Week 6
Weekly Presentation

Farzain Majeed

Last Presentation

2D Test Set



3D Test Set



Finished up 2D Experiments

- Trained on Udacity dataset (over 200k images, variety of scenarios, lots of driving in traffic, a lot of straight driving)

- Ended up getting *very* similar results on training set. Test set as well, but harder to judge.



Going 3D

- Liked results that a 3D CNN gave. Made more sense conceptually as well to use convolve along the temporal region as well.

- Original 2D CNN converted to 3D worked and was trained from scratch.

- Wanted to take it a step further by fine-tuning the C3D model trained on Sports 1M.

Fine-tuning C3D

- Pre trained model/weights built in Caffe. Spent a good deal of time converting it over to a workable format (Keras / TensorFlow).

- Spent more time actually learning what it meant to “fine-tune” something and the best practices.

- Keras doesn’t different learning rates for different layers yet.

- Added my own dense layer to the end of C3D to predict steering angles. Froze all other convolutional and dense layers. Trained. Saved weights.

Fine-tuning C3D

- Unfroze all layers in C3D (including my final dense layer). Loaded weights from previous model (where all layers except the last were frozen)

- I trained this with a slower learning rate. Results TBA. Bug somewhere, tough to find.

- Good news - I coded up the boiler plate and now I have a good idea of how fine tuning works now.

Observations on 2D/3D

- Still need to run some experiments where we regress upon steering angles, brake values, and throttle values.

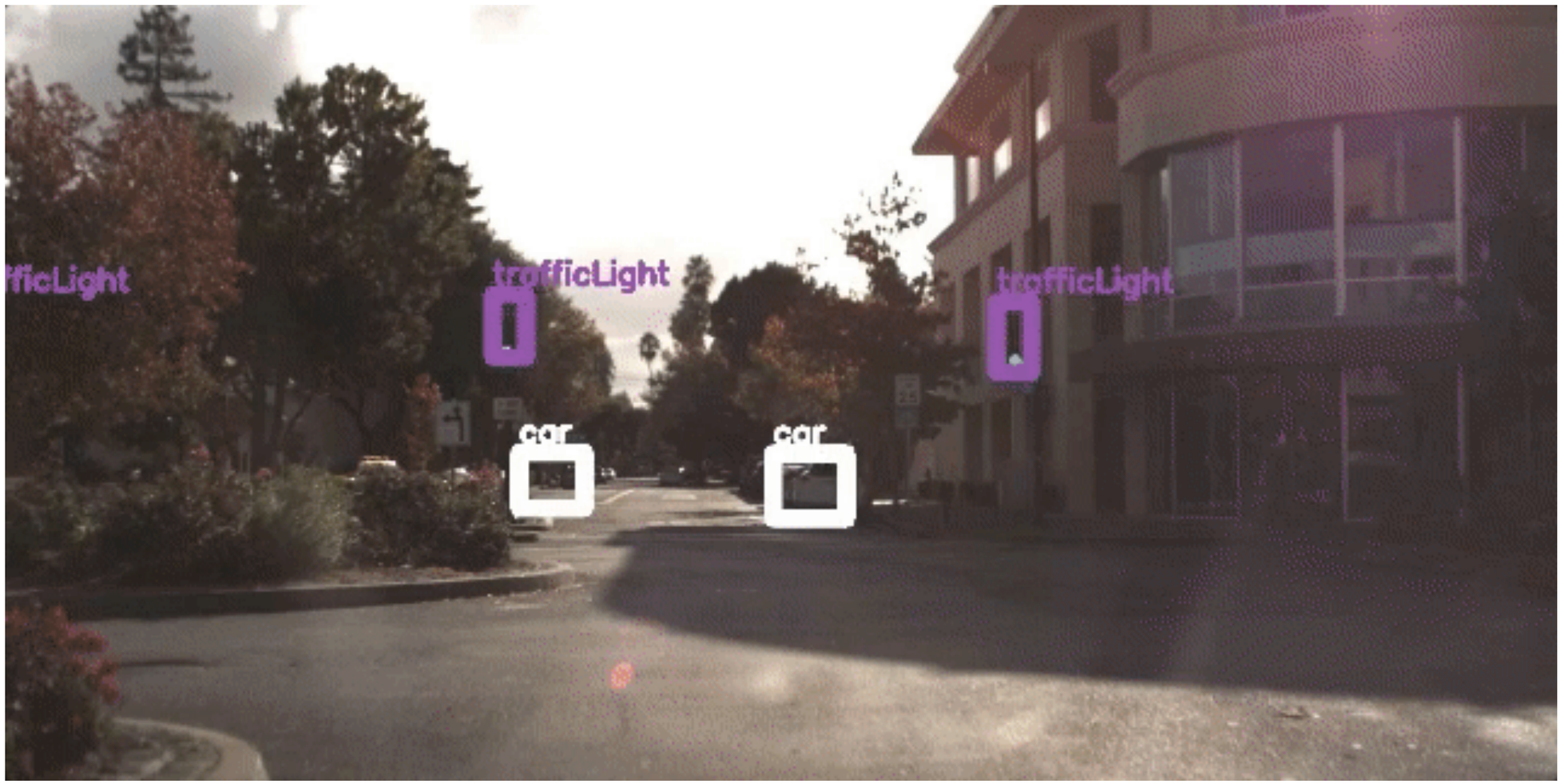
- I notice both CNNs sometimes predict to steer directly into another vehicle, causing a crash. Both types of models seem to have hard time paying “attention” to things like cars and traffic lights.

- This make some sense since we are only regressing upon steering angles for now and thats all the guidance the model has to figure out whats important.

- We want the model to pay explicit attention to certain things to see if performance increases.

New model

- Detect objects we care a lot about, for now - cars, pedestrians, and traffic lights. We can do this using something like the YOLO detector.



New model

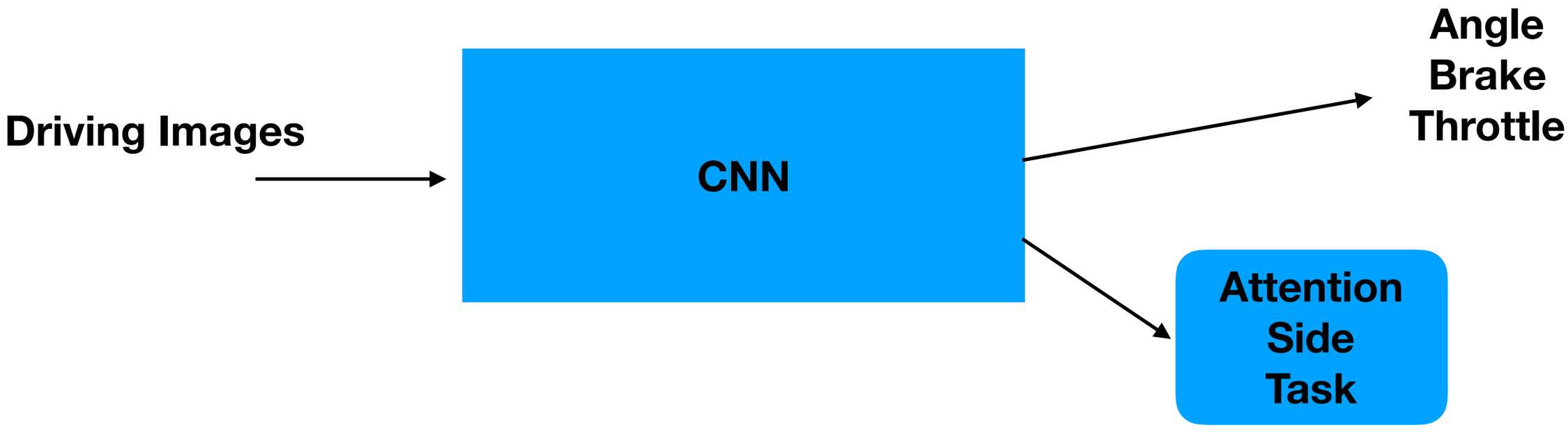
- YOLO gives us bounding boxes. Lets use these as the ground truth for our SIDE TASK where we are looking to teach the network to pay greater attention to certain things.





New model

Still working out many specifics with Kishan! ex. 2D vs 3D, LSTMs, etc



End