Application of Deep Learning in Automatic Driving

Wei Su, Xiaoyan Chen*

Tianjin University of Science and Technology,300222, China;

E-mail: 965756192@qq.com

www.tust.edu.cn

Abstract

With the continuous development of science and technology, the field of artificial intelligence has become a research hotspot, especially deep learning, which has attracted much attention from all walks of life. Starting with the automatic driving solution, this paper mainly expounds the important role and technical route of the application of deep learning, and finally looks forward to the development direction and application of automatic driving technology based on deep learning.

Keywords: Automatic Driving, Deep Learning

1. Introduction

Automatic driving is developing rapidly in the direction of intelligence. As a cutting-edge branch of artificial intelligence, deep learning has promoted the explosive growth in the field of artificial intelligence, and has been widely used in the perception, decision-making, control and other related fields of automatic driving. The current research focus is computer vision. The research hotspot is the application of convolutional neural network to robot objects. Deep learning is also used to assist decision-making. The integration of deep learning and autopilot has a good prospect, which will promote the development of autopilot to a higher level of intelligence.

Deep learning is a kind of machine learning. Its algorithm is based on representation learning of data, which can discover the complex structure in big data, and use back propagation to guide the machine how to calculate the representation from the previous layer network, so as to change the internal parameters of each layer[1]. The purpose of deep learning is to make the unmanned vehicle have the same analysis and learning ability as people, because it can enhance the perception, decision-making and control ability of unmanned vehicle, it has a wide application prospect in image recognition, target recognition, path planning, human-computer interaction and so on.

2. Development status of automatic driving

2.1. *Automatic driving level division standard*

The automatic driving level division standard formulated by the SAE (Society of Automotive Engineers) is shown in Fig.1.

Level 1: system is mainly used to improve driving safety. Such as ABS(antilock brake system).

Level 2: partial automatic driving. The vehicle provides driving for multiple operations in steering wheel, acceleration and deceleration, and the human driver is responsible for other driving actions.

Level 3: conditional automatic driving. Most driving operations are completed by the vehicle, and human drivers need to pay attention in case of emergency.
Level 4: highly automatic driving. cars can drive automatically in special scenes such as highways, but cars still need human intervention under complex road conditions.
Level 5: fully automatic. human beings become passengers completely

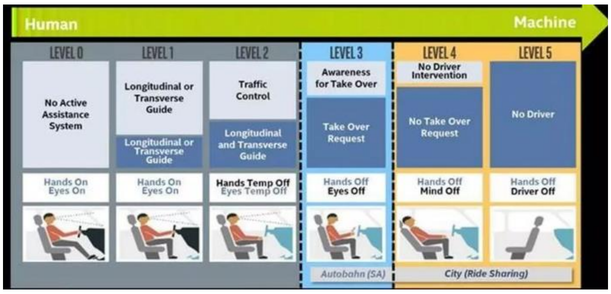

Fig. 1. The automatic driving level division standard

2.2. *Key technologies of automatic driving*

2.2.1 Environmental perception

Commonly people use sensors for vehicle environment sensing include lidar, millimeter wave radar, camera, etc. Cameras include monocular camera, binocular camera, and look around camera. Monocular camera has simple structure and mature algorithm, but its field of view depends on lens and ranging accuracy is low. The binocular camera has higher ranging accuracy. The look around camera can have a field of view of 360 degrees, but the sensing range is only 5 to 10 m, and the edge of the photo is seriously deformed.

In order to solve the problem of inaccurate camera ranging, engineers introduced lidar sensors. The lidar sensor emits a laser beam externally. After the laser beam meets the object, it returns to the laser receiver through diffuse reflection. The computer calculates the distance between the transmitter and the object according to the time difference between sending and receiving signals. The fast-rotating laser sensor can provide millions of data points per second, which can create 3D maps of surrounding objects and environment. Lidar is not affected by lighting conditions, but it is vulnerable to smoke or bad weather, and the cost of lidar is very high.

Millimeter wave radar is affordable, less affected by smoke, strong penetration and small volume. Its detection range can reach 200m. However, the data stability is poor and the algorithm is more complex.

2.2.2 High precision position

High precision position provides global path planning for autonomous vehicles, and obstacle avoidance planning based on perception results, also known as real-time navigation planning. At present, the two major challenges that positioning technology still faces are covering blind areas and high cost.

2.2.3 Collaborative decision making

The decision-making needs to integrate the current scene, path planning, user needs and other information, adapt to the real-time situation as much as possible, and make friendly and efficient decisions on the premise of ensuring driving safety[2].

3. Deep learning application on automatic driving

3.1. *Lane line detection*

Many scholars began to use the methods of deep learning to solve the problem of lane detection.

In 2018, Davy Neven[3] and others proposed a lane line detection neural network model LaneNet+H-Net, which can carry out end-to-end lane line detection. The model structure mainly including two network models LaneNet and H-Net, which is shown in Fig.2. LaneNet is a multi task model. One branch is used for binary semantic segmentation of lane lines to distinguish whether they are

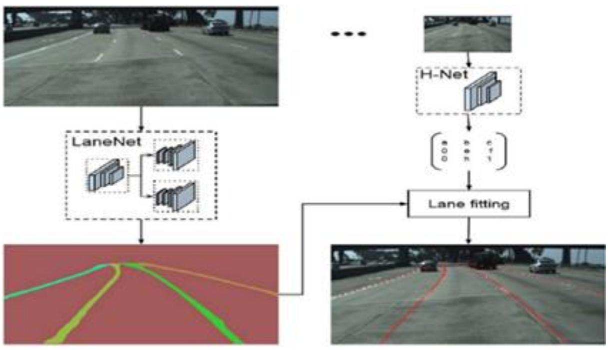

Fig. 2. LaneNet + H-Net network structure[3]

lane lines or background. The other branch uses the lane line division output of the first branch as input. During training, the segmented lane line mask is clustered by learning the designed clustering loss function, and the two branches are combined. LaneNet + H-Net network realizes end-to-end lane line instance segmentation.

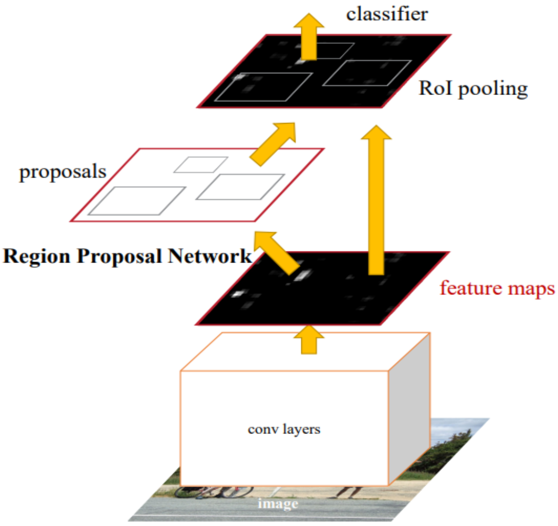

Fig. 3. Faster RCNN network structure[5]

3.2. *Pedestrian detection*

RCNN(Region CNN) algorithm is a pioneering work for target detection using deep learning. It is designed by Ross Girshick, a famous fair programmer under Facebook. It is also a great milestone in applying the traditional CNN method to target detection. Based on CNN's excellent feature extraction and performance classification, the transformation of target detection problem is realized by region proposal method.
However, RCNN has fatal shortcomings. Repeated feature extraction leads to a waste of computing resources. The later fast RCNN is improved on this basis to speed up the algorithm and no additional storage[4].

Faster RCNN network structure is shown in Fig.3. It integrates feature extraction, target extraction, candidate box and classifier into a network, which is completed by deep neural network and runs on GPU, which greatly improves the operation efficiency. Compared with fast RCNN, the comprehensive performance of RCNN is greatly improved. It is one of the mainstream algorithms of target detection. The faster RCNN network framework is shown in the Fig.3.

3.3. *Multisensor fusion*

Automatic driving based on deep learning has higher requirements for sensor data acquisition. As the data source of unmanned vehicle, various sensors are of great importance. However, as the input of the deep learning model, any sensor has its own shortcomings. So the

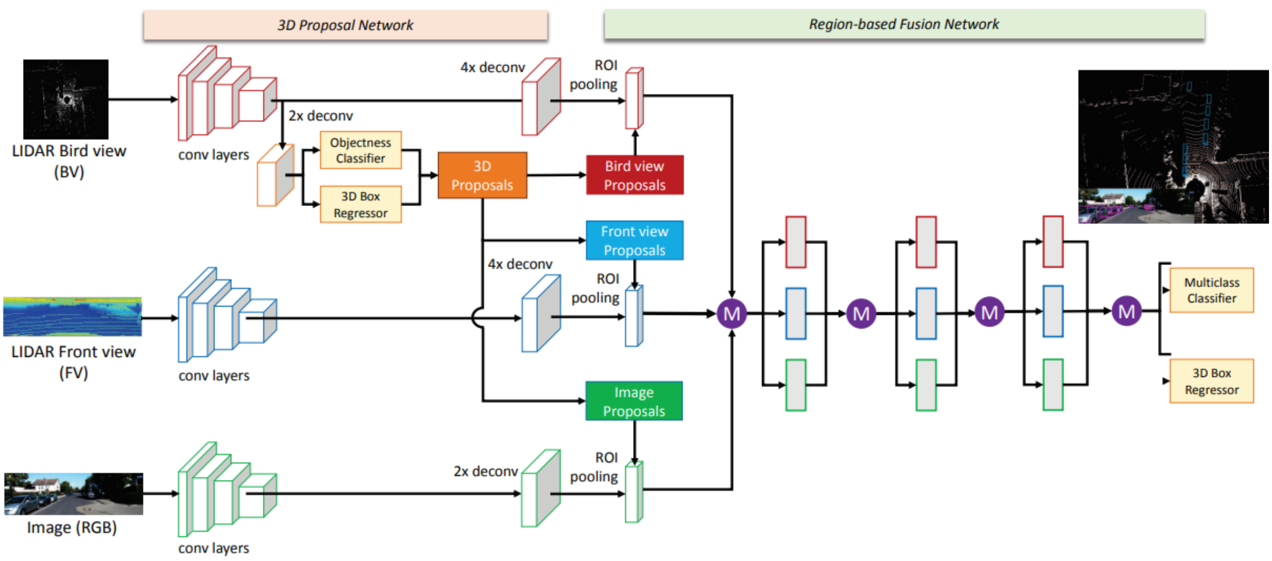

Fig. 4. MV3D object detection network structure[8]

multi-sensor fusion technology is of great significance for the accurate perception and cognition of autonomous vehicles.

At present, the target recognition scheme of multi-sensor fusion mostly integrates the traditional camera image and laser point cloud 2D processing scheme into one network, such as VeloFCN[6], Vote3D[7], etc. X Chen et al. proposed MV3D (multi view 3D object detection network) to fuse LIDAR point cloud data with RGB image information[8]. MV3D is used as the perceptual fusion framework. The 2D processing scheme of laser point cloud data represents the 3D point cloud information with the front view and aerial view of laser point cloud, and is fused with RGB image to predict the directional 3D boundary box[9]. The network structure is shown in Fig.4.

4. Conclusions

At present, deep learning shows a broad development prospect in the field of automatic driving. It can operate the original sensor data without manual assistance, infer the key features, and greatly shorten the preliminary engineering time. At the same time, it is also good at fusing multi-sensor data to capture the relationship between the original data, which is helpful for the autonomous vehicle to work in a complex environment. At present, the research focus of in-depth learning is computer vision, which will develop towards a higher level of cognition in the future. The purpose is to make autonomous vehicles reach the level of level 5 and realize real autonomous driving.

References

1. Yann LeCun, Yoshua Bengio , Geoffrey Hinton, Deep learning, *Nature*, 2015,521(7553):pp.436–444.
2. Peng Guo, Xueyi Wu, Hui Rong , et al, Local path planning of driverless cars based on cost function. *China Journal of Highway and Transport*, 2019,32(06):pp.79-85.
3. Davy Neven; Bert De Brabandere; Stamatios Georgoulis, et al.,Towards End-to-End Lane Detection: an Instance Segmentation Approach,*2018 IEEE Intelligent Vehicles Symposium (IV) IEEE*, 2018, pp. 286-291.
4. A. Ullah, H. Xie, M. O. Farooq, et al., Pedestrian Detection in Infrared Images Using Fast RCNN, *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*,2018:pp.1-6.
5. S. Ren, K. He, R. Girshick, et al.,Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017,39(6):pp.1137-1149.
6. M. Engelcke, D. Rao, D. Z. Wang, et al. Vote3Deep: Fast Object Detection in 3D Point Clouds Using Efficient Convolutional Neural Networks, *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017,pp. 1355-1361.
7. J. Lee, S. Walsh, A. Harakeh, et al.,Leveraging Pre-Trained 3D Object Detection Models For Fast Ground Truth Generation, *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*,2018,pp. 2504-2510.
8. X. Chen, H. Ma, J. Wan, et al., Multi-view 3D Object Detection Network for Autonomous Driving,*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6526-6534,
9. C. R. Qi, W. Liu, C. Wu, et al., Frustum PointNets for 3D Object Detection from RGB-D Data, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 918-927.

Authors Introduction

Mr. Wei Su

He received his bachelor's degree from the college of computer and Communication Engineering of Zhengzhou University of Light Industry in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Prof. Xiaoyan Chen

She is professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5). She had been in RPI, USA with Dr. Johnathon from Sep.2009 to Feb.2010 and in Kent, UK with Yong Yan from Sep-Dec.2012. She has researched electrical impedance tomography technology in monitoring lung ventilation for many years. She is in charge of the TUST-EIT lab and guides young researchers and graduate students to improve the electrical data acquisition hardware platform, to study the traditional and novel reconstruction algorithms with the prior structural information. Recently, her research team is focus on the novel methods through deep learning network models.