ABSTRACT

This thesis mainly studies the two key technologies for autonomous ground vehicles, namely scene perception and local path planning. Among them, scene perception is a very challenging task since the input data has a high dimensionality and contains a lot of noise. The current scene perception algorithms mainly utilize computer vision and pattern recognition techniques. By using machine learning, especially deep learning, one can dramatically improve the performance of scene perception. However, the three issues of (a) large data demand, (b) large demand for manual labels, and (c) lack of interpretability in deep learning severely limit its application in scene perception. To this end, the thesis proposes three different deep learning architectures for three key tasks in scene perception, respectively, which can better alleviate the above three issues while ensuring good performance for scene perception. In order to further improve the adaptability of autonomous driving to the scene, the thesis also conducts an in-depth study on the local path planning algorithm. The main works and novelties of this thesis are as follows:

(i) For dynamic scene prediction, a Tensorized Recurrent Neural Network is proposed. The prediction Mean Squared Error (currently the most predominant video prediction performance measure) of the model on the KTH, UCF101 and KITTI datasets are respectively reduced by 22.11%, 5.52%, and 13.57% based on the best existing algorithm (MCnet deep neural network). Compared to MCnet, the model also has the following advantages: first, the performance can be efficiently improved without introducing additional parameters and computation time, and its parameter number is only 74.47% of MCnet, reducing the amount of data required to solve the parameters and thus alleviating Issue (a); second, the Memory Cell Convolution is introduced, and by visually analyzing the information flow in memory cells, the interpretability is increased, alleviating Issue (c).

(ii) For unsupervised multi-object detection, a Memory-Based Recurrent Attention Network is proposed. The Average Precision (currently the most predominant detection performance measure) of the model reaches 94.78% of the best existing algorithm (RRC deep neural network) on the DukeMTMC dataset, and reaches 95.16% of the best existing algorithm (CRAFT deep neural network) on the TUD dataset. Compared to RRC and CRAFT, the advantages of the model are: first, it makes full use

of the parameter sharing mechanism, so that its parameter number is only 19.14% of RRC, and 12.58% of CRAFT, reducing the amount of data required to solve the parameters and thus alleviating Issue (a); second, the parameters are trained by minimizing the image reconstruction error, thus no manual labels are required, overcoming Issue (b); third, input memory and attention mechanisms are introduced, and by visually analyzing these mechanisms, the interpretability is increased, alleviating Issue (c).

(iii) For unsupervised multi-object tracking, a Reprioritized Recurrent Attention Network is proposed. The Identification F-measure (currently the most predominant tracking performance measure) of the model reaches 92.38% of the best existing algorithm (DeepCC deep neural network) on the DukeMTMC dataset, and reaches 93.39% of the best existing algorithm (AM deep neural network) on the TUD dataset. Compared to DeepCC and AM, the advantages of the model are: first, it makes full use of the parameter sharing mechanism, so that its parameter number is only 22.07% of DeepCC, and 16.33% of AM, reducing the amount of data required to solve the parameters and thus alleviating Issue (a); second, the parameters are trained by minimizing the video reconstruction error, thus no manual labels are required, overcoming Issue (b); third, mechanisms of input memory, attention, reprioritization, etcetera are introduced, and by visually analyzing these mechanisms, the interpretability is increased, alleviating Issue (c).

(iv) A hierarchical local path planning algorithm based on Dual Process Theory is proposed. The algorithm consists of multiple path planning layers and can automatically select an appropriate planning layer according to the complexity of the environment; through the task-oriented planning layer scheduling mechanism, the algorithm is applicable to local path planning in a variety of specified tasks, including the urban road (located in Lugu Industrial Park in Changsha), the highway (Changsha City Ring Expressway, with a total length of about 85 kilometers, and a total experimental distance of about 1000 kilometers), the country road (located in an island of Xiangjiang in Changsha, the island covers an area of about 1.67 square kilometers), and the off-road road (located at an armored vehicle test base); the planning success rate of the algorithm reaches 100.00% in the above urban road task, and reaches 99.99% in the above off-road road task, demonstrating its adaptability and effectiveness.

The research results of this thesis have contributed to the successful completion of the National Natural Science Foundation of China (NSFC) Major Research Projects' Integration Project "Key Techniques and Integrated Verification Platform for Autonomous Ground Vehicles" (belonging to the NSFC Major Projects), where the local path planning module of this platform is mainly implemented by the hierarchical local path planning algorithm proposed in this thesis.

Key Words: Scene Perception, Autonomous Driving, Recurrent Neural Network, Deep Learning, Scene Prediction, Object Detection, Object Tracking, Local Path Planning