Semantic Segmentation

In this project I had no time to do the enhancements so no video. Anyway will continue trying as it is very interesting.

Build the Neural network

Does the project load the pretrained vgg model?

Yes, in lines 31-45 it loads the pretrained model and extracts the relevant layers.

Does the project learn the correct features from the images?

It builds all the layers in lines 68-96 but previously scales poll 3 and pool 4 before the 1x1 convolution and addition.

Does the project optimize the neural network?

I lines 116-125 computes the entropy loss and extracts the l2 values and adds it to the loss. Also generates the optimizer and the train operation.

Does the project train the neural network?

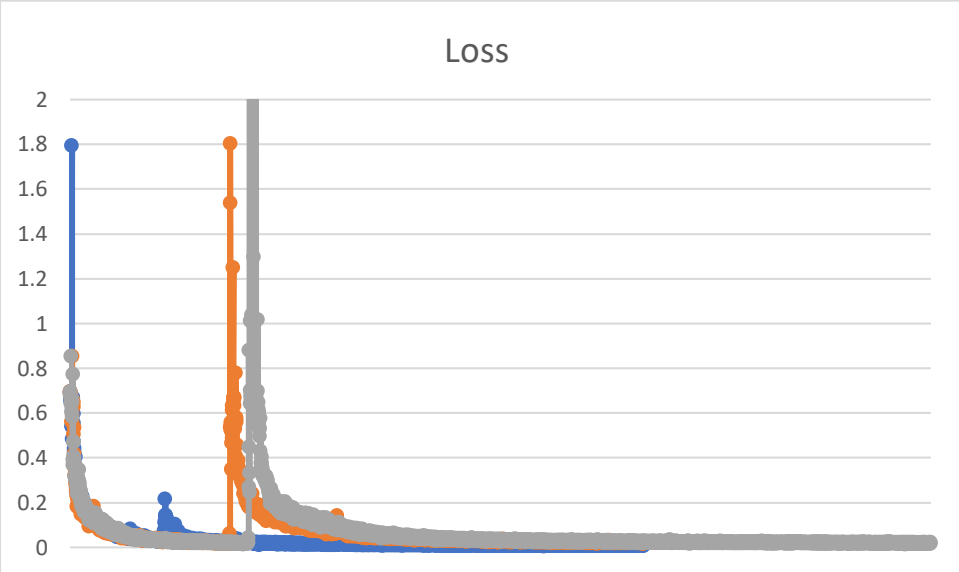
In lines 148-153 it trains the network and print the losses

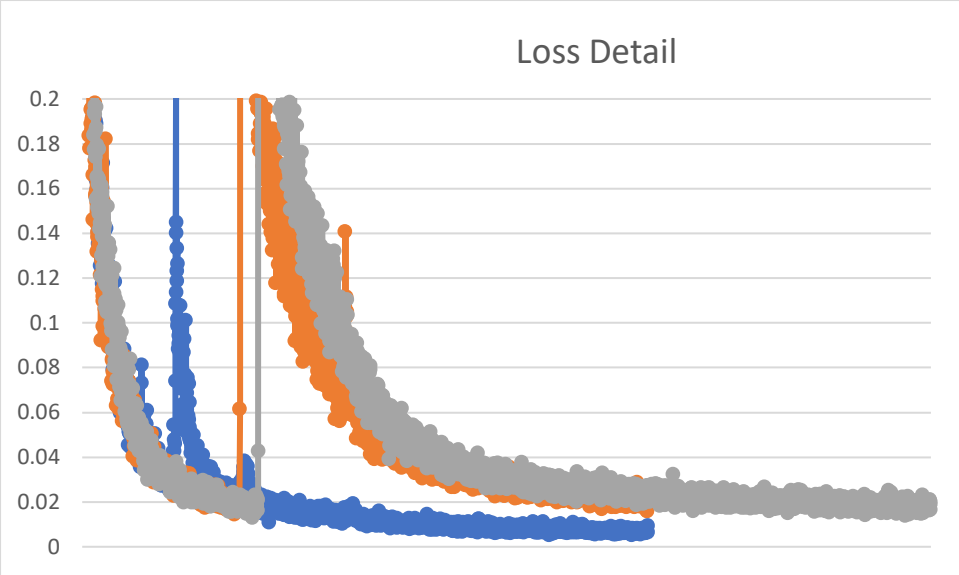
Neural network Training

Does the project train correctly?

The loss decreases over time as seen in following features. Anyway there is some artifact around batch 720 which generates an increase in the loss. I don't know the cause. Then it recovers and continues the decrease.

It happens mainly in the models with the l2 collection but also may be seen in one without.

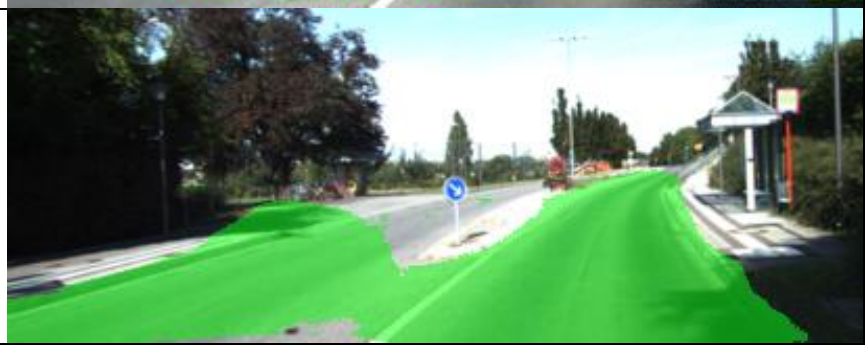


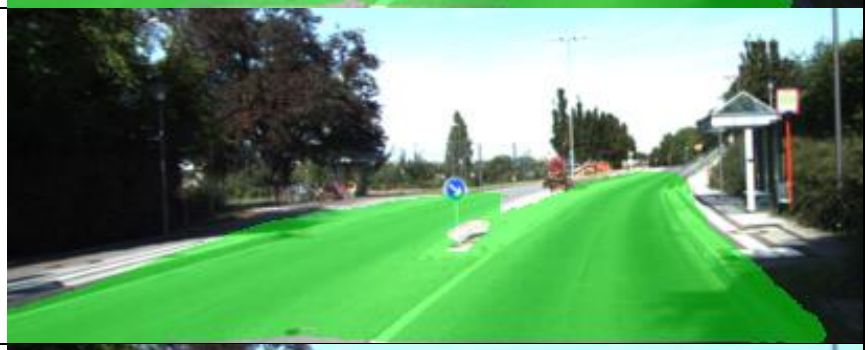
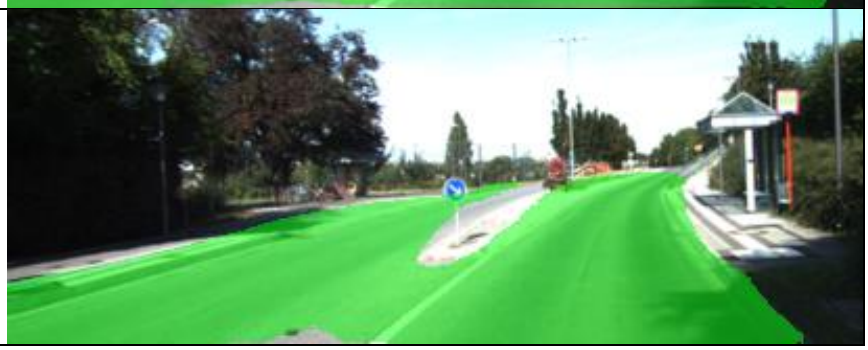


Does the project use reasonable hyperparameters?

First test was with 50 epochs. Result was quite bad. Then increased epochs to 200 and 300 with much better results.

Batch size began with 3 because the GPU wasn't able to process more. Then changed to an AWS p2.xlarge and augmented the batch size to 25.

Epochs	Reg_rate	Comments	Example
40			
80			

200		Layer 3-4 scaling	
200	0.003	Layer 3-4 scaling	
300	0.001	Layer 3-4 scaling	

Does the project correctly label the road?

Let's say that most of the times it does it reasonably but I should need much more work for a real case.

It sometimes gets with sidewalks and some "gray" road extensions.

Also the "border" around cars could be better.

I have made some small videos from all the images with um, umm and uu prefixes.

The third model without adding the collection and addition of the l2 losses sometimes is better but possibly because of overtraining.

Unfortunately I had no more time to try the video processing or the cityscapes database.