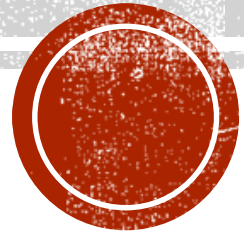
AUTONOMOUS DRIVING

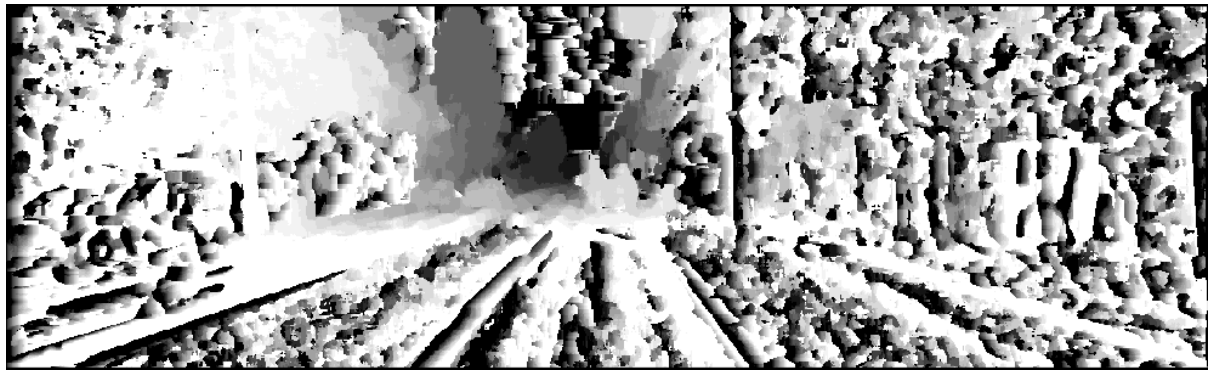


Shiqi Lin & Fanxuan Guo

COMPUTE DISPARITY

Method 1: Implementation based on lecture

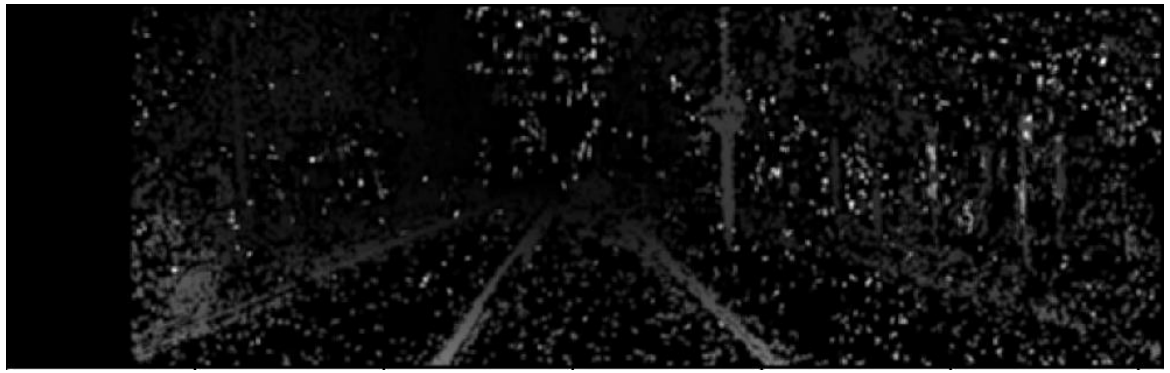
- Patch Size 10
- Max disparity 16
- Store seen right image vectorized patch in matrix
- Bad: Slow when patch size or max disparity is large



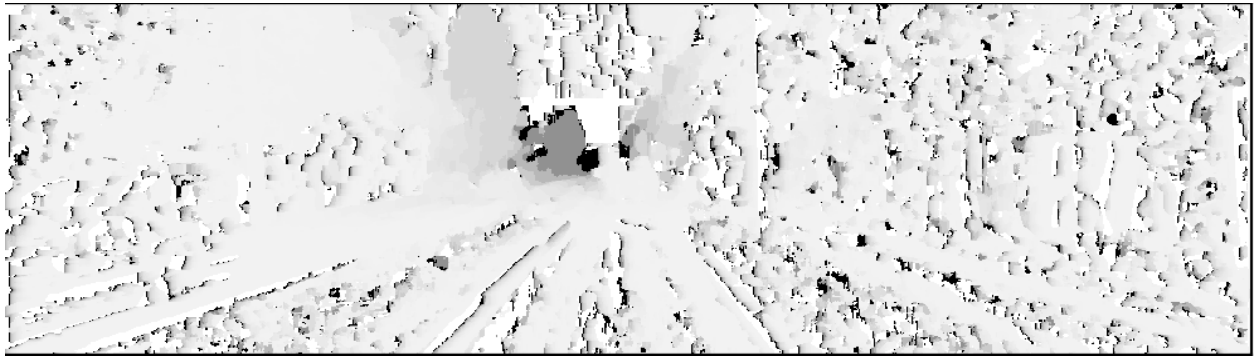
Method 2: Opencv Library

- Patch Size 5
- Disparity Range (4, 132)

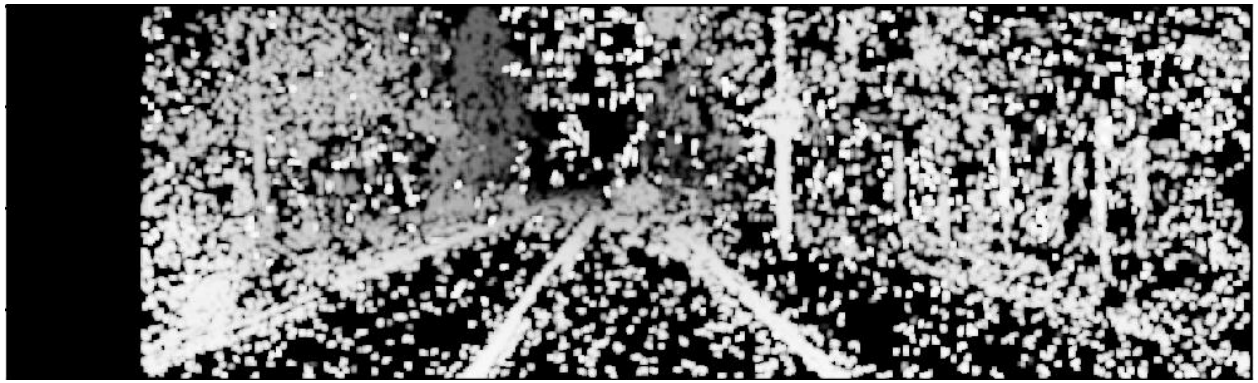
- Good: Fast under any condition



Method 1 Disparity -> Depth



Method 2 Disparity -> Depth



COMPUTE DEPTH

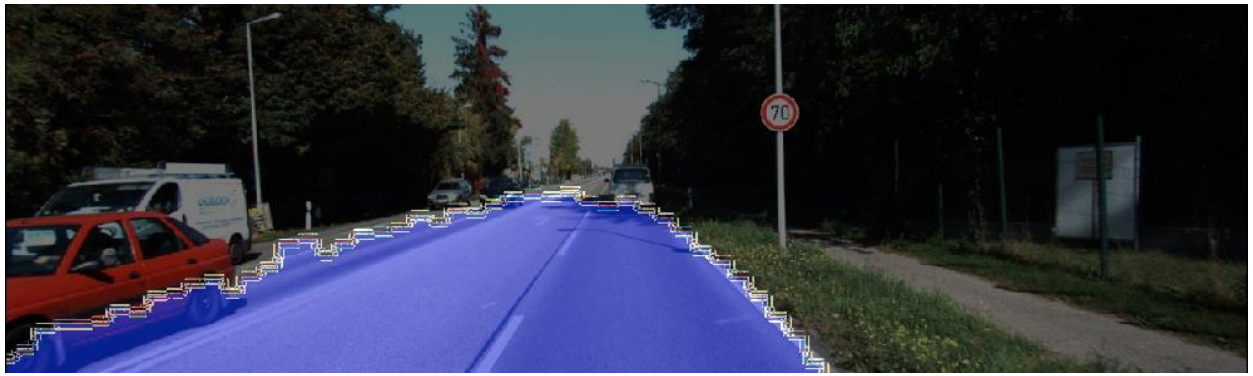
Formula From Lecture:

$$Depth = \frac{\text{focal length} * \text{baseline}}{\text{disparity}}$$

ROAD DETECTION

METHOD 1

- Proposed Result:
  - Enhances feature extractions
  - Highest Accuracy and Less Processing Time among existing road segmentation algorithm: MAP, StixelNet, Up-Conv-Poly
- Input:
  - 5-channel Image (RGB + XY coord)
  - Processed by pyramid prediction scheme
- TRAINING:
  - CNN – LSTM model
    - Encoder → Feature Processor → Decoder
- Output:
  - Road boundary vector

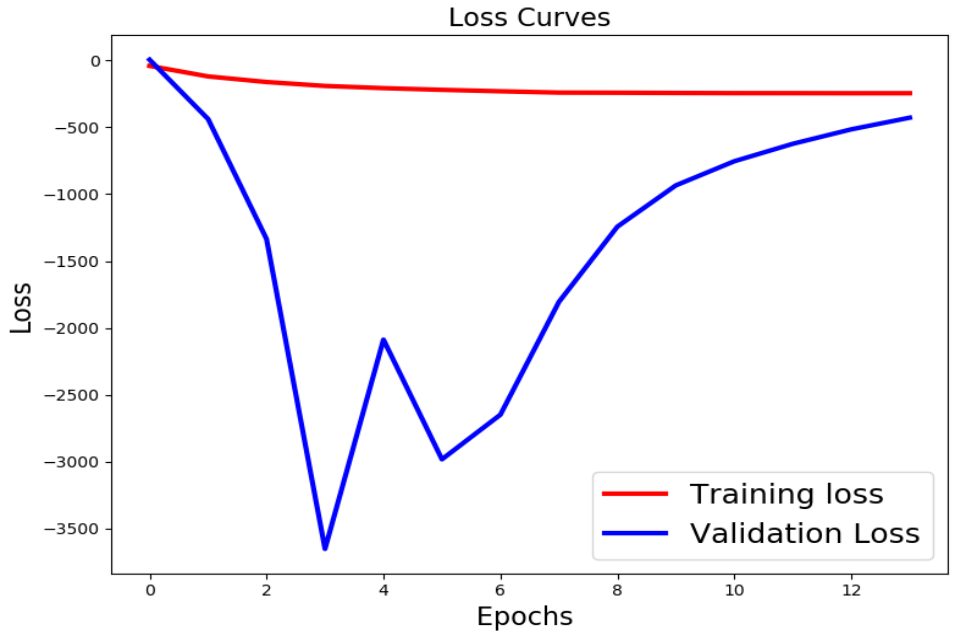


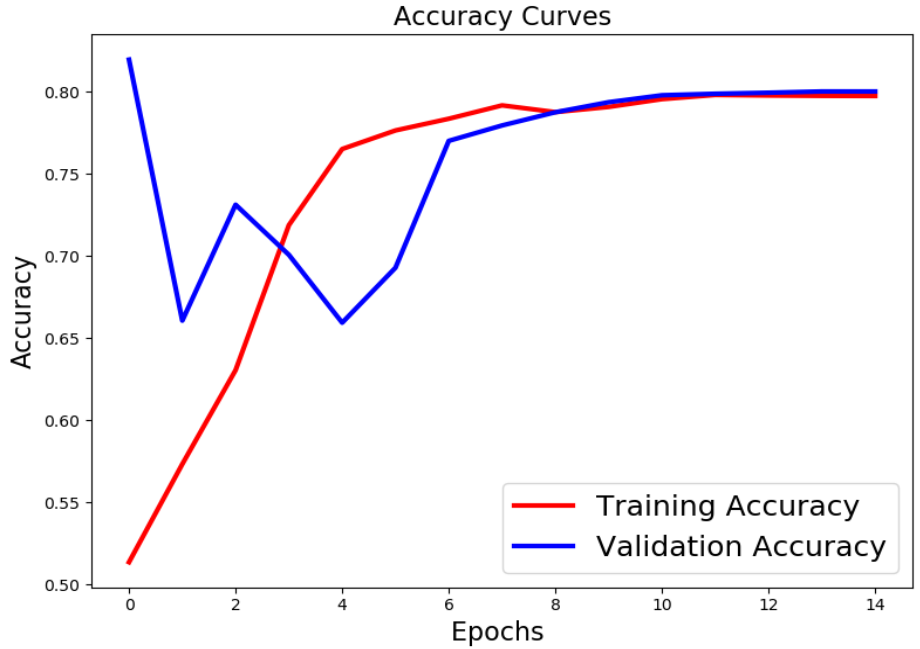
ROAD DETECTION

METHOD 2

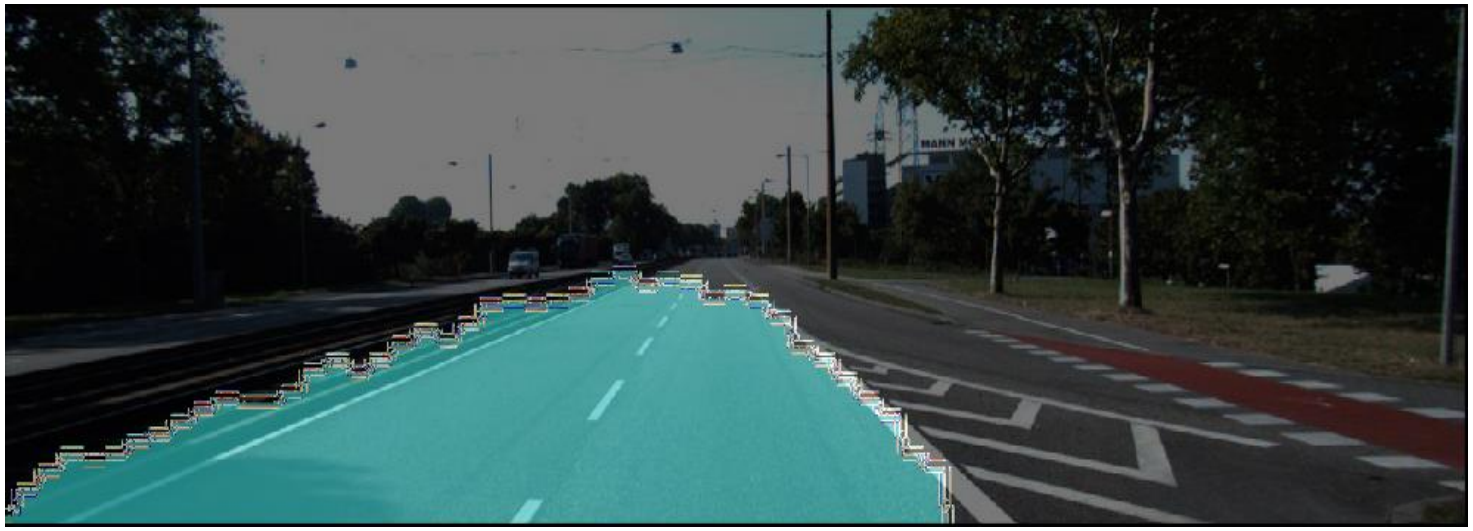
- U-NET Convolutional Network:
  - It has encoder and decoder
  - Takes in grayscale image and generate a ground truth mask

- Batch size = 20
- Data split = 80% training + 20% validation





CNN-LSTM:



U-NET:

COMPARISON

- CNN-LSTM :

long training time (3 hours) and is biased on the road location

- U-Net:

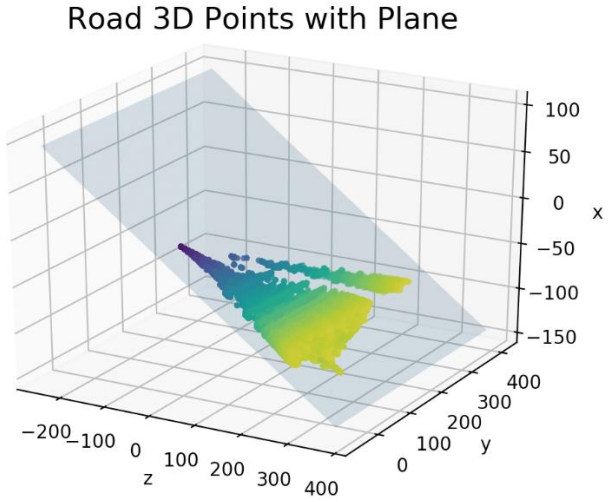
fast to train (10 minutes) but prediction is unstable, 80% accuracy

Why CNN?

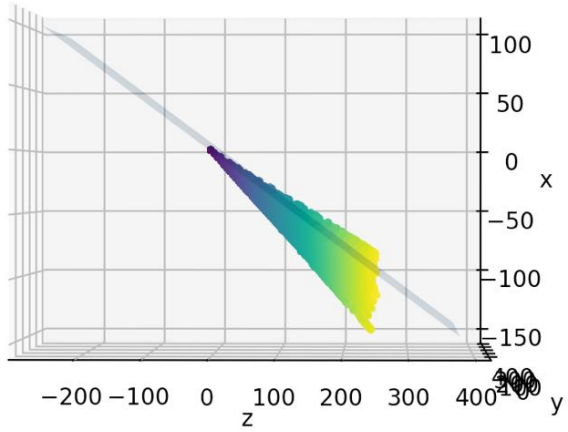
Old-school algorithms use rigid algorithms and require human intervention.

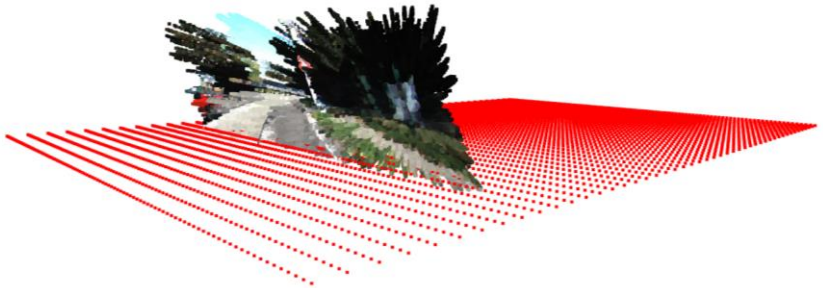
FIT GROUND LANE

- Disparity → Depth → 3D Location
  - Formula: $X = \frac{Z - x * Px}{f}, Y = \frac{Z - y * Py}{f}$
- Filter point cloud through road prediction mask
- RANSAC with 3000 iterations and 0.02 distance threshold
  - Each time fit a plane with build-in lstsq
  - Input 5 random chosen points
  - Keep the list of inliers index if the current plane produces the max number of inliers so far
- Fit the plane using lstsq with the best match
- Good: robust to outliers
- Bad: Not stable



Road 3D Points with Plane





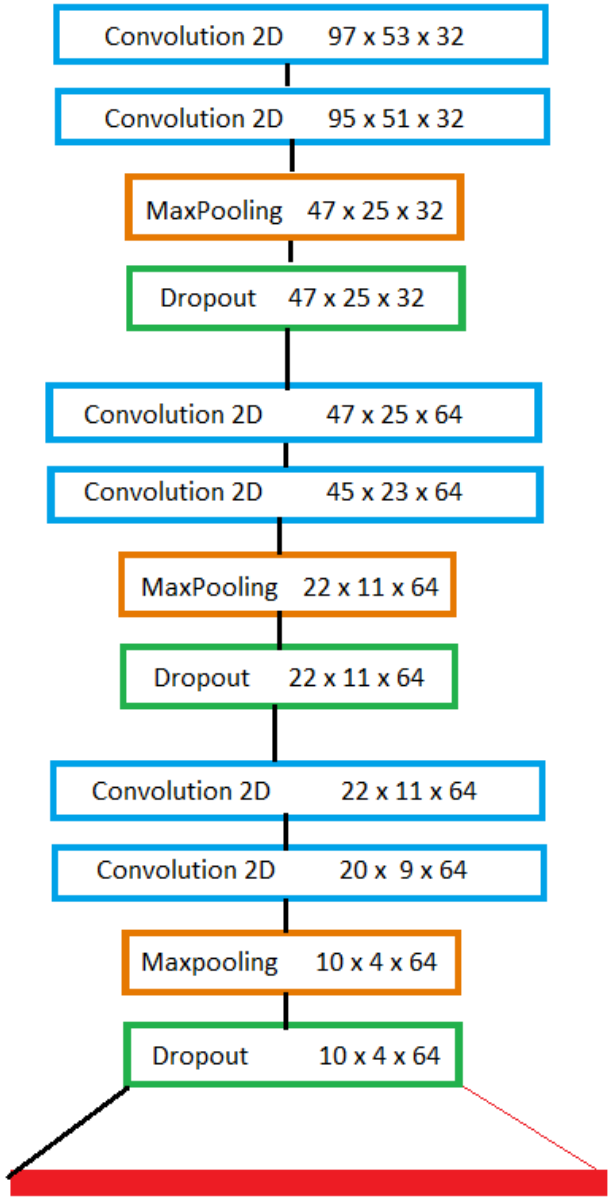
3D POINT CLOUD

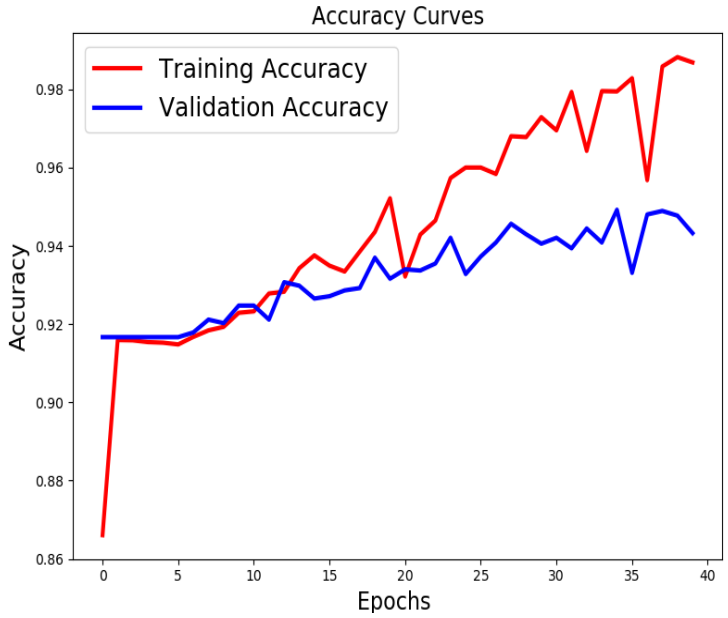
Plot 3D point cloud of all image pixels and the estimated ground plane with Open3d

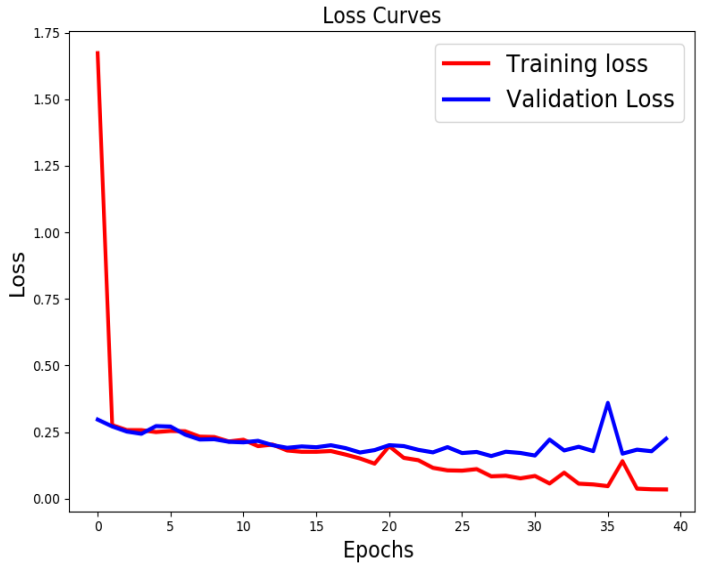
CAR DETECTION

- Did research for pre-trained model only for KITTI dataset

- Faster CNN from Torchvision pre-trained model



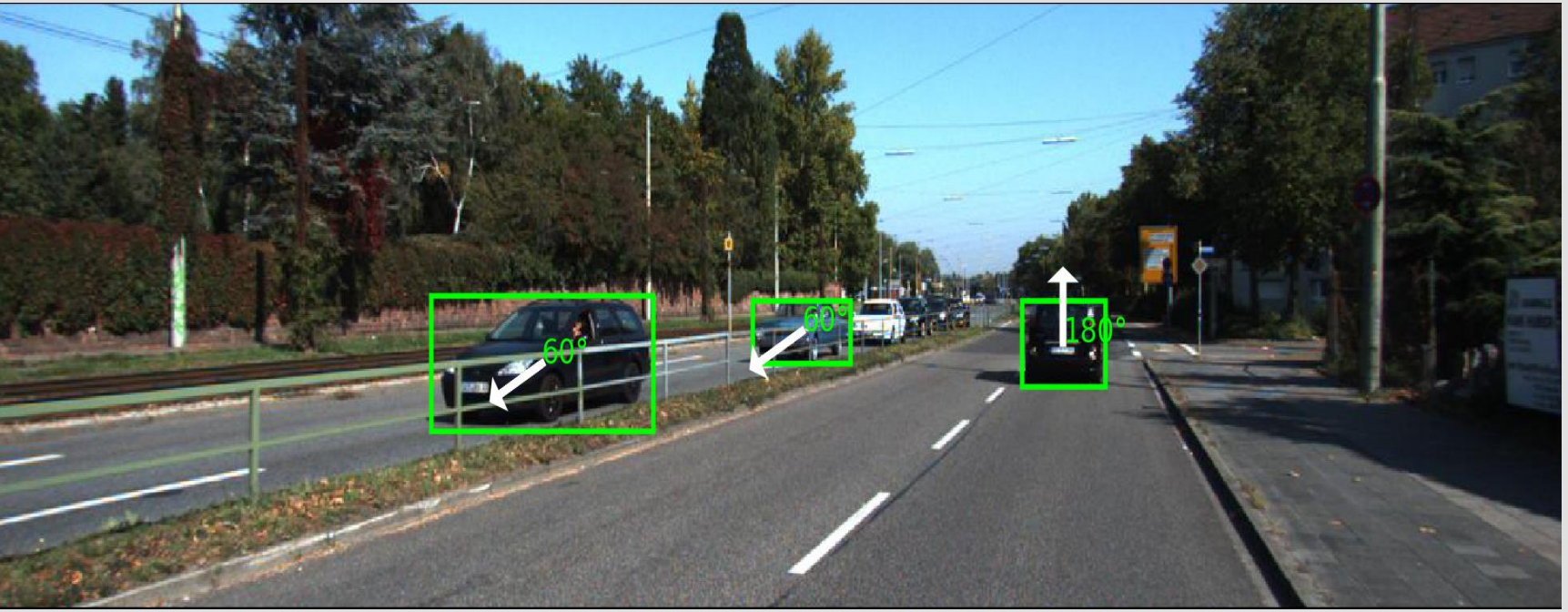




VIEWPOINT CLASSIFIER

- Using common CNN for image classification
- Validation accuracy is around 94%, training accuracy is 98%

VISUALIZATION



Using car detection to get image patches for cars, then feed in patches into pre-trained viewpoint model to get predicted angle.