Automating Processing of Videos from Online Proctoring

Stony Brook University
Electrical and Computer Engineering

Faculty Advisor: Dr. Ridha Kamoua
Abrar Abdurrob, Sajidul Chowdhury, Khondaker Sadiq

Habibi

Overview

With the increasing proliferation of online programs and courses offered by universities and other institutions, identity verification and monitoring of students while taking exams is a critical and challenging problem. Some of the approaches taken, consist of requiring proctors and monitoring students through webcam video recording. The goal of this senior design project is to automate the analysis of the video recording by employing computer vision, artificial intelligence, and digital signal processing techniques to extract image frames that require further examination.

Background

CNNs have currently been the most effective method of image detection and classification. They have this edge on normal neural networks and MLPS due to their ability to receive matrices as inputs as opposed to vectors. We did not end up using CNN, as we found the “Faster RCNN 2018” was better suited to our technical needs.RCNNs work differently from normal CNN’s as they use a method known as selective search to split the image into 2000 regions. These regions are then passed to the CNN which will go through the object detecting process.

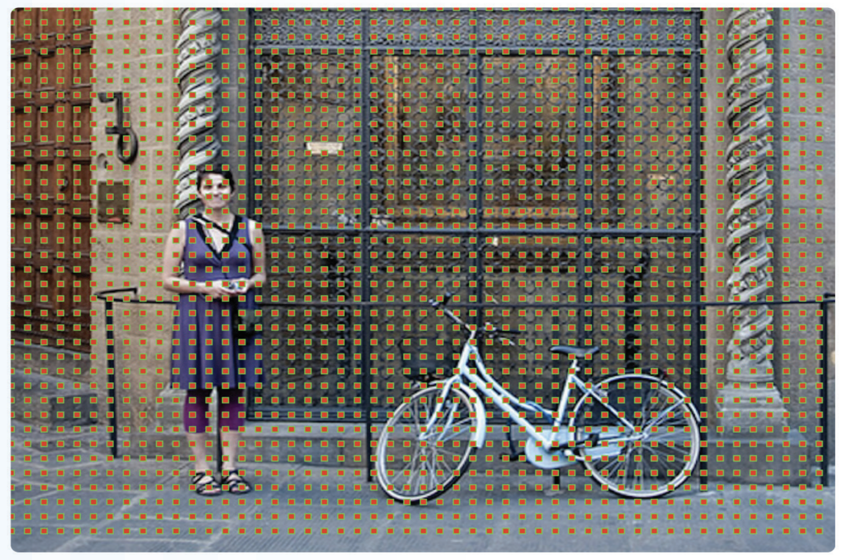

Fig 1: Shows how an RCNN views a photo in segments

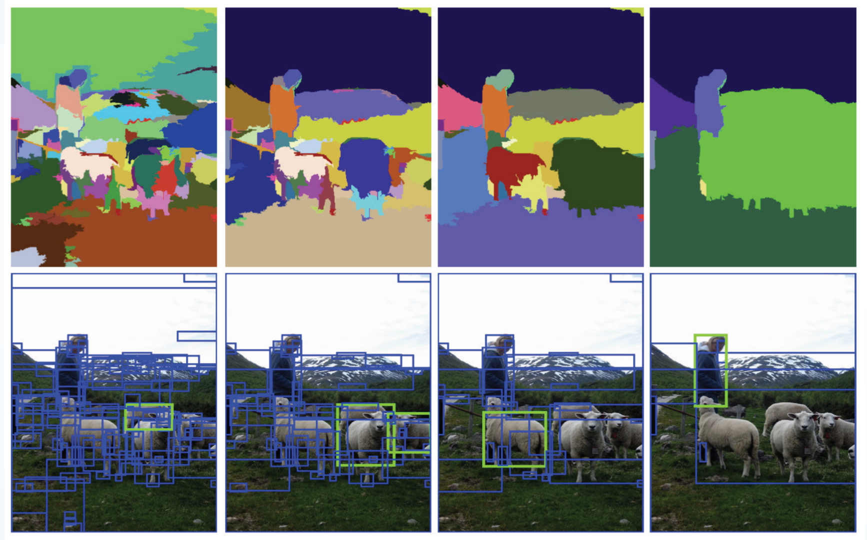

Fig 2: shows model attempting to group segments with like objects together.

Fast RCNN go one step further by not initially breaking the image into regions but instead passing the image to a CNN first. The feature maps that are produced from the image then go through the selective search method. This works quicker as the whole image can be passed through instead of 2000 smaller regions. Finally, “Faster RCNNs” work the fastest by replacing the selective search method with a region proposal network that allows for much faster classification.

Our Detection System

We have created a system where it monitors the students taking a test and detect the suspicious behavior. The suspicious behavior is flagged and a screenshot is taken with the time stamp and a report is generated with the screenshot and time of incidence. The suspicious activies include: looking at hand, looking down, phone on table, calculator on table and book on table.

We needed to have a good data set for the object detection to work. Although, most of the photos could be taken randomly, some photos were needed that could assure that the model can still identify rarer positions of the object.

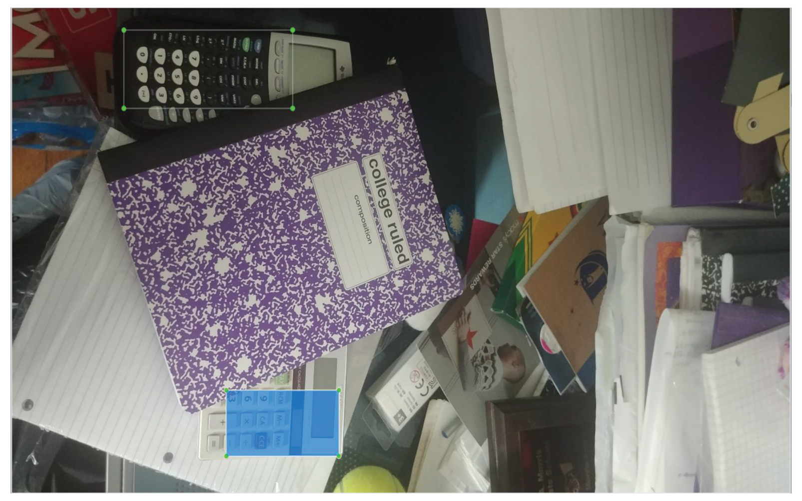

Fig 3: shows labelled calculators overlapped by a notebook while also having other objects in the surroundings

Figure 4: Several Objects of similar shape and color to a calculator with one labelled calculator

A phone looks similar to a calculator but should not be detected as one during the process. These extraneous photos help assure that the model is not overfitted and can detect the photo even in more difficult instances.

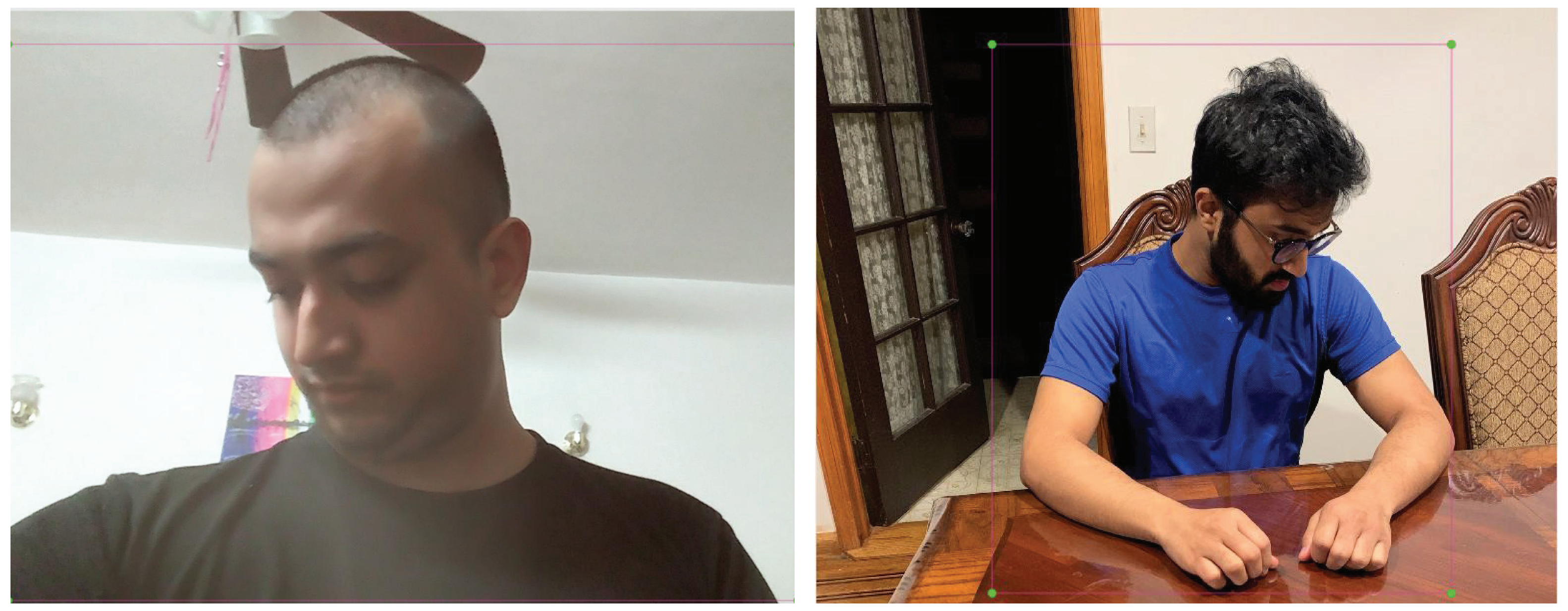

Fig 5: Images depict person looking right and left to cover different angles in which a person can look down

Similar to the object dataset the gesture dataset also needed extraneous photos for rarer instances. For example, a person looking down and to the left or right can also qualify as looking down.

Below are the entire detection system. It can be seen the system detects the raising of hand as well as looking down, the calculator and phone. To better our detection we would need to get more imags and train it.

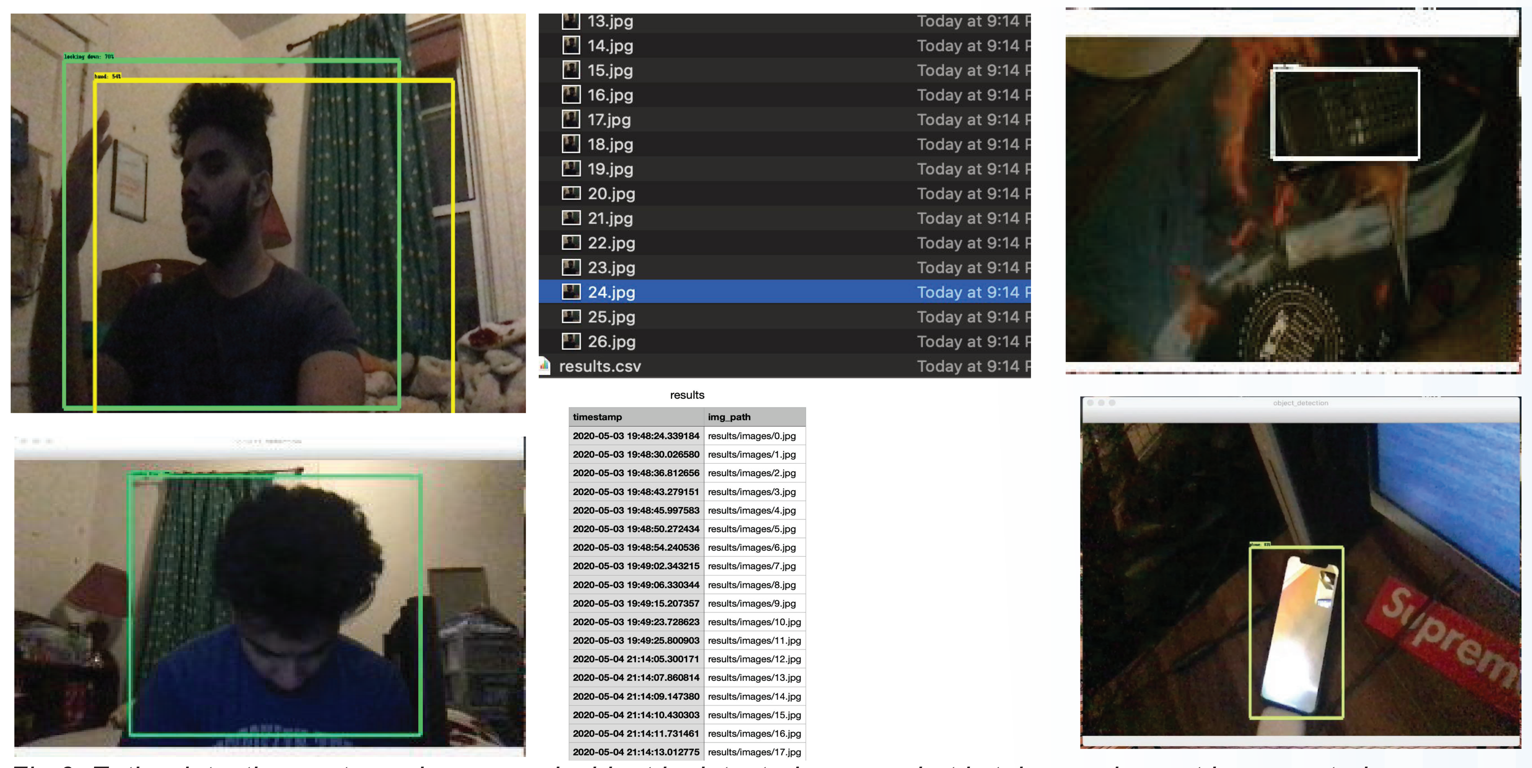

Fig 6: Entire detection system where once incident is detected screen shot is taken and report is generated.

Shown below are the loss and accuracy graphs for the model. Multiple graphs come from the different epochs that are run to train the model. The loss graphs are steadily declining which means that the model is slowly learning to be more accurate which is good for our product. The accuracy graphs for the models show a steady increase until sixty percent. This value could be higher with the introduction of more images and epochs.

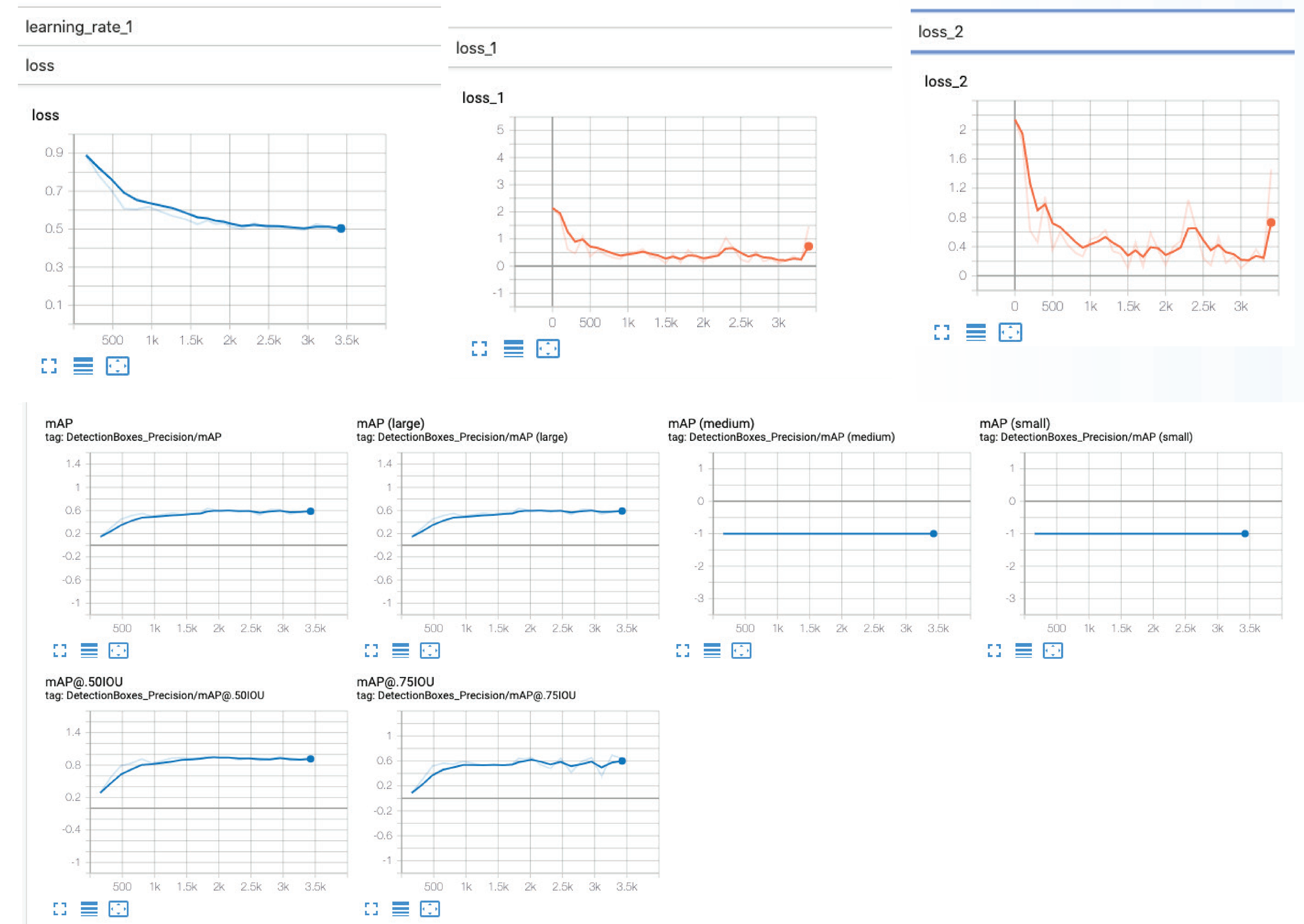

Fig 7: Loss and Accuracy Graph

Acknowledgements

We would like to thank Professor Dhadwal for working diligently to get us through the Senior Design even with the COVID-19 dilemma. We would like to extend our thanks and gratitude to Professor Kamoua for guiding us through this project.

Reference:
Xu, Y., 2018. Faster R-CNN (Object Detection) Implemented By Keras For Custom Data From Google’S Open Images Dataset V4. [online] Towards Data Science. Available at: <https://towardsdatascience.com/faster-r-cnn-object-detection-implemented-by-keras-for-custom-data-from-googles-open-images-125f62b9141a> .