Proposed Method

Since the relative position of camera and the bus is fixed, we adapt a background subtraction related method. We first apply frame by frame absolute difference to find out moving objects, then try to focus on door movements to identify the state of the door.

The moving objects can be classified into the following categories: people, bags, windows, and door. For people and bags, we import the YOLOv8 model to detect them and filter out their movements. For windows, we design a two pass method to tackle them. Since the lighting condition varies significantly on the windows, their average value after applying absolute difference tend to be high. Therefore, in the first pass, we will first calculate the average value of absolute difference in each pixel. Then the average value will later be used as mask to mitigate the noise of flickering windows.

After those processing, we can obtain a clean frame which only reflects the door movements. Since the door only moves during opening and closing, we can identify door state by simply summing up the pixel value of the processed frame. With value higher than a certain threshold we'll define them as opening and closing.

Techniques

Preprocess

- Two pass method to obtain average pixel map

Detection

- YOLOv8
- Frame by frame absolute difference
- Mask out middle area for noise reduction

Postprocess

- Clipping value of outliers by upper bound and lower bound
- Gaussian filter to smooth the curve
- Clustering on candidates in case making duplicate prediction on single opening/closing