I. Background & Problem

During the Covid-19 outbreak, we have to conduct the online exams. In order to keep fairness, I plan to develop the exam surveillance system to detect the cheating behaviours and four cases will be detected in my system including people counting, mouth opening, eyes movement, and cell phone detection.

II. Method

a. People counting problem

The system will detect the number of people by detecting the number of faces. Like the figure 1 below, the system will generate the warning message for multiple people and no person scenarios.

Figure 1. People counting scenario

Facial Landmark by Dlib

I apply the facial landmark from Dlib to detect the face. The facial landmark is used to label the critical facial attributes in the given image. And the first step is to locate the face in the image and the second step is to detect the critical parts including mouth, eyes, eye bows, nose, and jaw. Finally, total 68 points

are visualized to display the coordinates of facial structures. But for the people counting problem, only the first step is needed by using *get_frontal_face_detector()* function from the Dlib to detect the amount of faces.

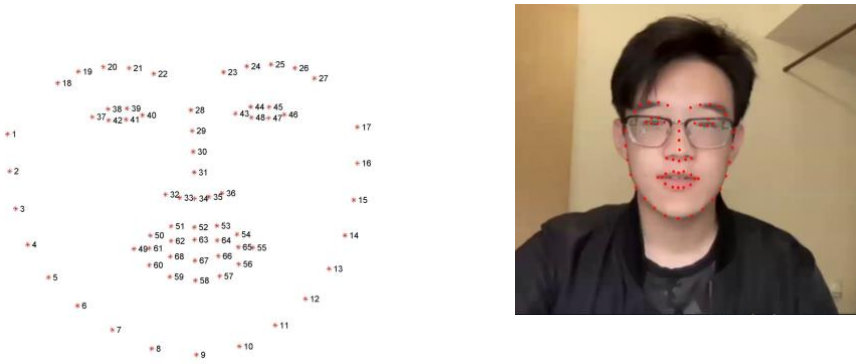

Figure 2. Facial Landmark

b. Eye moving

The abnormal eye moving is one of the cheating behaviours. Like the figure 3 below, if the eye moves frequency and distance are more than the thresholds defined. The abnormal behaviour will detected.

Figure 3. Abnormal eye moving behaviour and normal behaviour

Like the figure 4 below, the blue points are the parts of the eye bows. Record the X coordinate in the empty list for each frame, and if the list has more

than 6 components , the absolute value of the difference of the sixth element and the first element will be calculated. And if the value is larger than 30, the warning will be generated.

Figure 4. Facial Landmark detected for eye moving

c. Mouth opening

The abnormal mouth opening is one of the cheating behaviours. Like the figure 5 below, if the mouthing opening frequency and distance are more than the thresholds defined. The abnormal behaviour will detected.

Figure 5. Abnormal mouth opening behaviour and normal behaviour

Like the figure 6 below, the green points are the parts of the inner mouth. Get the value of the difference of the y coordinate of the upper point and lower point. And if the value is more than 8, the frequency will be added 1. Finally, if the frequency is more than 40, the warning will be generated.

Figure 6. Facial Landmark detected for mouth opening

d. For cell phone detection

The cell phone appearing is one of the cheating behaviours. Like the figure 7 below, if the cell phone appears , the abnormal behaviour will detected.

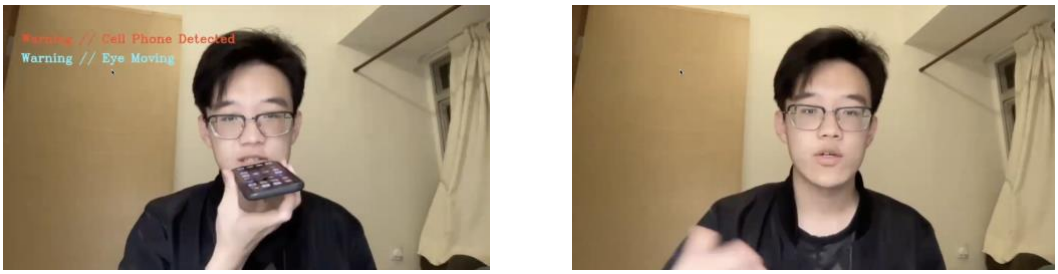

Figure 7. YOLO v3 detection for cell phone

For this operation, I apply YOLO v3 as the object detection algorithm. Like the below figure 8, the YOLO v3 has 106 layers of FCN. And there are 3 different scales to make predictions.

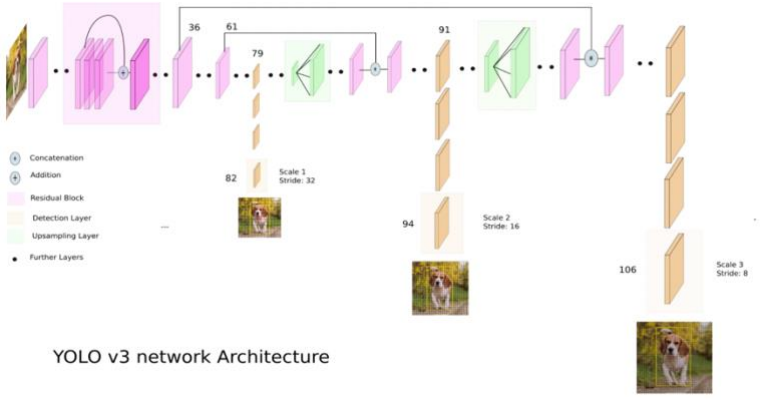

Figure 8. YOLO v3 architecture

III. Experimental Results

a. People Counting

The accuracy is 0.81 , while some frames cannot be detected like the below figure.

Figure 9. People counting failed

b. Eye Moving

I conducted several experiments on various thresholds of y_coordinate. And the accuracy is based on frame computation. And I select the 30 as the x_coordinate to achieve 0.72 accuracy.

X_coordinate threshold	Accuracy
10	0.60
20	0.68
30	0.72
40	0.70
50	0.69
60	0.67

c. Mouth Opening

I conducted several experiments on various thresholds of frequency. And the accuracy is based on frame computation. And I select the 8 as the frequency to achieve 0.78 accuracy.

Frequency threshold	
6	0.6
7	0.71
8	0.78
9	0.73

d. Cell Phone Detection

The accuracy is could reach to 0.92 by YOLO v3 and I tried R-CNN with 0.8 accuracy. And I applied the YOLO v3 for this project. But also there are some situations which cannot be detected like the below figure

Figure 10. Cell Phone Detection failed

IV. Discussion

For the eye moving and mouth opening, the key point is how to set the threshold value to control the abnormal behaviour. And the related experiments need to be conducted to test various threshold to get the best one. While another difficulty is to avoid the suspicious but not the cheating behaviour such as glance and yawn, for the current system, the solution is also using the threshold to avoid similar situation. Another problem is that how to combine the method to one model to detect the cheating behaviour without knowing any abnormal behaviour before the detection. I think it could achieve by YOLO algorithm, which could divide the frame into multiple cells and we could predict the object based on location and probabilities, and if the object location has been changed between the frames, it could be considered as the cheating behaviour.

V. Conclusion

For this project, I applied Dlib and YOLO v3 to detect four cheating behaviours including people counting, eye moving, mouth opening, and cell phone detection with the satisfied accuracy. And in the further work, more cheating behaviours should be detected such as test-taker, using other cheating devices and voice cheating. Also, the detecting method should be combined as one model to detect all of the cheating behaviours.