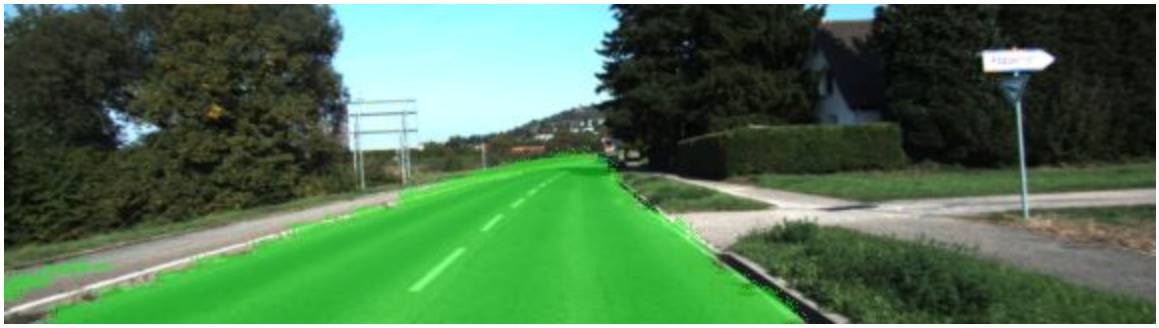
Term 3 Project 2 - Semantic Segmentation

The purpose of this project was to use Tensor Flow on road images and classify pixels as “road” or “not-road”. We used images from the Kitti Dataset (http://www.cvlibs.net/datasets/kitti/eval_road.php). We used a Fully Convolutional Network mimicking the VGG-16 architecture. We implemented a 1x1 convolutional layer to replace the fully connected layer. We also used upsampling, 1x1 convolutions to previous VGG layers and skip layers. Typically a fully connected network doesn’t maintain spatial data, so we used all these techniques to maintain the spatial data and output images.

For the GPU environment, I used a g3.4xlarge Amazon EC2 instance with the Udacity Deep Learning AMI. Using trial and error, I was able to determine hyperparameters that output the best images. For my model, I trained using a batch size of 16 over 15 epochs. For training, I set my learning rate to .0001. Also, an Adam optimizer was used to optimize the training and a cross entropy loss function was used to measure loss.

Below are some sample output images:



This project took a lot of effort to complete, especially with unforeseen complications with the AWS GPU. I’m happy with the final results, but it’s likely results could be improved with further hyperparameter tuning. Thanks for checking out my project!