# Research on the Application of Deep Learning in Automatic Vehicle

Jiarui Bai[1, *]

[1]School of electrical and information engineering, Tianjin University. Tianjin 300072 China

**ABSTRACT:** In recent years, various kinds of applications of artificial intelligence technology to the automatic driving vehicle were widely reported by the media, also aroused public interest. According to a survey by the Ministry of Public Security, there are more than 500 million motor vehicles in China in 2021 and more than 200 thousand traffic accidents occurred in 2021, which is a great challenge to traffic safety protection. Human reaction time, which consists of perception time and judgment time, is too long to take measures. Besides, human drivers may find it difficult to even make the right response in some complex road conditions. To improve this situation, the Intelligent Transportation System (ITS) may transform the traditional passive safety that takes protective measures after accidents into active safety that focuses on prevention. Deep learning network receives information through the vehicle-loaded sensor then makes judgments by its computing unit and prompts drivers the possible danger. This article reviewed the performance of different deep learning models in autonomous driving vehicles.

## 1. INTRODUCTION

With the development of machine learning, automatic driving has gradually come into people's vision and has been applied in specific environments like mine cave in recent years. In any case, the current ITS is only certified to a level three, which means it can only be used in a few specific areas. Traditional automatic driving vehicles are limited by road conditions, the number of vehicles, etc. So it is difficult to ensure service quality. And as is widely acknowledged, auto-driving should not replace a human driver.

So the main goal of the system is to prevent the occurrence of accidents by analyzing road conditions information and helping the driver to make judgments. It perceives road conditions information outside the vehicle through the sensor. Compared with the previous machine learning model, deep learning does not require too many sensors and chips that meet the specified requirements of the vehicle. Through a lot of training, deep learning can perceive the external environment of the vehicle in a more accurate way, alleviate the demand for computing resources, and respond faster. Using a deep learning model to detect a possible accident and provide the information to the driver. In this way, drivers can make their judgment easier and reduce the possibility of an accident.

The key to this design is how to build a deep learning model and there are many achievements in this respect. The application of deep learning in autopilot is summarized in detail, including the problem of target detection in autopilot, the advantages and disadvantages of existing methods are discussed, and the problems and challenges to be solved in the existing work are have prospected.

## 2. THE DEVELOPMENT STATUS OF DEEP LEARNING

Compared with traditional machine learning methods, deep learning has powerful information extraction and processing ability, and has greater demand for computing resources. Many deep learning models can be applied to automatic driving technology. This section mainly introduced some typical deep learning models widely used in automatic driving.

### 2.1 Restricted Boltzmann machines

Restricted Boltzmann machines (RBM) [1] is a kind of stochastic neural network model with a two-layer structure, symmetrical connection and no self feedback, with the full connection between layers and no connection within layers. The Boltzmann machine is fully connected. Let's compare here:

*Corresponding author. Email: 3019234093@tju.edu.cn.
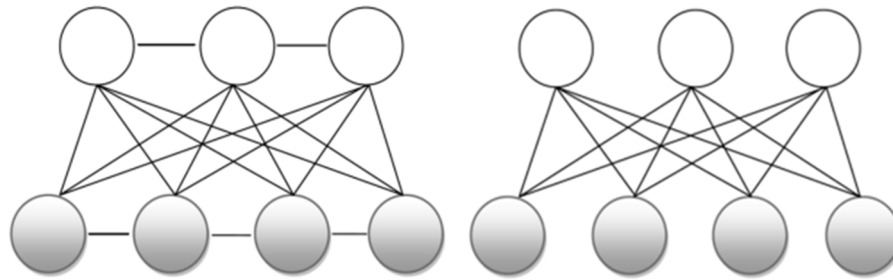
**Figure 1** Boltzmann machines and Restricted Boltzmann machines

The Boltzmann machine can learn internal complex representation and can provide a good solution for the application of signal processing in the process of vehicle networking. In addition, after multiple restricted Boltzmann machine layers are stacked, a deep confidence network composed of visible layers and multiple hidden layers can be formed, which is widely used in fault and anomaly detection on the Internet of vehicles.

## 2.2 Deep reinforcement learning

Deep reinforcement learning focuses more on reinforcement learning. Different from traditional reinforcement learning, deep reinforcement learning has a strong representation ability of approximation value function or direct strategy, and it uses a deep neural network to represent strategy. Deep reinforcement learning algorithms can be divided into two categories: value based model and strategy gradient based model. With the powerful representation ability of the deep neural network, the value function or strategy is fitted to solve a series of state behavior space problems, and then complete a series of path selection and vehicle control decisions in the field of automatic driving [2].

## 2.3 Deep forest

In order to obtain better service performance, a neural network needs more parameters. A large number of parameters means higher requirements for the storage capacity of the edge server. Therefore, the training of neural network usually requires researchers to spend a lot of energy on fine-tuning the super parameters. Zhou Zhihua and others proposed integration of the traditional tree based methods in breadth and depth in deep forest GC forest, Compared with neural network, deep forest can effectively deal with data of different scales and has more stable and good learning performance. Under the same hyperparametric setting as deep neural network, deep forest can obtain excellent performance in dealing with different data in different fields [3].

# 3. DEEP LEARNING IN AUTONOMOUS VEHICLE APPLICATIONS

## 3.1 Main sensors

Common cameras are divided into monocular and binocular [4]. The monocular camera has a simple structure and camera calibration, but it is difficult to determine the real size of the object in a single photo. Only when it is moving can the depth information be inferred. Binocular cameras use two cameras to locate. For a point on an object, you only need to know the exact position of the two cameras to know the coordinate value of the feature point in the same coordinate system, that is, the position of the feature point. Vision sensor is easy to install, has a huge amount of information and various algorithms, but it is easy to be affected by lighting and meteorological conditions.

3d-Lidar is mostly used for three-dimensional target detection. Compared with the two-dimensional image obtained by the camera, it can also obtain the depth information of the environment. At the same time, it has extremely high resolution and range detection accuracy, up to centimeter level. Therefore, almost all autonomous 8 unmanned system technology Volume 4 sensor systems include lidar. Lidar uses laser beam for detection. It is a high-precision detection equipment combined with laser technology and modern photoelectric technology. Millimeter wave radar refers to the radar with the wavelength of millimeter, and the main working frequency bands are 24 GHz, 77 GHz and 60 GHz[5]. 24 GHz is used for short-range collision warning and blind spot detection. 77 GHz has better range resolution and can improve the ranging accuracy on the road. Millimeter wave radar is not affected by the weather. Its detection distance is long, its performance is stable, its range resolution is high and its cost is low. Its disadvantage is that the pedestrian perception ability is weak, the building cannot be modeled, and the angle resolution is poor. Deep learning methods are mostly used to detect and recognize objects from 2D images (obtained by cameras) and 3D point clouds (obtained by lidar, etc.).

## 3.2 Target detection technology

Autopilot is edge calculation and combination of deep learning along with the rapid progress of science and

technology, one of the typical applications of the wisdom of the urban traffic system and from ideal to reality not out of reach. As is known to all, automatic driving technology is to positioning sensors, video processing, target recognition, radar, road decision-making techniques such as effective combination. In the real world city traffic is a real time change, sensors in a moving vehicle every moment will be a large amount of data received from the surrounding environment. Combining deep learning and edge computation can greatly reduce the time delay of data produced in the process of transmission, to improve the safety of the car networking system. Some researchers at home and abroad have automatic calculation and deep learning together the edge of the driving field carried out detailed research [6]. This section will be from the target awareness, path planning and collision detection and avoid dwelt on in recent years, the research status in this field.

Target detection is one of the most important research problems in the field of automatic driving. As described in document [7], automatic driving technology can realize the simultaneous interpreting of intelligent vehicles, that is, the perception of surrounding targets. Different vehicle sensors will use different sensors to perform corresponding sensing tasks, such as road detection, Vehicle detection and pedestrian detection, etc. The results of different detection tasks will be used for the realization of subsequent tasks such as path planning and vehicle reception control. Due to the original limitations of the detection algorithm, the complexity of the road environment and multi-level constraints, the detection algorithm that meets the requirements of engineering practice accuracy and operation efficiency is also limited.

In recent years, people have done a lot of research on vehicle detection and counting and pedestrian detection in roads mainly based on shallow learning. Shallow learning relies on manual feature extraction. The basic steps of vehicle detection are described in the literature [8], starting from selecting areas where there may be cars, For the vertical and horizontal filtering directions, the two groups of histogram gradient features are extracted, and the differences between vehicles and objects are mainly completed through mutual information measurement, standardized cross-correlation, and the combination of correlation measurement and support vector machine. Then, by connecting the direction value with the points classified as vehicles, the points belonging to the same vehicle are merged, So as to complete the detection of vehicles. In the development process of automatic vehicles, they are equipped with a variety of emerging sensors, and also put forward high requirements for the accuracy and real-time performance of detection. Due to the huge sensing data, intelligent vehicles are facing a huge computational burden, Computing power will become a bottleneck that prevents vehicles from benefiting from the high system accuracy brought by high-resolution cameras. At this time, applying deep learning to target detection will help to improve the accuracy of detection. However, the process of deep learning training requires a lot of computing and storage resources. Performing the above tasks in the cloud server will lead to high bandwidth consumption With the development of

edge computing, target detection based on deep learning can be migrated near the data source, that is, to the terminal device or edge node. In the terminal layer, on-board radar, high-definition on-board camera and other devices are responsible for the collection of image and video resources, and use the intelligent device of the terminal for compression, preprocessing and image segmentation. Then unload the data to be calculated to the edge node. By reducing the unnecessary filters in the convolution neural network layer, we can effectively reduce the resource consumption of the edge layer and improve the overall performance while ensuring the analysis performance. Next, we will further elaborate the target detection based on deep learning

### 3.2.1 Road detection

The moving vehicle needs to detect the lane line in real time to determine the forward direction. The lane marking detection algorithm first removes the pavement constituting the lane marking background, and then uses a group of waveform generation areas from local images. The experimental results show that the detection error rate is only 0.63% in the daytime and 1.14% even at night. But the disadvantage is that the algorithm can not prove that it can maintain low error rate in complex scenes. In order to test the accuracy in complex scenes, it's a nice option to use the data of various sensors such as lidar and high-speed cameras, and use deep neural network to detect lanes in three-dimensional space. The proposed method shows good performance in complex scenes such as blocking, bifurcation, merging and intersection. An end-to-end method to train Lane detectors is also a good choice. Firstly, the weight map of similar segments of each lane line is predicted by depth network, and then the parameters of the best fitting curve are returned for each lane line by weighted least squares. Compared with the traditional two-step method, the result is significantly improved under the condition of 70 frames. Aiming at the problem that the lane line and boundary of the road are blurred, a recursive neuron layer for structured visual detection can automatically detect the lane boundary. However, the model is relatively large, and the training time may be too long. In order to further shorten the training time, the full convolution network algorithm takes the position a priori as a feature map and directly adds it to the final feature map to improve the detection performance by learning more road boundary recognition features. Compared with the traditional model, its convergence speed is increased by 30%, which can effectively save the training time.

### 3.2.2 Vehicle and environmental detection

In order to avoid accidents, the self driving vehicle needs to detect and track other vehicles on the road and obstruct the suspicious obstacles that drive the vehicle. In this task, it is necessary to estimate the shape of surrounding vehicles or obstacles, relative speed with the vehicle, relative three-dimensional position and other factors. The vehicle counting system introduced in reference [9]

mainly uses a convolutional neural network to regress the vehicle spatial density map on aerial images. The evaluation results on Munich and high-altitude image research data sets show that it has high precision and recall based on convolution neural network, the vehicle trajectory can be extracted by using fast feature points to obtain the data of different vehicles, such as the number, driving direction, vehicle type, vehicle number and so on. Compared with the traditional method of monitoring vehicle flow with hardware, it has low cost and high stability. In the environment of mobile edge computing, there is no need to carry out large-scale construction or installation of existing monitoring equipment. Document [10] proposes a fusion strategy of camera and lidar for target recognition. By projecting the lidar 3D onto a 2D image plane, and then using the up sampling strategy to generate a high-resolution 2D range view, the convolution neural network is used for three channel color image classification and depth image classification, and the actual distance from the identified vehicle and environment is included in the sensing system. High complexity algorithms not only shorten the response delay of the system, but also put forward new requirements for the computing power and energy consumption of the edge server. Spike neural networks can use time coding for target recognition. Its advantage is that it can effectively reduce the energy consumption and delay of the system when recognizing targets in the real world environment, but the recognition accuracy still has room for further improvement. The question of how to balance recognition accuracy, system delay, and energy consumption in an edge computing environment will be the wind vane of future research in this field.

### 3.2.3 Pedestrian detection

Compared with other objects, pedestrians have a higher level of importance, so it is necessary to distinguish ordinary objects to be detected from pedestrians The automatic driving vehicle is used to detect, track and identify pedestrians through visual cameras, so as to avoid collisions with pedestrians. Although the recognition framework proposed in reference [11] can obtain higher accuracy of pedestrian detection, the disadvantage is that the processing time is significantly higher than that of other models Reference [12] proposes a hybrid local multi system based on convolutional neural network and support vector machine, which divides the complete image into multiple local sub regions, uses principal component analysis to screen the discriminative features, and applies empirical minimization and structural risk minimization methods to multiple support vector machines, with an average accuracy of pedestrian detection of more than 90% Using partial context network to detect pedestrians through body semantic information and context information, a strong complementary pedestrian detector can be designed, especially for blocked pedestrians, with low bit error rate and high positioning accuracy, so as to improve the detection effect of driverless vehicles on pedestrians and improve the safety factor

## 4. CONCLUSION

This paper introduces some typical deep learning models such as deep Boltzmann machine, deep reinforcement learning and deep forest, and then discusses the literature on the application of deep learning in the field of automatic driving target detection from four aspects: hardware facilities, road detection, vehicle and environment detection and pedestrian detection.

Self driving vehicles emphasize the real-time nature of computing tasks and require ultra-low latency interaction and powerful computing. At present, it is still necessary to further study how to analyze the image and video data collected by vehicle mounted sensors in time, and transfer the processing results to the automatic driving system in real time.

## REFERENCES

1. HUANG W, SONG G, HONG H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning[J]. IEEE Transactions on Intelligent transportation Systems, 2014, 15(5):2191-2201.

2. Liu Quan, Zhai Jianwei, Zhang Zongchang, Zhong Shan, etc. A summary of deep reinforcement learning. Journal of computer science, 2018

3. ZHOU Z H, FENG J. Deep forest: towards an alternative to deep neural networks [C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. 2017:3553-3559

4. Caselitz T, Steder B, Ruhnke M, et al. Monocular camera localization in 3D LiDAR maps[C]. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, October 9-14, 2016.

5. Zheng Fuyu Automatic mobile robot based on low-cost two-dimensional lidar [D] Guangzhou: Guangdong University of technology, 2018

6. Wang Yingfeng, Xu Yansong, Chen Yanhai Overview of intelligent vehicle multi-target detection technology based on multi-sensor fusion [J] Journal of automotive safety and energy conservation, 2021,12 (04): 440-455

7. Zijiang Zhu,Zhenlong Hu,Weihuang Dai,Hang Chen,Zhihan Lv.Deep learning for autonomousvehicle and pedestrian interaction safety. Safety ScienceVolume 145, January 2022, 105479

8. WANGQ, GAOJ, YUAN Y. Embedding structured contour and location prior in siamesed fully convolutional networks for road detection[J]. IEEE TransIntell, 2018:230-241.

9. TAYARAH, KIM G S, KIL T C. Vehicle detection and counting in high-resolution aerial images using convolutional regression neural network[J]. IEEE Access, 2017(6): 2220-2230.

10. MELOTTI G, PREMEBIDA C, GONÇALVES N. Multimodal deep-learning for object recognition combining camera and LIDAR data[J]. In 2020 IEEE

International Conference on Autonomous Robot Systems and Competitions(ICARSC). 2020:177-182.

11. BARBA-GUAMAN L, JOSE E N, ANTHONY O. Deep learning framework for vehicle and pedestrian detection in rural roads on an embedded gpu[J]. Electronics, 2020(4): 1-17.

12. UCAR A, DEMIR Y, GJZELI C. Object recognition and detection with deep learning for autonomous driving applications[J]. 2017,93(9): 759-769.