FastDraw: Addressing the Long Tail of Lane Detection by Adapting a Sequential Prediction Network

Jonah Philion
ISEE.AI
jonahphilion@isee.ai

Abstract

The search for predictive models that generalize to the long tail of sensor inputs is the central difficulty when developing data-driven models for autonomous vehicles. In this paper, we use lane detection to study modeling and training techniques that yield better performance on real world test drives. On the modeling side, we introduce a novel fully convolutional model of lane detection that learns to decode lane structures instead of delegating structure inference to post-processing. In contrast to previous works, our convolutional decoder is able to represent an arbitrary number of lanes per image, preserves the polyline representation of lanes without reducing lanes to polynomials, and draws lanes iteratively without requiring the computational and temporal complexity of recurrent neural networks. Because our model includes an estimate of the joint distribution of neighboring pixels belonging to the same lane, our formulation includes a natural and computationally cheap definition of uncertainty. On the training side, we demonstrate a simple yet effective approach to adapt the model to new environments using unsupervised style transfer. By training FastDraw to make predictions of lane structure that are invariant to low-level stylistic differences between images, we achieve strong performance at test time in weather and lighting conditions that deviate substantially from those of the annotated datasets that are publicly available. We quantitatively evaluate our approach on the CVPR 2017 Tusimple lane marking challenge, difficult CULane datasets [8], and a small labeled dataset of our own and achieve competitive accuracy while running at 90 FPS.

1. Introduction

Previous models of lane detection generally follow the following three-step template. First, the likelihood that each pixel is part of a lane is estimated. Second, pixels that clear a certain threshold probability p_{min} of being part of a lane are collected. Lastly, these pixels are clustered, for instance with RANSAC, into individual lanes.

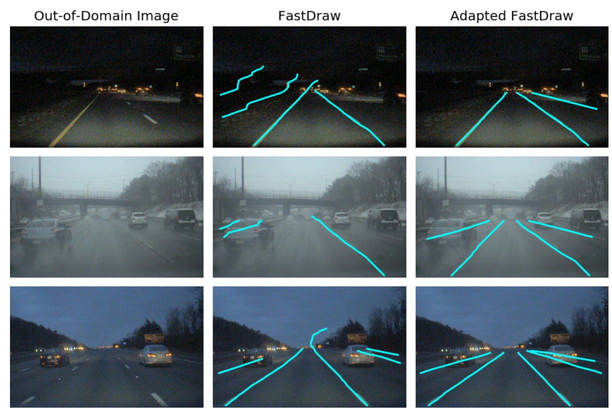

Figure 1. Best viewed in color. We train a novel convolutional lane detection network on a public dataset of labeled sunny California highways. Deploying the model in conditions far from the training set distribution (left) leads to poor performance (middle). Leveraging unsupervised style transfer to train FastDraw to be invariant to low-level texture differences leads to robust lane detection (right).

Because the second and third steps in which road structure is inferred from a point cloud of candidate pixels are in general not differentiable, the performance of models of lane detection that follow this template is limited by the performance of the initial segmentation. We propose a new approach to lane detection in which the network performs the bulk of the decoding, thereby eliminating the need for hyper-parameters in post-processing. Our model "draws" lanes in the sense that the network is trained to predict the local lane shape at each pixel. At test time, we decode the global lane by following the local contours as predicted by the CNN.

A variety of applications benefit from robust lane detection algorithms that can perform in the wild. If the detector is iterative, the detector can be used as an interactive annotation tool which can be used to decrease the cost of building high definition maps [6, 1]. For level 5 systems that depend on high definition maps, online lane detection is a useful lo-