Tasks:

1. Identify which frames in the clip have the staff present?

2. [Bonus] Locate the staff xy coordinates when present in the clip.

The proposed solution for the first and second questions is as follows:

1) Detect every person in the video.
2) Track and assign a unique ID for each person.
3) Index their bounding box's x and y axes coordinates.
4) Crop every person according to their bounding box and label them according to the number of frames and unique ID.

To accomplish these steps, different techniques are used:

- YoloV5: 1 and 3
- DeepSORT: 2

For the results, shown in the screenshots below, the staff (person wearing the tag) is detected and assigned ID number 48 on the left and ID number 116 on the right.

However, the tracker is not perfect as the staff was assigned different IDs when he is walking and another different IDs when wearing the black jacket.

- The ID when not wearing the jacket: 48,64,65,71,83.
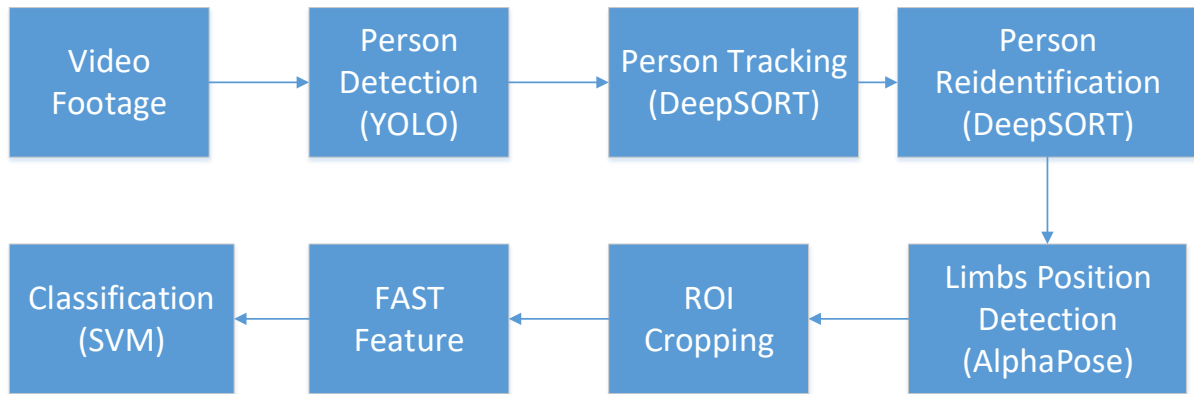- The ID when wearing the jacket: 102,113,116,121,131,138.

The frames where the staff is detected while walking according to ID number:

- ID 48: 366-392
- ID 64: 446-475
- ID 65: 452-469
- ID 71: 525-564
- ID 83: 593-594
- ID 102: 829-866
- ID 113: 913-933
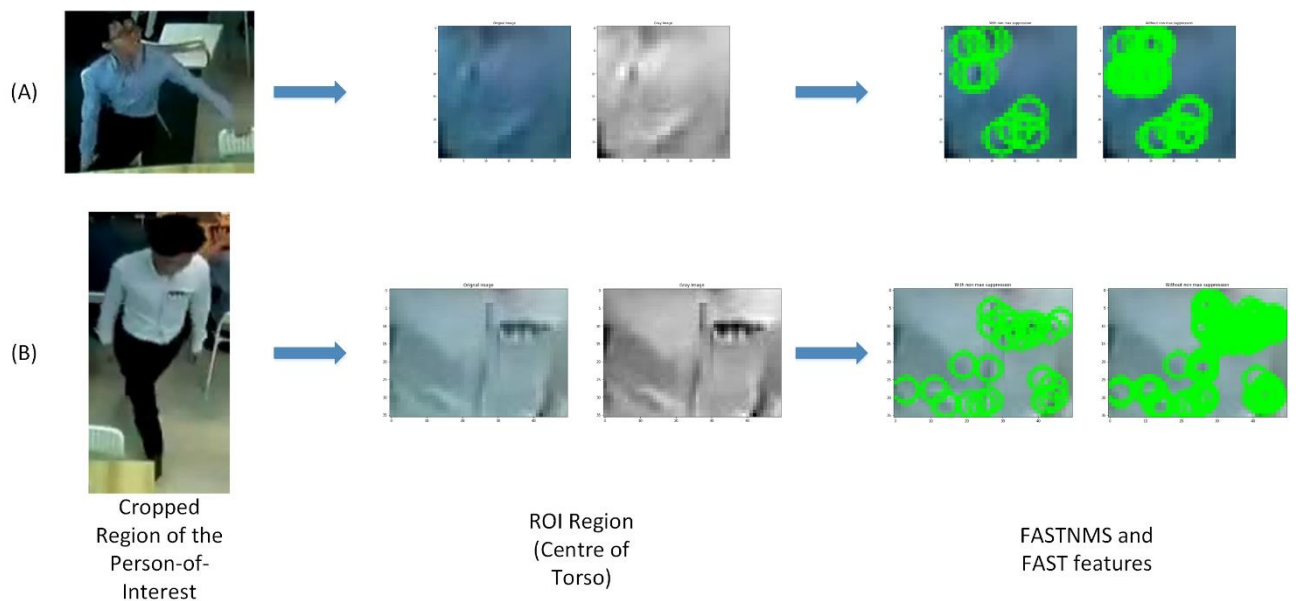- ID 116: 936-944
- ID 121: 992-1055
- ID 131: 1080-1127

- ID 138: 1178-1305

The coordinates of x and y axes are in the Excel file.

Assuming that we do not know which ID has the tag, it is necessary to detect the tag first before classifying the person as the staff. Here is the proposed framework to accomplish it:
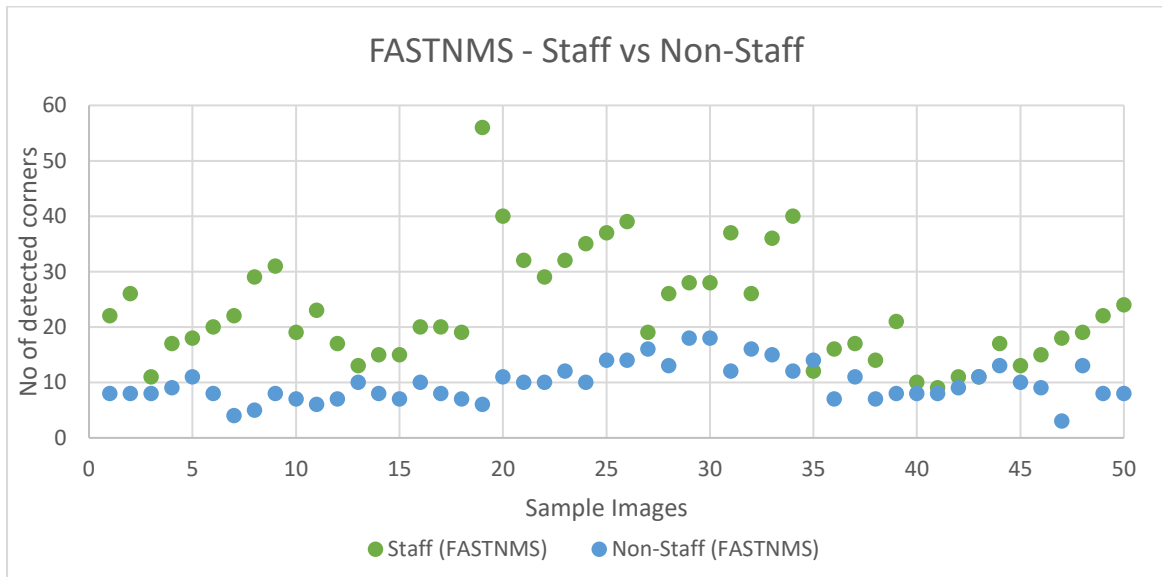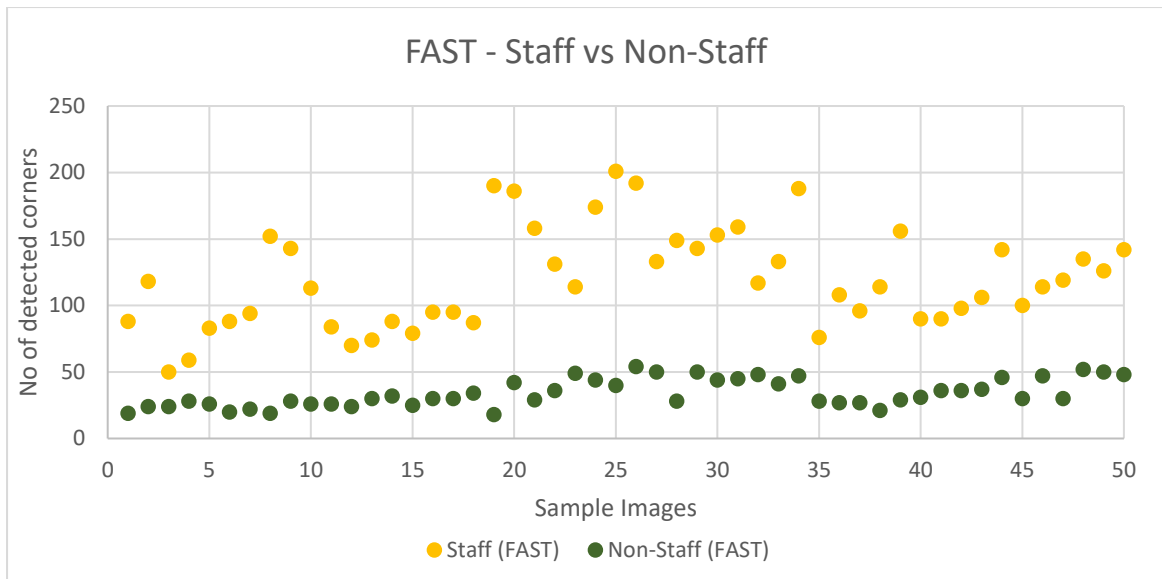

The Features from accelerated segment test (FAST) features refers to the number of detected corners from applying FAST and FAST with Non-Maximal Suppression (FASTNMS).

(A)

(B)

Cropped Region of the Person-of-Interest

ROI Region (Centre of Torso)

FASTNMS and FAST features

The idea is that the one wearing the tag will generate more corners as shown in the example above:

a) Number of Keypoints Detected In The Image With Non Max Suppression: 9
   Number of Keypoints Detected In The Image Without Non Max Suppression: 26
b) Number of Keypoints Detected In The Image With Non Max Suppression: 23
   Number of Keypoints Detected In The Image Without Non Max Suppression: 95

FAST - Staff vs Non-Staff

● Staff (FAST) ● Non-Staff (FAST)

FASTNMS - Staff vs Non-Staff

● Staff (FASTNMS) ● Non-Staff (FASTNMS)

The two graphs above shows the interclass separation between different classes and allows high classification rate when the features are utilized by the classifier such as SVM. This also mean that it is possible to identify the staff by determining if he is wearing the tag or not by using corner features.