

CMP\_SC 8690: Computer Vision

## Homework 4A: Semantic Segmentation Using Pre-Trained Deep Learning Networks

By: Mikey Joyce

Due: 4/2/2024

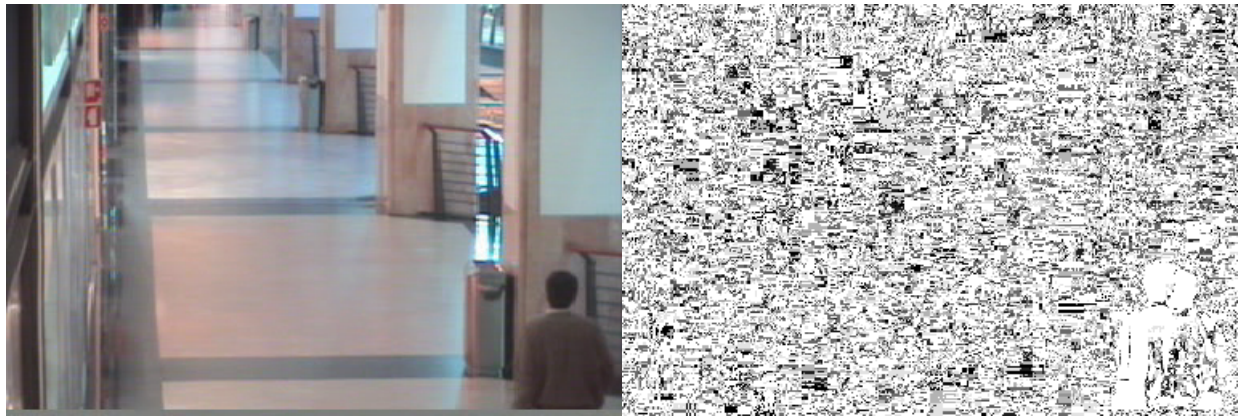
### Abstract:

The goal of this assignment is to dabble in video analysis using computer vision techniques. More specifically to segment motion from the image utilizing a simple background subtraction algorithm that utilizes a gaussian to model the background of the image and to detect differences of the pixels with the variance, high differences mean something in the frame is moving.

### Experiments and Results:

Alpha = 0.01 and TM = 2:

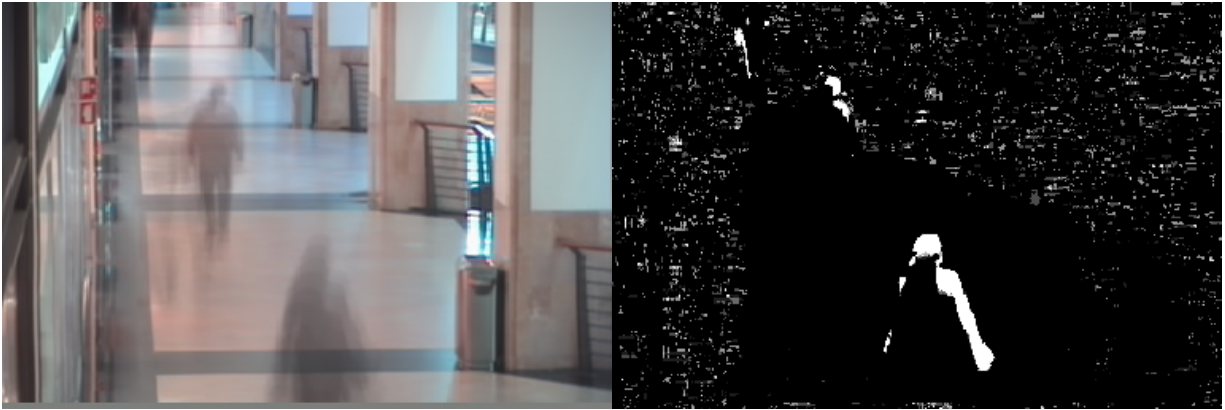
Frame 5



Frame 100

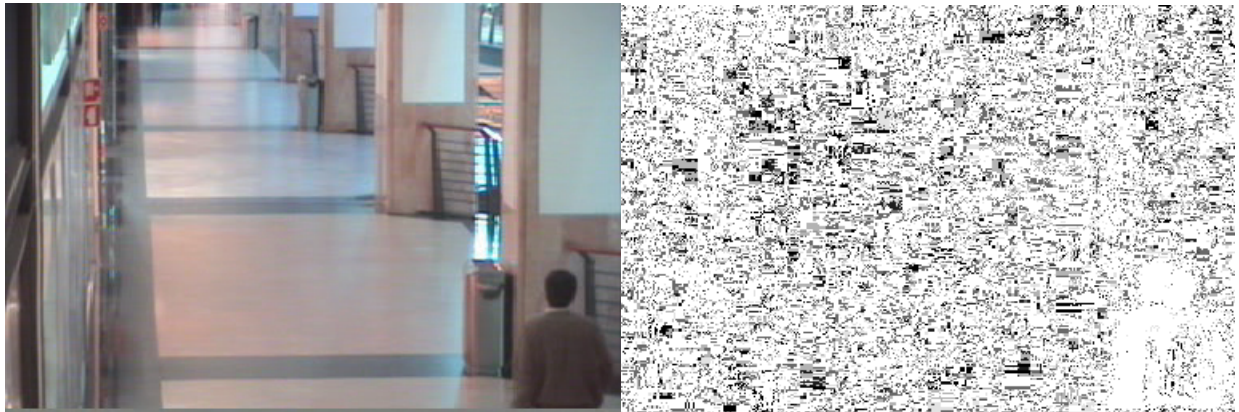


Frame 400

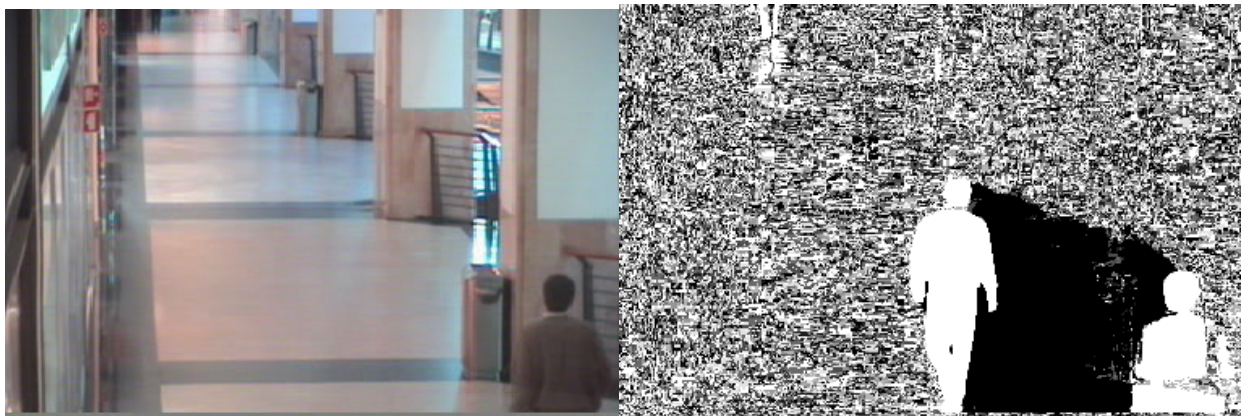


$\text{Alpha} = 0.001$  and  $\text{TM} = 2$ :

Frame 5



Frame 100



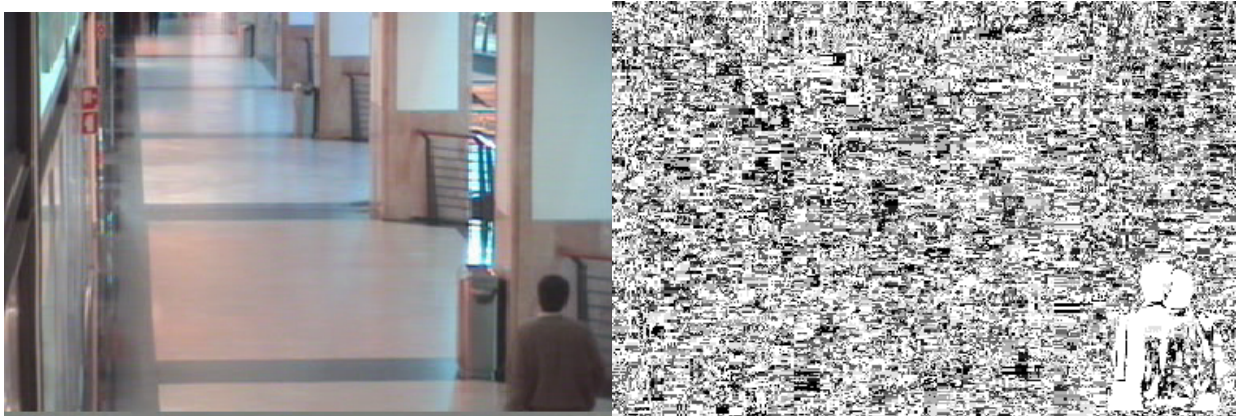


Frame 400



Alpha = 0.01 and TM = 3:

Frame 5



Frame 100

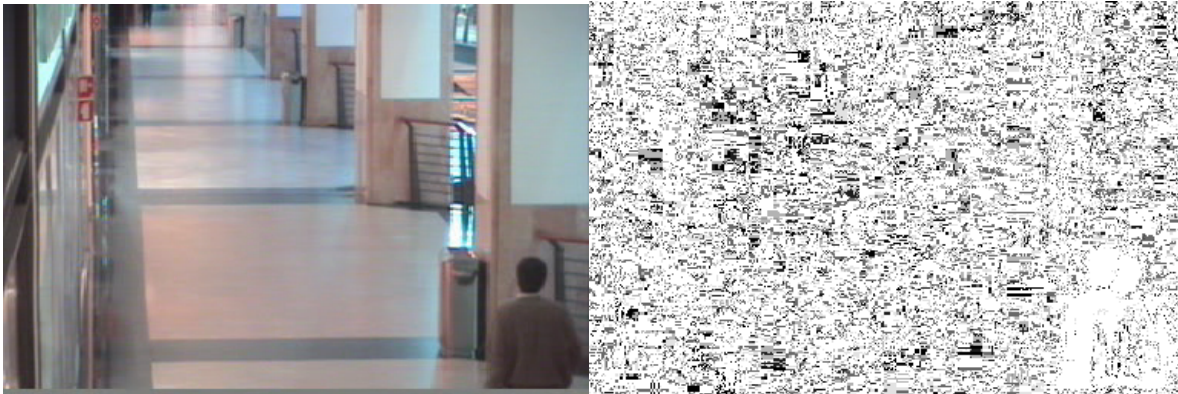


Frame 400

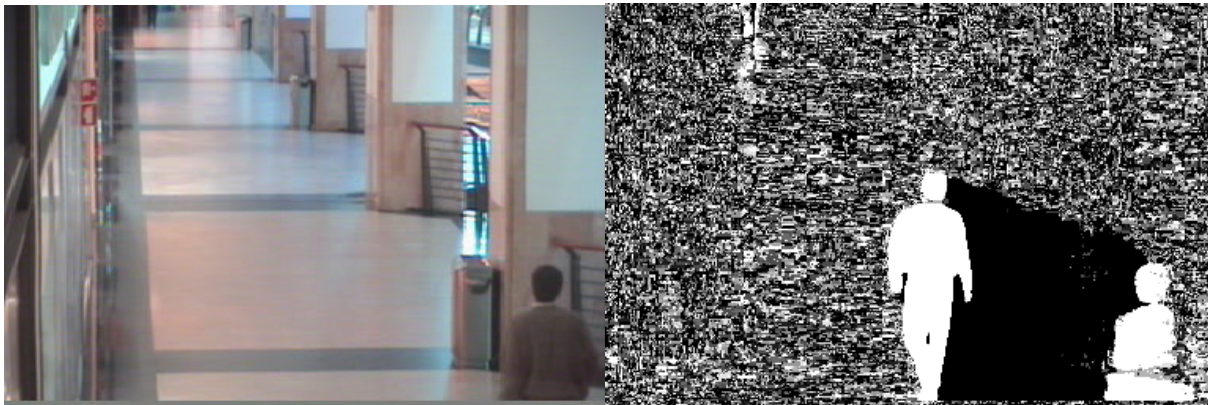


Alpha = 0.001 and TM = 3:

Frame 5



Frame 100



Frame 400



### Conclusion:

In this assignment we were asked to implement a simple background subtraction algorithm given the boiler plate code over the CAVIAR1 video set. We were then supposed to experiment with a couple of different parameters for the model to see how it effects the result. I believe my algorithm had an adequate result. Based on the results it is clear that the alpha parameter, or the learning rate parameter, effects how much the BG model actually changes through time. A larger alpha allows for more change, meaning the BG model changes a lot. You can see proof of this because the person on the bottom right of the video stays in the BG model for longer when the alpha is smaller. The TM thresholding parameter effects the resulting mask. If we have a TM that is lower, there is more noise in the resulting mask. Check the 1<sup>st</sup> and 3<sup>rd</sup> experiment on frame 100, as you can see the first one (the one with a lower TM) had more noise in the resulting image. The best parameter setup for the given model based on the parameters tried was probably  $\alpha=0.001$  and  $TM=3$ , however there are probably different parameters that were not tried that give a clearer result. There seems to be a lot of noise in the resulting mask, especially at the beginning of the video, this could possibly be combatted by using a more complex background subtraction model, or utilizing another method to perform the classification of the pixels.

### References:

- Libraries and tools: PyCharm, OpenCV, NumPy, Preview.
- Lec18\_MotionAnalysis.pdf
- CV2024\_HW4B\_SimpleBackgroundSubtraction\_Python.pdf