# Classroom Attendance Auto-management Based on Deep Learning

## Dan Wang[1,a], Rong Fu[1,b], Zuying Luo[2,c,*]

[1]College of Information Science and Technology, Beijing Normal University, Beijing, China

2 Research Center of Virtual Reality and Visualization Technology and Engineering, Beijing Normal University, Beijing 100875, China

[a] 201521210005@mail.bnu.edu.cn, [b] 201621210003@mail.bnu.edu.cn [c] luozy@bnu.edu.cn

*corresponding author

**Keywords:** classroom evaluations, attendance, face detection, face recognition.

**Abstract:** Attendance is an important part of daily classroom evaluation. This paper develops an automatic attendance system by integrating two deep learning algorithm Faster R-CNN face detection algorithm and SeetaFace face recognition algorithm. The results of numerous experiments indict that: (1) the system can record such five violations of classroom, that is absence, later arrival, early departure, free access, and carelessness for attendance, and give the attendance table which can reflect the learning situation of all students after school. (2)  For small classrooms with length of less than 6 meters, 1080P classroom monitoring video can meet the needs of classroom attendance; but for 9 meters of big classrooms, you should use 4K classroom surveillance video.

## 1.    Introduction

As an important part of class teaching, attendance plays an important role in the evaluation of teaching. At the beginning and ending of class, it is usually checked by the teacher, but it may appear that miss someone or some students answer multiple times. Based on high-definition monitor video, how to use face recognition and other information technology for attendance is a practical research problem.

Firstly, we use Faster R-CNN [1] for face detection to determine the number of students, and then use face recognition algorithm SeetaFace to automatically identify the students to cover absence, later arrival, leave early, free access these four attendance functions. The face detection ability of SeetaFace is lower than Faster R-CNN, so we use the difference to achieve carelessness this new attendance function, that is, when Faster R-CNN can detect a student's face, but SeetaFace cannot detect that explains the student is sitting in a particular way and not listen carefully. We measure this attendance system in the classroom, and use 720P, 1080P camera, 4K HD camcorder to conduct tests; the classroom camera is almost of 720P at present. We design an experiment to verity the minimum resolution required for our analysis at first. And then we use the picture from 4K camera for recognition, the size of a single picture is $3840 \times 2160$.

The contributions of our work are: (1) with comparative experiment, we obtain that in real classroom scene, the resolution of the camera should be at least 1080P to detect the classroom-wide faces. (2)The face recognition system SeetaFace is successfully applied to classroom evaluation, which combined with Faster R-CNN face detection can achieve classroom automatic attendance.

## 2.    Related Works

Face detection is a basic technology of human-computer interaction [2]. It can get information from the faces in pictures or video. Face recognition technology analyse the face image to extract the facial feature, and then identify specific target. The development of deep learning technology further improves the accuracy of face recognition [3]. Deep Learning CNNs (Convolutional Neural Networks) have made significant breakthrough in image classification [4] [5]. The deep learning model DeepID developed by the Computer Vision Research Group, have achieved a recognition rate of 99.15% on LFW database, which is higher than that of human eye with 97.52%.

Faster R-CNN is the best method of R-CNN series for target detection, in this paper face detection training is performed on the Faster R-CNN network using the face detection benchmark data set Wider Face [6]. The data set is ten times larger than the existing data set.

SeetaFace face recognition consists of three core modules, including face detection, face feature point positioning module and face recognition module. It uses traditional shallow network for face detection and feature point positioning to improve the speed, and uses the deep network for face recognition to ensure the accuracy of face recognition in the last step.

## 3.    Automatic Attendance System Based on Faster R-CNN and SeetaFace

### 3.1.    Algorithm Introduction

Figure 1 is the attendance system based on face detection algorithm Faster R-CNN and face recognition algorithm SeetaFace. In this paper, we propose an automatic attendance system, which includes five modules (1) Capture HD video module. Record a video to ensure that every student is appeared in the video. (2) Image separate module. Separate a frame per minute for class attendance; (3) Faster R-CNN face detection module. Detect all students' faces and output coordinates. (4)SeetaFace face recognition module for face recognition. (5)Automatic attendance analysis module. The seats of students are fixed so we can contrast their faces' coordinate, and determine the identity of each student to achieve automatic attendance.
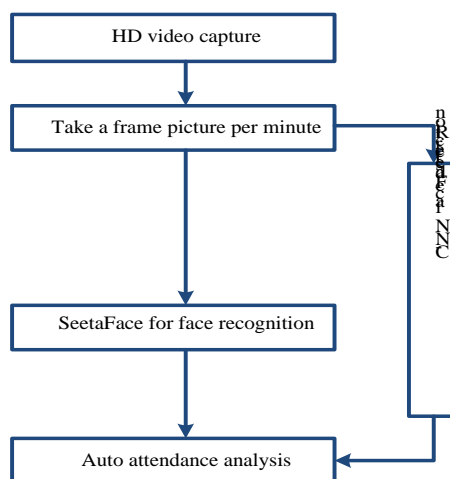


Figure 1 Algorithm of the automatic attendance system.

### 3.2.    Definition of Five Class Attendance Indicators

For automatic attendance, we define three kinds of learning states in the i-th frame image (1) if SeetaFace can recognize the face of the student, we regard the student as attendance and indicate it by ' √ ' in attendance table;(2) Faster R-CNN can't detect the face, we regard the student as absence

and indicate it 'x';(3) Faster R-CNN can detect the face of the student, but SeetaFace can't recognize the face of the student. In means the student is not in a normal sitting position. We regard him as carelessness and indicate with '●'; Based on the learning states, our system can realize automatic attendance about the five indicators: absence, later arrival, early departure, random access and carelessness. The following are definition of the indicators:

- Absence. Faster R-CNN and SeetaFace neither detect the student in a forty-five minute class, where appears forty five times 'x' in the attendance table. We regard him as absence.
- Later arrival. Faster R-CNN and SeetaFace can't detect the student in a few minutes at the beginning of the class, but can detect him later, that appearing a few times 'x' at the beginning, and always appearing'√' later in the attendance table, and we regard him as late.
- Early departure. Faster R-CNN can detect the student in ahead of class, but then can't detect him, that appears a few times '√', then always appearing 'x'. We regard him as early departure.
- Random access. We can detect the student in a few minutes at the beginning and ending of class, but can't detect him during the middle of class. That is appearing a few times '√', then appearing 'x', and appearing '√' at last. We regard him as random access.
- Carelessness. We can detect the student in a few minutes at the beginning and ending of the class, but only SeetaFace can't detect him during the middle of class. That is appearing a few times '√', and then appearing '●'. In this situation, we regard him as carelessness.

### 3.3. Face Detection for Faster R-CNN



Figure 2 Result of face detection by Faster R-CNN. Figure 3 Face identification result of SeetaFace.

As shown in Figure 2, Fast R-CNN obtains good face detection results for 720P frame images that were separated from classroom video downloaded from China's public class. Most of the students' faces were detected in the classroom. A few have not been detected are those with serious obstruction. In order to detect the faces of all students, the camera must be installed in a suitable location.

### 3.4. Results of Face Recognition by SeetaFace

SeetaFace is the open source algorithm of face recognition published by the Institute of Computing Technology of the Chinese Academy of Sciences, based on its learning parameters, we directly use SeetaFace on a 4K video from a real classroom for face recognition, the result is shown in Figure 3, there have been accurate face recognition for all students in the classroom, which illustrates that SeetaFace can meet the need of real classroom attendance face recognition.

## 4. Experiments

### 4.1. Setup

We place three different resolution ratio video in the classroom of 6mx10m. And then collect experimental data, Firstly we collect multiple face images for every student, arrange six students seated across row to ensure that even the student at the last row can be detected and recognized. And then attach their name to their face images, that is, we establish the face database for our automatic attendance system.



Figure 4 results of the 720P camera.



Figure 5 results of the 1080P camera.



Figure 6 result of the 4K camera.



Figure 7 detection Results when head left.

### 4.2. Comparative Experiments of Different Resolution Ratio Video

Figure 4 to figure 6 show that when using SeetaFace for face feature point extraction, for 720p frame picture, we can obtain the feature points from 4 meters away; 6 meters for 1080P, and for 4K frame image we can obtain the feature points 9 meters away from the face, that is, 720P camera, 1080P camera, 4K camera can obtain the feature points 4 meters, 6 meters, 9 meters away from the faces to realize recognition respectively. This indicates: (1) 1080P camera can be used for attendance of the 6-meter classroom; (2) 4K camera can be used for a 10-meter classroom, the smallest size of a single face is $63 \times 63$ and the maximum is $156 \times 156$; (3) the 720P camera widely used in current classroom cannot meet the use of classroom automatic attendance.

As shown in figure 7, the SeetaFace face feature point detection has good robustness to human face attitude. When hardware device is invariant, the feature point detection can be realized when the head is deflected at an angle, it can meet the automatic attendance needs in real classroom.

### 4.3. Comparative Experiments of Faster R-CNN and SeetaFace

We further compare the robustness of Faster R-CNN and SeetaFace by increasing the variation range of face posture. Experimental results are shown in Figure 8: blue box is the detection result of SeetaFace, while red box is of Faster R-CNN. The results can be concluded as follows: (1)

Figure 10 (a) illustrates that SeetaFace has the ability to detect front faces almost the same as Faster R-CNN; (2) Figure 10 (b), (c), (d) show that SeetaFace can't detect the face when the student has abnormal classroom gestures, but Faster R-CNN can do it. It indicates that SeetaFace is not as good as Faster R-CNN when face attitude is too large. We use this difference to achieve automatic attendance about students who don't listen carefully and sitting in an irregular position.
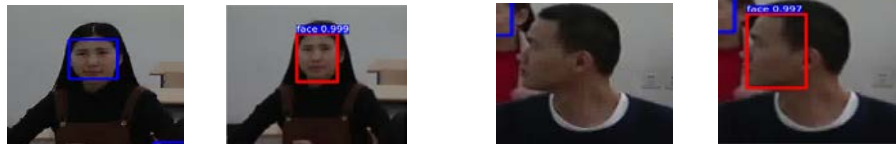


(a)Comparison of front face detection       (b) Comparison of side face detection
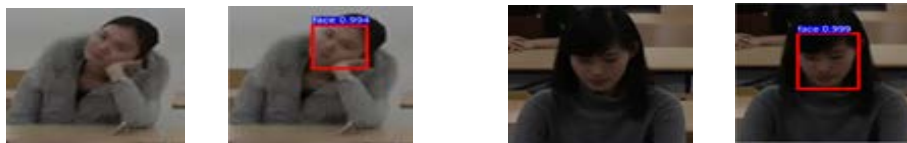


(c)Comparison of tilted face detection   (d) Comparison of face detection in the case of bow

Figure 8 comparison of face detection result between SeetaFace and Faster R-CNN.
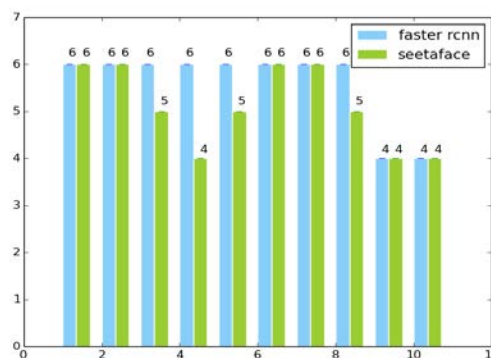


Figure 9 the comparison of Faster R-CNN and SeetaFace face detection results.

We use Faster R-CNN and SeetaFace for face detection, the results are shown in Figure 9, Faster R-CNN has a strong detection ability to detect all six faces with an exaggerated gesture, SeetaFace with a shallow network only detects 5 faces at 3, 5, 8 minute, and detect 4 faces at 4, 9, 10 minute. In face detection, the attitude robustness of SeetaFace is significantly lower than Faster R-CNN.

## 4.4.    Attendance Results

Finally, we sample every minute for a ten-minute class, and give the statistical attendance table, which is shown in Table 1, from which we can see: (1) Chang Wenquan achieves a full attendance, basically ensure facing the blackboard; (2) Li Huichao and Fu Rong have the problem of early departure, one minute before departure Fu Rong appeared carelessness.  (3) Liu Heng, Wang Dan, Lu Jiao appeared carelessness, where Liu Heng is in the first 3-5 minutes didn't listen carefully, Wang Dan is in the first 4-6 minutes, and Lu Jiao only in the sixth minute didn't listen carefully.

Table 1 Attendance statistics of six students.

| Time(min)<br>student | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Li Huichao | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | X | X |
| Liu Heng | ✓ | ✓ | ● | ● | ● | ✓ | ✓ | ✓ | ✓ | ✓ |
| Wang Dan | ✓ | ✓ | ✓ | ● | ● | ● | ✓ | ✓ | ✓ | ✓ |
| Chang Wenquan | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fu Rong | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ● | X | X |
| Lu Jiao | ✓ | ✓ | ✓ | ✓ | ✓ | ● | ✓ | ✓ | ✓ | ✓ |

## 5.    Conclusions

Based on deep learning, SeetaFace and Faster R-CNN have a strong capability for face recognition and face detection. To address the problem of low resolution, we use 4K HD video for face detection and face recognition. This paper applies Faster R-CNN and SeetaFace to class attendance, and achieves satisfactory results. It can not only give the four traditional attendance indicators about absence, later arrival, early departure and random access, but also gives the new indicator of carelessness. It's a promising class attendance technology.

In order to apply this class attendance system to practice in real class as soon as possible, we will continuously improve the class attendance system, increase the accuracy and speed, and ultimately achieve the high-precision real time attendance to meet the need of automatic classroom evaluation.

## Acknowledgements

## References

[1] R. Girshick, "Fast R-CNN," in IEEE International Conference on Computer Vision (ICCV), 2015

[2] Yang B, Yan J, Lei Z, et al. Convolutional Channel Features for Pedestrian, Face and Edge Detection [J]. Computer Science, 2015:82-90.

[3] Michał Dolecki, Paweł Karczmarek, Adam Kiersztyn, Witold Pedrycz, "Face recognition by humans performed on basis of linguistic descriptors and neural networks", Neural Networks (IJCNN) 2016 International Joint Conference on, pp. 5135-5140, 2016, ISSN 2161-4407.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held on December 3-6, 2012, (Lake Tahoe, Nevada, United States), pp. 1106–1114, 2012.

[5] Soniya, Paul S, Singh L. A review on advances in deep learning[C]// IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions. IEEE, 2015:1-6.

[6] Shuo Yang; Ping Luo; Chen Change Loy; Xiaoou Tang,"WIDER FACE:A Face Detection Benchmark," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Year: 2016 Pages: 5525 - 5533, DOI: 10.1109/CVPR.2016.596