Footfall Cam Assessment – Detecting Staff

The proposed approach composes of 3 key components which break down into **detecting people in the frame, extracting the frame, and determining whether the people detected is wearing a tag**. In the code implementation, the first component, which is the people detector used is Histogram of Gradients (HOG) provided by the OpenCV is selected due to its fast performance and decent accuracy. It is effective in detecting objects with distinguishable textures and patterns, which makes it a good choice for detecting people. Using the HOG detector, the Region of Interest (ROI) where the people are detected can be extracted using the coordinates of the boxes returned by the detector model. This concludes the processes of the first 2 components in determining whether the staff is in the video. To determine whether the person extracted is a staff, contour detection method is used to determine whether the tag is on the staff or not. This is because the tag has a very contrasting texture, and patterns compared to the people's clothes. Therefore, using thresholding provided by OpenCV, the tag will be marked as white area in the threshold image. Then, the "findContours" function can easily identify the tag in the video. If the contour's area is smaller than a 50-pixel square, then the contour detected is classified as the tag. To further enhance accuracy, the contour within certain region determined will only be classified as tag because the tag will be put on by the staff in the same place in cloth area. Thus, the program will classify the frames where the tag is detected within certain area as the staff is present.

Thought Process:

1. Why use HOG from OpenCV?
   - Fast, can detect people very quickly compared to other models like YOLO
   - Decent performance although not the best.
   - Easy to implement

2. Why use Contour Detection approach compared to ML/AI methods?
   - Computational efficient, it does not require much in terms of hardware specifications

- Performs quite well when there is no labelled data to build ML model to distinguish staff and customers.

Limitations

1. This project does not implement object tracking that records the movement of the object (people) detected within the video.
2. The accuracy of the contour detection is highly dependent on the quality of the video as it is affected by the contrast of the tag compared to the neighboring pixels.
3. The video has fisheye lens issue which distort the frame can affect the performance of the model detector as the model is most likely trained on flat image.

Recommendations:

1. Given more data of staff and customers, a more intelligent approach can be used such as ML/AI methods by building classifiers to distinguish between staff and customers.
2. Given access to the camera, can use checkerboard method to re-calibrate the camera video to get the flatten image for better performance.
3. Super-resolution techniques can be attempted with the aim to increase the quality of the image by eliminating blurry frames.

References:

1. https://machinelearningmastery.com/opencv_hog/ [HOG]
2. https://www.tutorialspoint.com/white-and-black-dot-detection-using-opencv-python [Contour detection]
3. https://stackoverflow.com/questions/39441935/how-to-get-x-y-position-of-contours-in-python-opencv [Calculating contour coordinates]