Journal Report 2
9/7/19-9/16/19
Avyuk Dixit
Computer Systems Research Lab
Period 4, White

Daily Log

Monday September 09

Finished downloading entire dataset. Attempted to run code mentioned before through entire dataset locally on my computer. This ended up overwhelming my memory and making it so that I could not use my computer.

Tuesday September 10

Attempted to make a number of changes to my code to make it use less memory space. For example, instead of looking at every frame in every video, I experimented with looking at 1 in every 5 frame in non-violent sections of videos and cutting large videos (5 minutes plus) closer to the time frames in which violence was occurring.
Even after such changes, my computer couldn't handle running the code which made me think about uploading all my data to remote servers and running code on them.

Thursday September 12

Attempted copying entire dataset to remote server. This did not work on the first attempt since the size of the dataset surpassed the 4GB quota. After my quota was increased, I was able to copy all of my data to the remote server. I installed necessary python libraries for my work and ran my code through the dataset.
Ran code remotely and it crashed after three videos. Moved everything to hpc12 and received an error related to not having built the C distributions necessary to run code. Spent some time trying to resolve this error but was not able to and probably need help doing so now.

Timeline

Date	Goal	Met
Today minus 1 week	Settle on a dataset to use for the project.	Yes, received access to the NTU CCTV-Fights Dataset
Today	Run pre-processing code to decompose surveillance feed into individual frames marked as violent or non-violent through entire dataset.	A number of memory related errors made it so that I was not able to meet this goal.
Today plus 1 week	Review OpenPose code and determine best statistical representation for poses.	
Today plus 2 weeks	Run OpenPose code through dataset of frames marked as violent or non-violent.	

Reflection

I'm not sure why my code is causing so many memory errors. When I was working on this step of the project for my science fair project last year, I didn't have any memory related issues. Granted, the dataset I used then was much smaller and consisted of like 1-2 second videos whereas the ROSE dataset consists of videos ranging from 1-10 minutes in length. It could still be the case that the memory related problems are coming from a more subtle part of my code.
Worst case, I should be able to run my code on parallel machines. I don't have any prior experience doing this but will probably learn about it soon in Parallel Computing. The fact that there's already memory errors in just decomposing videos and classifying their frames as violent and non-violent makes me worried about how much memory OpenPose will take in running through the new dataset, given the immense computational power and amount of data it generates.