Traffic Environment ~~Semi~~-Supervised Learning

Atul Gandhi, Elliot Silva, Ziv Schwartz
(ag6776/egs345/zs1349)@nyu.edu

Problem Definition

- Self-driving vehicles is the latest tech craze!
- Data collection for this problem has already taken off on a massive scale.
- Many decision are necessary in building an autonomous vehicle:
  - Making sense of the world around you ←[we are here]
  - Learning the mechanics of driving
  - Ethical decision-making
- This project processes real-world images collected from six cameras on a vehicle, and creates a 2D top-down representation of the world.
- This representation would be used to help the car make decisions (brake, accelerate, turn, etc.)

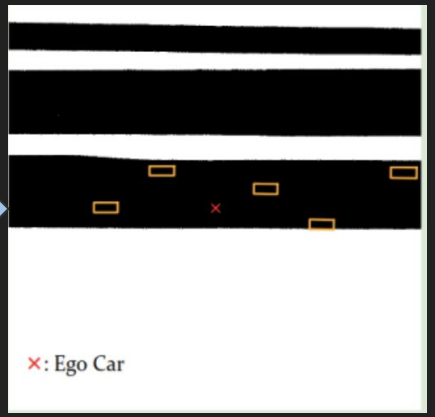
Our Intuition

6-Image Sample

Pre-processing

Stitched Representation

Training



Binary Road Map

Bounding Box

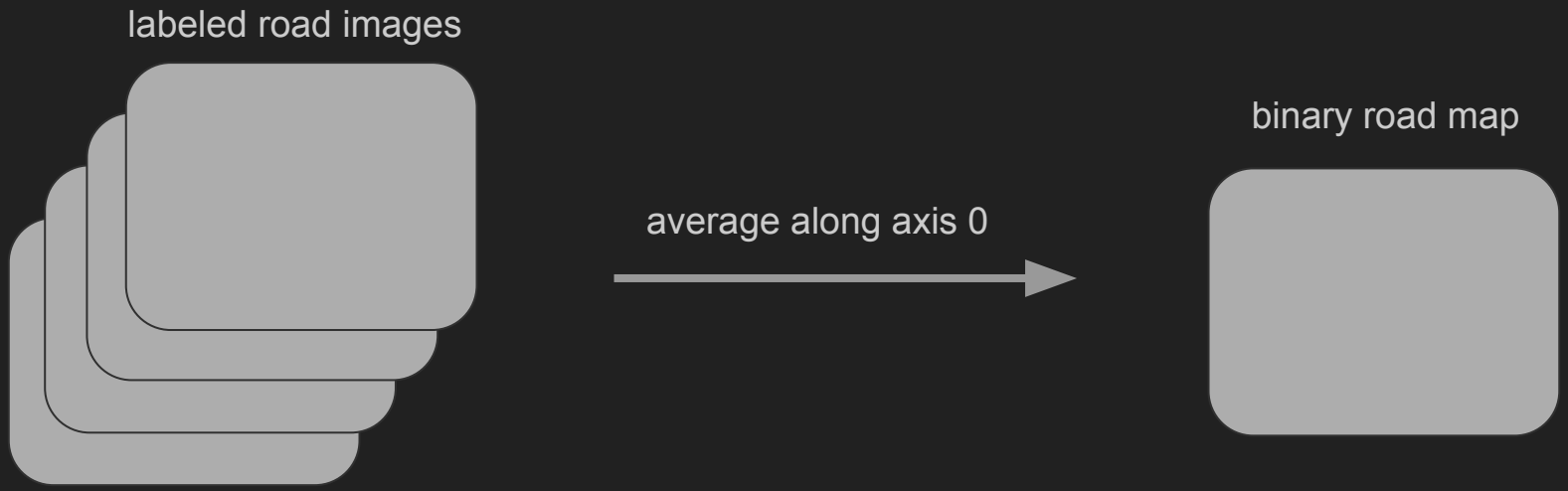
- Used the YOLOv3[1] architecture to place bounding boxes over the stitched representation
  - Uses Darknet-53 as the feature extractor
  - Comprised of 3 YOLO layers i.e. object detection is done at 3 different scales
  - Trained from scratch to identify and classify 9 types of objects

- Labels passed to the YOLOv3 architecture are bounding box coordinates & object category
  - Model learns transformation to the binary road map space
  - Model also classifies the detected object

1. https://github.com/ultralytics/yolov3

Binary Road Map

- The labeled set consists of 3528 road images. For the binary road map, our model utilizes these road images and returns the binary average representation (thresholded at 0.5) as the prediction for future road maps.



Results

- Our model obtains average threat scores on run_test.py for the tasks:
  - Bounding Box: 0.003217
  - Binary Road Map: 0.652

- Our model performs much better on the road map task
- Low threat score for bounding box task can be explained by the difficulty in training model to transform object point-of-view