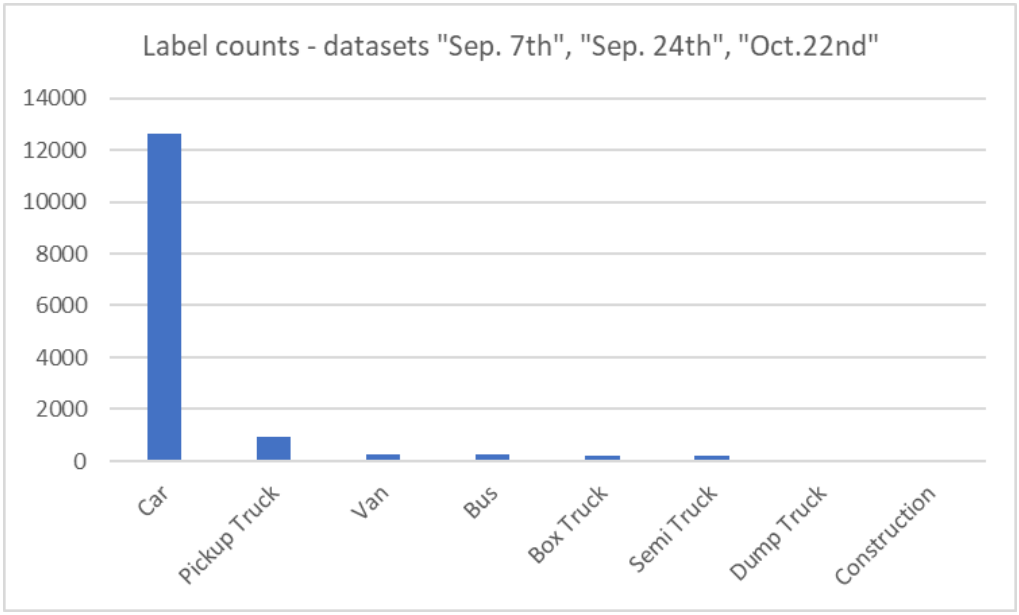
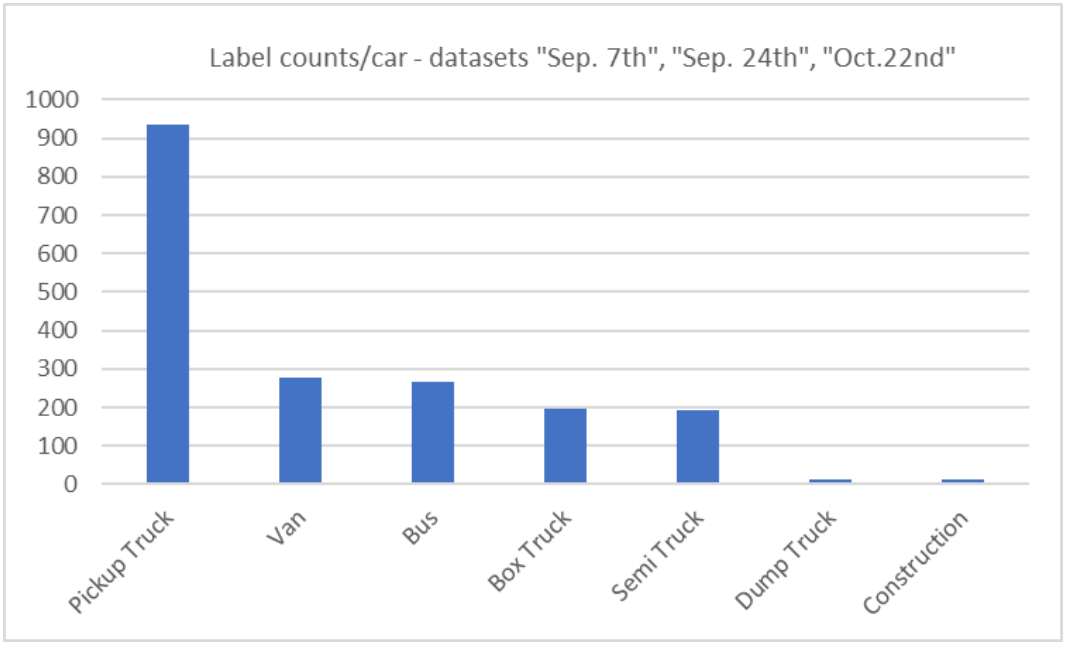
Distribution Report

Below is a chart showing how many instances of each label were found in the dataset, giving us an idea of the distribution of each vehicle type. As we can see, it is heavily biased towards car, which is to be expected.

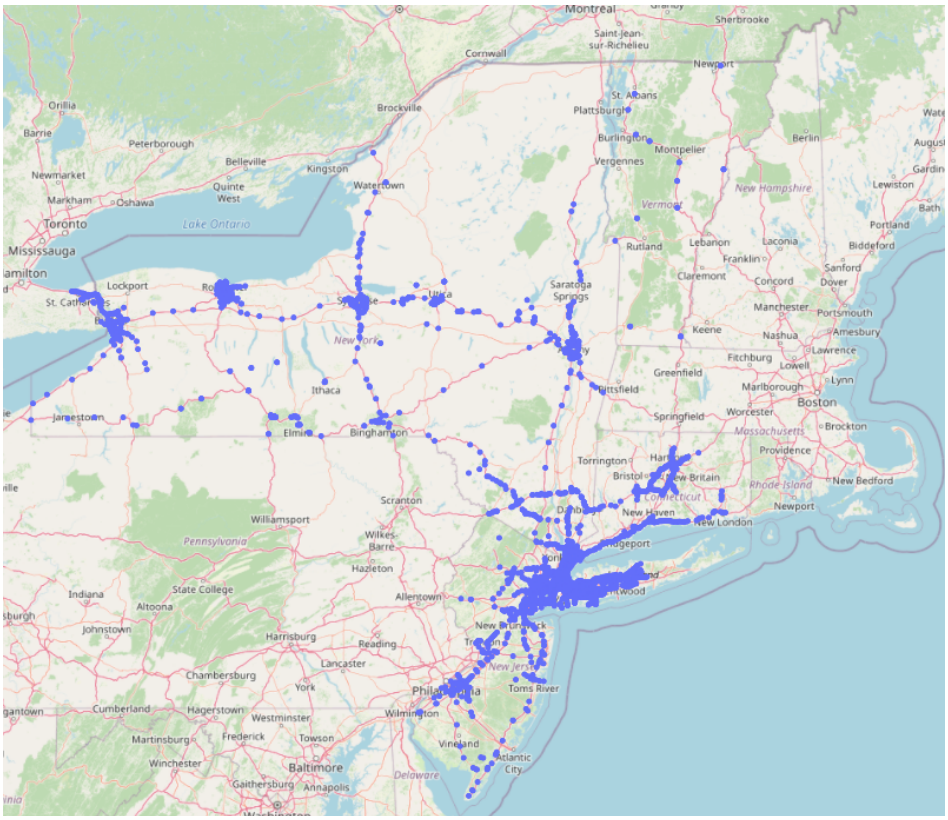


Below is a chart of the same data as above, with cars excluded.



There is a severe dearth of examples of vehicles that are not cars, and to an extent pickup trucks, in the data naturally obtained from traffic footage. YOLO typically expects 1500 examples per class for good performance, and so there is a need to augment the data, eg. with web scraping. The issue with web scraping is that the images obtained tend to be very different from the use case. Depending on the particular query, they may be either isolated images of just a particular type of vehicle, or they may be too high definition compared to footage from traffic cameras, or they may be taken from angles that are atypical for traffic footage. One solution that perhaps could be used to resolve this issue is to create artificial instances in our dataset by performing small mutations on our existing instances. This can be done with convoluting images with random noise, or deliberate noise to challenge the model, in the vein of adversarial examples

Now, we look at the distribution of the cameras we gather our footage from. Below is a visualization of where these cameras are on a map:

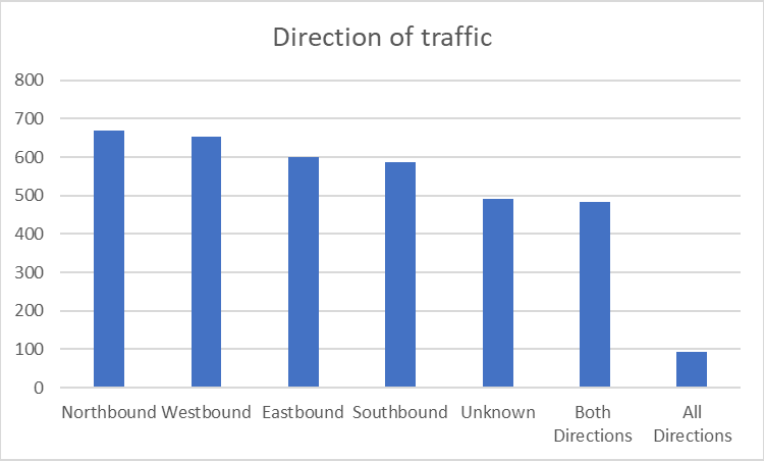


Here is a link to a Colab file to generate an interactive version of the above map (you need access to the JSON file of camera information): [Camera Locations Notebook]

Most of our cameras are focused around NYC and Long Island, with other locations largely being in upstate New York, New Jersey, and Connecticut. Cameras are focused on urban areas or major highways. Our footage is thus largely from busy roads. In fact, roughly 35% of the cameras record Interstate Highways. This is a conservative estimate since it only counts cameras with correctly inputted road names, which is not true in general.

This is backed up by anecdotal evidence while labeling the data, as most images seem to be of very crowded roads. Intuitively, this would be good for the use case, both because busier roads are likely more significant sources of emissions, and being able to identify vehicles in a crowded setting probably lends to good performance in less busy images as well (perhaps with issues in smaller than expected bounding boxes, something that can be noticed in the pre labeling on occasion)

Below is a table showing the directions of the roads we have cameras on:



The difference between the four directions is fairly negligible, but noticeably, most of our traffic footage is only from roads with cars going in only one direction, with only about 13.5% of cameras focusing on bidirectional roads, and only 2.5% of them focusing on crossroads. While direction of traffic is likely to have little bearing on the model, the lack of crossroads is potentially a problem for the model. From experience, almost all cameras are pointed such that cars are either driving towards or away from the camera. This means there is a lack of examples of cars oriented horizontally, i.e driving across the view of the camera.

A very brief review of labels from the September 24th Image dataset points to the fact that mistakes in labeling are rare but present. A review of approximately 15 images uncovered 7 mistakes out of a total 89 labels. These mistakes were split between correct boxes but incorrect labels; boxes drawn around objects that are not cars, or could not be feasibly identified as a car without context of other frames, and accidental double boxes around the same vehicle. A review carried out by Eric on 20 images saw similar results, outlined in the table below: A more detailed review would have to be conducted before an estimation of accuracy can be made.

	Car	Van	Pickup	Box	Semi	Bus
Attempts	73	8	4	5	11	3
Correct	81	11	7	5	13	4