is mainly used to detect the abnormal behavior of students in the classroom, so as to urge students to study seriously. Check, and warn students who often have abnormal behavior in the classroom.

# 3. Face recognition implementation process

This article uses the face recognition open source project to implement the face recognition function of the system, face recognition module is equipped with complete development documents and application cases, and provides a wealth of API interfaces, users only need to simply learn to achieve various functions of face recognition. Face recognition powerful, simple and easy to use features greatly simplify the development of face recognition systems.

## 3.1. Face detection

Before you can perform face recognition, you need to find all the face positions in the video frame. When used, first use the read method in the Opencv module to read the video frame, and when calling the face locator face_recognition. face_location() recognizes all the faces in the video frame and returns its position information.

## 3.2. Feature point detection

Feature points are important feature points of various parts of the face, usually contour points and corner points, which can reflect the facial features of each part, and the Dlib official feature point extraction model "shape_predictor_68_face_landmarks.dat" can achieve face 68 feature points. Through the linear transformation of image space, the various feature points can be aligned to achieve the effect of improving the accuracy and speed of face recognition.

## 3.3. Facial coding and comparison recognition

After identifying the faces in the video frame and their location, call the face_recognition. face_encodings() method to predict the 128-dimensional vector value of each face. The 128-dimensional vector value extracted from the video frame is compared with the 1 28-

ISSN: 2664-9640

dimensional vector value pre-saved in the face library, and if the similarity is greater than a certain threshold, the identity of the detected face is determined to be the identity of the image in the database.

# 4. Implementation of anomalous behavior detection

## 4.1. Deep learning framework Pytorch

The wide application of deep learning benefits from many excellent open source deep learning frameworks[2]. Computer Simulation,2005,22(5):109-110. Widely used deep learning frameworks include Keras, Caffe, Torch, and Tensorflow. Pytorch is a python version of torch, an open source neural network framework by Facebook, and Pytorch is more user-friendly to Python language speakers. PyTorch's design follows the tensor→variable (autograd) →nn. Module has three low to high levels of abstraction, the source code is only about one-tenth of TensorFlow, and the concise design makes the Pytorch source code easier to reference and learn. And under the same algorithm program, PyTorch is more likely to outperform frameworks such as TensorFlow and Keras, and Pytorch is also very supportive of GPU The training mode greatly improves the efficiency of the training model.

## 4.2. Improved yolov5 algorithm and training

The yolo series of algorithms is one of the most effective target recognition algorithms, and many ingenious convolutional neural network ideas are included in the entire computer object recognition field literature. Compared with other two-stage algorithms such as Faster R-CNN and R-CNN, yolov5, as a one-stage algorithm, is faster in recognition speed and better in target recognition problems that need to be detected in real time[3].

(1) Yolov5 network analysis

Yolov5s network is the network with the smallest depth and the smallest width of the feature map in the Yolov5 series, Yolov5m, Yolov5l

Yolov5x is constantly deepening on this basis. This article uses Yolov5s, Yolov5s is mainly composed of four parts: Input input, Backbone backbone network, Neck and Head output[4]. Perform Mosaic data enhancement operations on the I nput side. Backbone contains the Focus structure and spatial pyramid pooling SPP, and the Neck network layer uses the FPN and PAN structures. Both Backbone and Neck contain a large number of Crossbowcock CSP structures to help extract different levels of features in the image, which plays a role in reducing the amount of computation and speeding up training. This article uses yolov5s' own Glou_loss as a loss function.

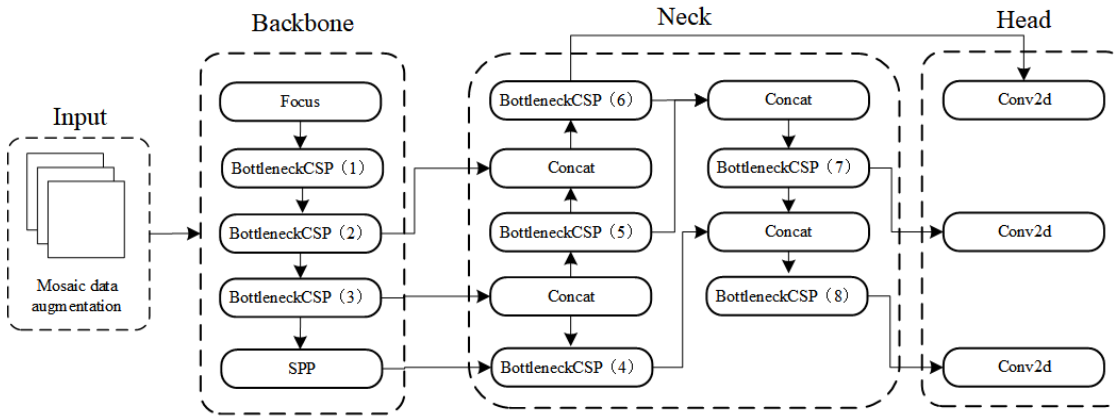


Figure 5 Yolov5s network structure diagram

(2) Yolov5s fusionC-A attention mechanism

ISSN: 2664-9640

In recent years, the development of target detection algorithms has been very rapid, and the position of one-stage algorithms in object detection has become more and more dominant. Attention mechanism in many object detection, classification of the application of the application scenarios appear, the current common attention mechanism is SE, CBMA, ECA, CA, etc. The previous attention mechanism such as SE, CBMA on the channel processing generally use two-dimensional global pooling, resulting in location information loss[5], and Yolov5s this lightweight object detection algorithm is difficult to bear the huge computing overhead of these two. The CA embeds the position information into the channel attention, it decomposes the attention into two 1-dimensional feature coding processes, and aggregates the features along two spatial directions at the same time, and encodes the generated feature map into a pair of direction-aware and position-sensitive feature maps, which can be applied complementarily to the input feature maps to enhance the representation of the objects of interest. Compared with the previous attention mechanism, CA can not only retain position information, but its extremely small computational overhead will also make the attention mechanism work better in the lightweight object detection algorithm of Y olov5.

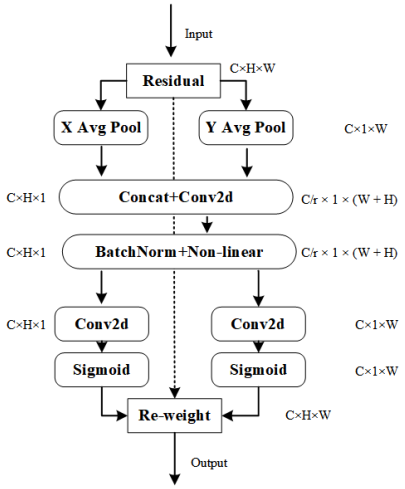


Figure 6 CA structure diagram

Theoretically, the attention mechanism can be added to any feature map, but Backbone is the most critical place to extract features, and the way the attention mechanism randomly initializes the weights will cause damage to the pre-trained weights of the backbone network, so this paper adds CA to the feature layer extracted from the trunk network, The experimental results show that after introducing the CA, the accuracy of the algorithm is improved by 5.34%,and the specific structure is shown in Figure 7.

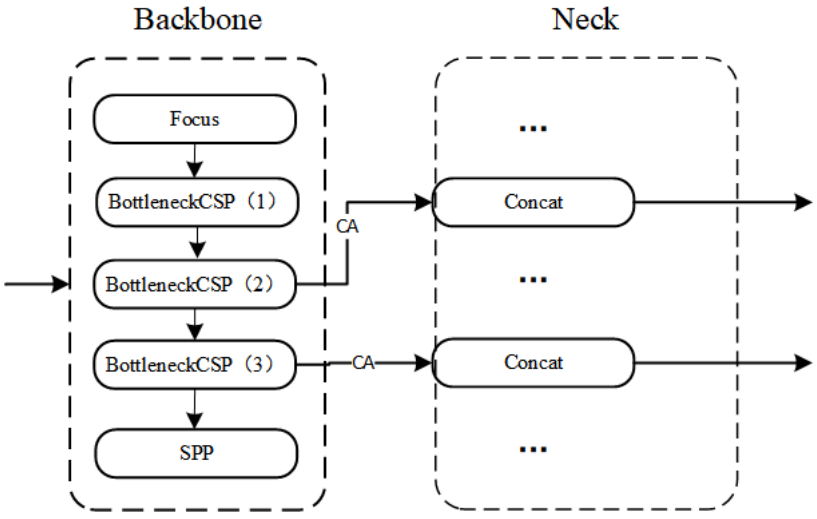


Figure 7 Integrating CA between Backbone and Neck

ISSN: 2664-9640

## 5. Conclusion

Based on face recognition and object detection technology, this paper proposes a set of classroom management system for student attendance and abnormal behavior detection, and after development, deployment and testing, it is found that its performance can better meet the needs of classroom management. By using the system, it can solve the drawbacks of traditional attendance methods and improve the quality of students' class time, which greatly improves the efficiency and level of student management, and the system can be generalized to scenarios with high similarity with this article, with good generalization ability.

## Acknowledgements

Supported by: Beijing Wuzi University Students' Innovation and Entrepreneurship Program Training Program (2022103026).

## References

[1] SUN Yue, YANG Guowei. Research on student attendance system based on face recognition [J/OL]. Modern Electronics Technique, 2020, 43(10): 116-118+123.
[2] LI Xiong, WEN Kaifu. Developing a Face Recognition Attendance Management System Based on Deep Learning [J]. Research And Exploration in Laboratory, 2019, 38(7): 115-118+123.
[3] XU Degang, WANG Lu, LI Fan. Review of Typical Object Detection Algorithms for Deep Learning [J]. Computer Engineering and Applications, 2021, 57(8): 10-25.
[4] LI Yongshang, MA Ronggui, ZHANG Meiyue. Traffic Monitoring Video Vehicle Volume Statistics Method Based on Improved YOLOv5s+DeepSORT [J]. Computer Engineering and Applications, 2022, 58(5): 271-279.
[5] WANG Lingmin, DUAN Jun, XIN Liwei. YOLOv5 Helmet Wear Detection Method with Introduction of Attention Mechanism [J]. Computer Engineering and Applications, 2022, 58(9): 303-312.