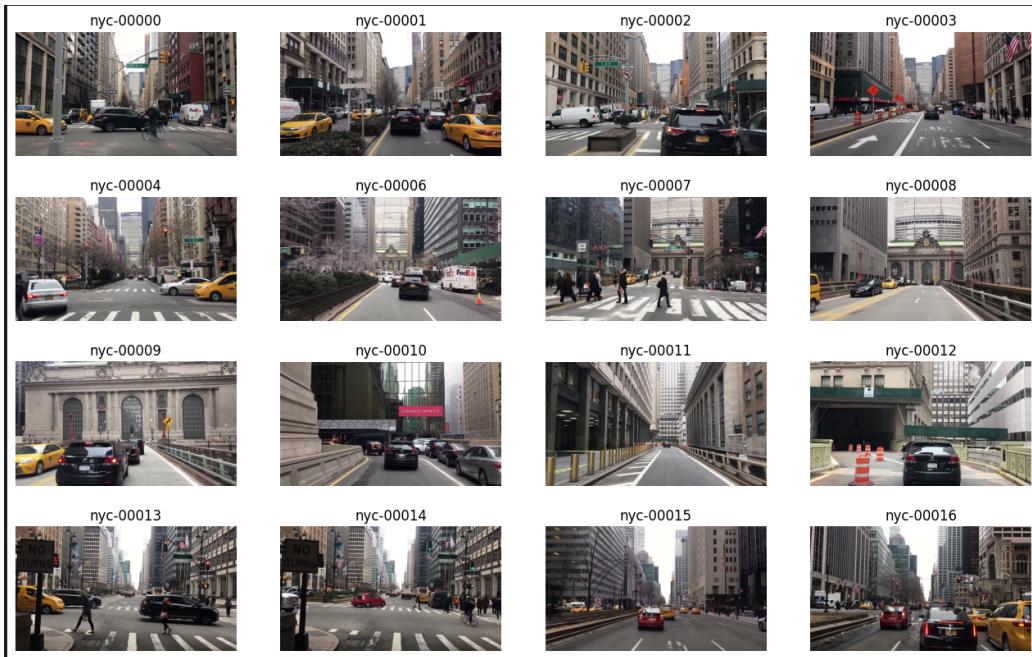


Using the “[0. Generating images.ipynb](#)” notebook, I developed the raw dataset by taking different frames from this [video](#) using the python **supervision** library like shown below.

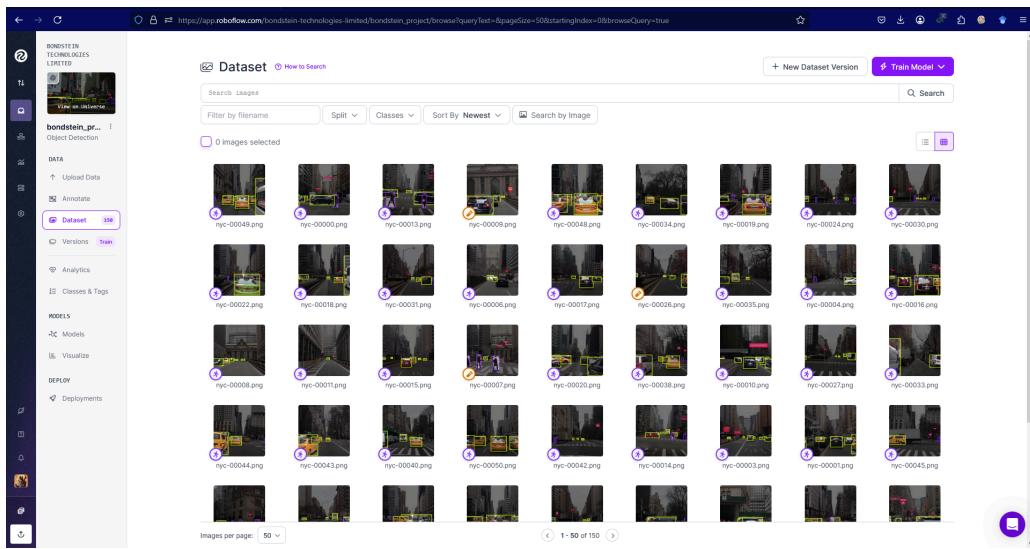


Then using the “[1. annotating dataset.ipynb](#)” notebook, I did automatic labeling/annotation of those videos by defining the 3 classes (vehicles, pedestrians, signboards). I used the autodistill’s **CaptionOntology** and **GroundedSAM** modules to automatically annotate the images like shown below.



Also, I used the ROBOFLOW platform to annotate the images. But while doing manual validation, I noticed some tags were missing and a few annotations of some images were

wrong or inaccurate so I fixed them manually and labelled them properly myself. Then did a 80:10:10 split as train-test-valid folds like shown below to finetune my model (YOLOV8).



I also applied different augmentation techniques to generate more samples for the final version of the dataset that I used to finetune the model.