Abstract

This thesis is written based on two main topics: Scene semantic recognition and road lane marks recognition. The thesis first reviews the studies of general scene understanding or recognition method using semantic segmentation approaches, and then focuses on a more specific topic of analyzing road scene images for detecting road lane marks.

Semantic segmentation has been a popular topic in the field of computer vision research. The main purpose of semantic segmentation is to label pixels of interest in an image with corresponding categories of the objects. This thesis mainly focuses on scene recognition, a branch of semantic segmentation which takes more contextual information into consideration. This thesis presents an experimental study in which a multi-task method for scene recognition is proposed. In this method, edge information is used in enhancing recognition performance. A network which outputs both edge detection map and pixel-wise segmentation is designed. The network is based on FCN and the prediction branches of the two outputs are parallel. Each branch uses multi-scale features concatenation as the image representations. The method expects that the information from edge detection could contribute to the ability of extracting image features for pixel-wise segmentation.

Modern approaches on multiple road lane marks detection are facing several problems. First, insufficient database make related solutions with machine learning technique difficult to train a robust model for application; second, current researches focus on single lane marks detection, which pays less attention to entire roads' condition. To solve the problems, a database with proper ground truth of marks' label set is constructed and a method is developed for detecting and classifying road lane marks of entire roads with Extreme Learning Machines (ELM). The implementation result shows promising performance and further improvement could be expected.