Road Boundary Detection in Image via Machine Learning

Intermediate Report

Team Member: Dongbo Wang(dwang49), Yibo Jiang(yjiang71), Bangqi Wang(bwang34)

Introduction

Road boundary detection in image is actually a segmentation problem in computer vision. The purpose of road boundary detection is to find the area of road, or in other word, find the area of specific class in images. We did some research on this topic during the last several weeks. The most common computer vision method in recent years is deep learning with convolution neural network. Deep learning neural network is the complicated and optimized version of traditional neural network with multiply hidden layers. In our final project, we decide to test the traditional neural network and the deep learning neural network. We will test, analysis, and compare the existing neural network layouts. Then, we will optimize the existing neural network, or design our own neural network.

The project has two important parts, image feature extraction and deep learning neural network construction. The paragraphs above have introduced the method to preprocess the dataset. The preprocessed dataset will provide the labels and features for neural network to learn. The output will be significant different with different preprocess methods. We will test different preprocessed datasets with same neural network layout and find the method that provides the highest accuracy.

Image Feature Extraction

Normally, the deep learning neural network does not require image feature extraction. The image is feed to the neural network directly and the feature extraction is somehow finished by the layers and neurons in the neural network. However, this is more practical in classification problems, not the segmentation problem we are dealing with. We think feeding the whole image directly to the neural network is not efficient enough. We are asking the neural network to understand the segmentation out of nowhere, which means we either need more complex neural networks or more resources like data and computation time to learn a good model. However, since we want to detect the road boundary in the image, we do know more information about the image. The image is not just some pixel values, there are a lot of information that could be extracted from the image, like the histogram, gradient magnitude and orientation, depth and even raw segmentation. These are all good potential features that could feed to the neural networks. We believe this is a good practice because we are using our knowledge to help elevate some difficulty for the neural network to learn.

We expect that we could either increase the accuracy of the road boundary or decrease the training time needed before the network converges. We are currently implementing different feature extraction algorithms.

Deep Learning Neural Network

Deep learning neural network can work on supervised and unsupervised learning. For this project, we will use deep learning as supervised learning. We will preprocess the images, extract features, and label road area. The deep learning neural network will learn the rule from labels and predict the road based on the road labeled in input dataset. For multiple hidden layers in deep learning neural network, each hidden layer goes into the next layer with non-linear combination of the layers below it. In other word, each layer is an optimally weighted, non-linear combination of the layer below it. Therefore, deep learning has great ability to extract the intrinsic regularity from dataset and it is a good method to predict the road.

We have read several papers about segmentation with deep learning. In the paper about semantic segmentation [1], people adapt contemporary classification network, such as, AlexNet, VGGNet, and GoogleNet, into fully convolutional network transfer the learned representation to segmentation task. This neural network can make pixel-wise prediction for different classes based on the input with arbitrary size. We believe this neural network also work for road detection if we train this network with our road dataset. The layout will be shown below.

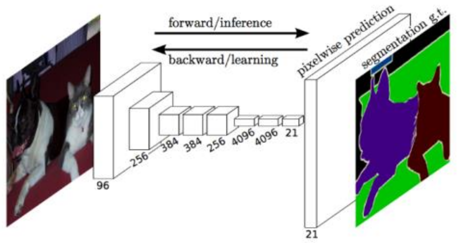

Figure 1. Fully convolutional networks can efficiently learn to make dense predictions for per-pixel tasks like semantic segmentation.

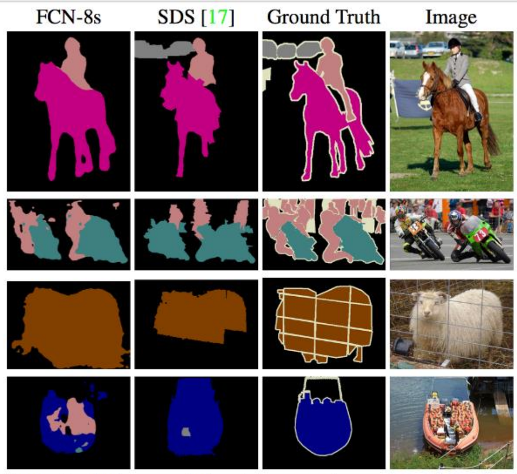

Figure 6. Fully convolutional segmentation nets produce state-of-the-art performance on PASCAL. The left column shows the output of our highest performing net, FCN-8s. The second shows the segmentations produced by the previous state-of-the-art system by Hariharan *et al.* [17]. Notice the fine structures recovered (first row), ability to separate closely interacting objects (second row), and robustness to occluders (third row). The fourth row shows a failure case: the net sees lifejackets in a boat as people.

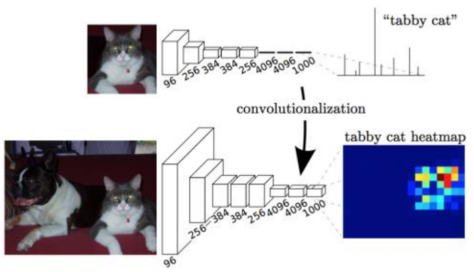

Figure 2. Transforming fully connected layers into convolution layers enables a classification net to output a heatmap. Adding layers and a spatial loss (as in Figure 1) produces an efficient machine for end-to-end dense learning.