Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28516

Integration of Deep Learning and Signal Processing in Autonomous Driving

Yuxiao Gu

College of Electronic and Information Engineering, Tongji University, Shanghai, China
18862871208@163.com

Abstract. With the rapid development of intelligent transportation, autonomous driving has become a core direction of technological innovation. Its safe and efficient operation relies on precise processing of multi-source sensor signals. Traditional signal processing methods, limited by manual feature design, often fail to extract complete features and suppress noise in complex traffic environments. This leads to perception biases and decision-making delays, posing safety hazards and restricting the implementation of autonomous driving technology. By combining literature review and technical analysis, this paper focuses on the integration of deep learning and autonomous driving signal processing, aiming to explore how to enhance the efficiency and reliability of signal processing through deep learning algorithms. It concludes that deep learning, through its adaptive feature learning ability, can significantly improve the real-time performance and accuracy of signal processing. However, challenges such as data scarcity, poor model interpretability, and high computational costs still exist. Future research should focus on technological breakthroughs such as lightweight model design and multi-sensor fusion to promote the large-scale application of deep learning in autonomous driving signal processing.

Keywords: deep learning, autonomous driving, signal processing, neural network

1. Introduction

As the core carrier of the intelligent transportation system, the development level of autonomous driving technology directly determines the safety and efficiency of future travel. The basis of all this lies in the accurate analysis and efficient use of multi-source sensor signals. At present, self driving vehicles are generally equipped with cameras, lidars, millimeter wave radars, ultrasonic sensors and other types of equipment. Each type of sensor generates tens to hundreds of MB of raw signals per second - the image signal output by the camera needs to identify lane lines, traffic signs and moving targets, the point cloud signal of the lidar needs to complete three-dimensional space modeling and obstacle positioning, and the RF signal of the millimeter wave radar needs to capture target speed and distance information. These signals together constitute the "sensory nerve" of the automatic driving system, and their processing quality directly affects the reliability of subsequent decision-making and control.

In recent years, deep learning has shown its unique advantages in adaptive feature extraction and nonlinear fitting. Its ability to automatically learn high-dimensional features from massive data

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28516

provides a new scheme for automatic driving signal processing. Existing researches have applied deep learning to target detection, path planning and other links, but lack of systematic analysis of its multi-level fusion mechanism with signal processing.

This paper combs the application of deep learning in signal processing, explores its technical path in automatic driving environment perception, decision planning, control implementation, and discusses the current challenges and future directions, aiming to provide reference for the industrialization of automatic driving.

2. Overview of deep learning

2.1. Overview of deep learning model

Deep learning models excel in adaptive feature extraction, simulating human neural networks to achieve hierarchical feature extraction through multi-layer nonlinear transformations. This section examines four key deep learning models used in autonomous driving signal processing, analyzing their architecture principles, adaptability, and typical applications.

The structural design of convolutional neural network (CNN) simulates the process of human visual perception, and realizes signal feature extraction through the cooperative work of convolution layer, pooling layer and full connection layer [1]. Among them, the convolution layer captures the local features of the signal by sliding convolution check, such as extracting the lane edge, vehicle contour and other features from the road image. The pool layer reduces the number of model parameters by down sampling, while retaining key features to avoid over fitting; The full connection layer maps the previously extracted multi-layer features to the output space, and finally completes the task of signal classification or regression. In automatic driving signal processing, convolutional neural network is especially suitable for processing signals with grid structure, such as image signals collected by vehicle camera, pseudo image signals after laser radar point cloud conversion, etc., due to the characteristics of local perception and weight sharing. For example, in a rainy road scene, the convolutional neural network can autonomously learn the color and texture features of lane lines through multi-layer convolution, without manually adjusting the feature template. Compared with the traditional edge detection algorithm, its effect on feature extraction of fuzzy image is better.

Both recurrent neural network (RNN) and long short term memory (LSTM) are time-series signal processing models, which are mainly used to process time-dependent signals, such as target velocity signals continuously scanned by millimeter wave radar and range time-series data collected by ultrasonic sensors. By introducing feedback connection into the network structure, the recurrent neural network can realize the memory of historical signal information, so as to capture the time sequence change law of the signal. However, when dealing with long sequence signals, the traditional cyclic neural network is prone to the problem of gradient disappearance or gradient explosion, resulting in the model can not effectively learn the long-term dependence. To address this issue, Long Short-Term Memory (LSTM) networks selectively retain and update historical signal information through a gating mechanism comprising the "forget gate," "input gate," and "output gate" [2,3]. The forget gate filters out useless noise signals in the sequence, the input gate controls the update weight of the effective signal at the current time, and the output gate adjusts the contribution of the network state to the output result. In the automatic driving signal processing, the long-term and short-term memory network has become the mainstream model of time series signal analysis. For example, in the target tracking task, the long-term and short-term memory network can learn the acceleration, deceleration and lane changing laws of the target vehicle from the continuous

DOI: 10.54254/2755-2721/2025.LD28516

scanning signals of the millimeter wave radar, and provide accurate trajectory prediction results for the subsequent path planning.

The generative adversarial network (GAN) is composed of two parts: generator and discriminator. The generation and optimization of signals are realized through the confrontation training of the two parts. Among them, the core function of the generator is to learn the distribution characteristics of the real signal, and then generate a forged signal highly similar to the real signal; The discriminator is responsible for distinguishing whether the input signal is a real signal or a forged signal generated by the generator [4]. In the process of repeated confrontation training, the performance of generator and discriminator is continuously improved, and finally the generator can output enough signals to confuse the true with the false. In the field of automatic driving signal processing, the application of generated countermeasure network mainly focuses on two directions: first, the enhancement of extreme scene data. In view of the difficulty of data acquisition in rare scenes such as Blizzard, rainstorm and strong light at night, the generated countermeasure network can generate training data simulating extreme scenes based on normal scene signals to make up for the scarcity of real data; The second is the repair of damaged signals. When some pixels of the camera fail due to overexposure or occlusion of strong light, the generation of countermeasures network can complement the signals of the damaged area based on the characteristics of the surrounding normal pixels and ensure the continuity of subsequent environmental perception.

2.2. Signal processing method based on deep learning

Compared with traditional methods, deep learning driven signal processing realizes end-to-end feature learning and reduces the dependence on expert experience.

Signal noise reduction is the primary link of automatic driving signal processing. The original signal collected by the sensor often contains useless information such as electronic noise and environmental interference. If it is not filtered in time, it will directly affect the accuracy of subsequent feature extraction and target recognition. At present, the signal denoising methods based on deep learning are mainly divided into two categories: self encoder denoising and generation countermeasure network denoising. The self encoder realizes signal noise reduction through the symmetrical structure of "encoder decoder". The encoder maps the noisy signal to the low dimensional feature space and filters the noise information through feature compression; The decoder reconstructs clean signals from low dimensional features. In the noise reduction of LIDAR point cloud signal, researchers often use convolutional self encoder to convert the point cloud signal into voxel grid and input it into the network. The encoder extracts the noiseless structural features, and the decoder restores the point cloud details. After this method, the accuracy of lidar target segmentation can be improved by more than 20%. The noise reduction effect of the generated countermeasure network is more prominent in the scene of "strong noise+signal distortion".

Signal classification is one of the core tasks of automatic driving environment perception. Its purpose is to divide the processed sensor signals into different categories such as "vehicle", "pedestrian", "bicycle" and "traffic signs", so as to provide basis for subsequent decision-making and planning. Deep learning significantly improves the robustness and accuracy of classification tasks through the discriminant characteristics of autonomous learning signals. Convolutional neural network is the most widely used model in image signal classification. For example, in the task of traffic sign recognition, lenet-5 model can extract shape and color features from traffic sign images with different lighting and angles through 5-layer convolution and pooling operation, and the classification accuracy on the German traffic sign data set (GTSRB) exceeds 98.77% [5]. In the time sequence signal classification, the long-term and short-term memory network shows its unique

DOI: 10.54254/2755-2721/2025.LD28516

advantages: the millimeter wave radar has different time sequence rules for the echo signals of different targets. The long-term and short-term memory network learns these rules through the gating mechanism, and the accuracy rate in the radar target classification task can reach 91%, which is 18% higher than the traditional hidden Markov model.

Autonomous driving sensors generate vast amounts of signals that need efficient storage and transmission. Traditional compression methods like JPEG and H.265, though effective, can distort signals, impacting subsequent processing accuracy. Deep learning-based compression achieves higher ratios while maintaining quality by learning sparse signal representations. Moreover, multi-source signal fusion is crucial for overcoming single-sensor limitations. Cameras excel at capturing color and texture but are light-dependent; LiDAR offers precise positioning yet performs poorly in adverse weather; millimeter-wave radar is robust but low-resolution. Deep learning fuses these signals by constructing a unified feature space. For instance, a CNN-LSTM model extracts spatial features from camera images with CNNs and temporal features from radar data with LSTMs, then integrates them via fully connected layers. This fusion significantly enhances target detection accuracy in dark scenes compared to using a single sensor.

3. Technical application of deep learning in automatic driving signal processing

The autonomous driving system's core functions—environmental perception, decision planning, and control execution—each have unique signal processing needs. This section reviews deep learning applications in these areas based on current research and industry practice.

3.1. Environmental perception level

Environmental perception is the "perception organ" of automatic driving system. Its main task is to realize target detection and tracking, lane line detection and traffic sign recognition through multi-sensor signal processing, and provide accurate environmental information for subsequent decision-making and planning. At present, deep learning has become the core technology to improve the accuracy and robustness of environmental perception.

Target detection and tracking is the core task of environmental perception. It is necessary to identify dynamic targets such as vehicles, pedestrians and bicycles from sensor signals, and predict their motion trajectories to avoid collision risks.

Lane line detection is the key to realize the function of vehicle lane keeping, which needs to accurately extract the position and direction of lane line from the image signal. Traditional methods such as Canny edge detection+Hough transform are prone to failure in lane wear, occlusion or complex lighting scenes, while the deep learning model based on semantic segmentation shows stronger robustness.

Traffic sign recognition needs to deal with interference factors such as light change and angle offset, and lightweight deep learning model has become the preferred vehicle scene. In the aspect of image signal processing, YOLO series models are currently the most widely used target detection models. Among them, the YOLOv8 model achieves multi-scale feature extraction through the three-tier architecture of "backbone network neck network head network": the backbone network uses the csparknet structure to extract the global features of the image, the neck network uses the pan-fpn structure to fuse the features of different levels, and the head network outputs the location and category information of the target [6].

DOI: 10.54254/2755-2721/2025.LD28516

3.2. Decision planning level

Decision-making and planning in autonomous driving rely on processed environmental perception signals, the vehicle's state, and destination routes to formulate driving strategies and optimal trajectories. Deep learning offers new approaches for dynamic and complex scenarios by learning human driving experiences or interaction patterns with the environment. This process is divided into behavioral decision-making and path planning.

Behavioral decision-making outputs driving instructions like acceleration, deceleration, lane changes, and parking based on real-time signals. Deep learning methods include reinforcement learning and imitation learning. Reinforcement learning involves agents interacting with the environment to learn optimal strategies through reward signals. The Proximal Policy Optimization (PPO) algorithm, which stabilizes training by constraining policy update steps, has a 92% success rate in automatic lane changes, outperforming earlier Deep Q-Network (DQN) algorithms by 10% and providing smoother transitions and higher passenger comfort [7]. Imitation learning combines human driving data with reinforcement learning to reduce trial-and-error costs.

Path planning aims to generate trajectories that ensure safety, efficiency, and comfort. Traditional algorithms, relying on preset cost functions, struggle with dynamic traffic. Deep learning models, like Graph Neural Networks (GNNs), represent road networks as graphs, using graph convolutions to learn traffic patterns and plan optimal routes. Reinforcement learning also excels in dynamic path optimization. For example, during rush hours, the Deep Deterministic Policy Gradient (DDPG) algorithm, incorporating path length, travel time, and collision risk into its reward function, interacts with virtual traffic environments to learn strategies that avoid congestion, reducing average travel time by 25% compared to traditional methods [8].

3.3. Control execution level

The control execution needs to convert the trajectory command output from the decision planning into the actual action of the vehicle, so as to ensure that the vehicle runs stably according to the planned trajectory. Traditional control methods rely on manual adjustment of parameters, and the control accuracy is insufficient under complex road conditions. However, deep learning realizes adaptive control by learning vehicle dynamics and driving data [9].

Longitudinal control is responsible for adjusting vehicle speed, including acceleration, deceleration, car following, start stop and other functions. The signals to be processed include the target speed output by decision-making and planning, the distance and speed of the front vehicle collected by millimeter wave radar, etc. CNN-LSTM hybrid model is a common model for longitudinal control: convolutional neural network extracts the spatial characteristics of the current speed and the distance ahead of the vehicle, and the long-term and short-term memory network captures the time sequence of speed changes, and finally outputs the throttle opening or braking pressure command [2].

Lateral control is responsible for adjusting the steering angle of the vehicle to ensure that the vehicle travels along the planned trajectory. The signals to be processed include trajectory curvature, the deviation between the current position of the vehicle and the trajectory, etc. The lateral control model based on convolutional neural network converts the planned trajectory into a two-dimensional curvature map, extracts the curvature change characteristics by using the convolution layer, and outputs the steering angle command. Fault tolerant control is an important supplement to ensure the safety of control execution. When a sensor fails, deep learning can ensure control stability through signal reconstruction.

DOI: 10.54254/2755-2721/2025.LD28516

4. Challenges and future prospects

Despite significant progress in autonomous driving signal processing, deep learning still faces multidimensional challenges from R&D to industrial implementation.

First of all, the performance of the deep learning model depends on massive annotation data, but there are three major pain points in the automatic driving signal data: one is the scarcity of extreme scene data. The probability of rainstorm, Blizzard, traffic accident and other scenarios is low, and it is difficult to collect enough annotation signals, resulting in a sharp drop in the processing accuracy of the model in these scenarios. The second is data distribution offset. Because there may be significant differences between actual environmental scenes, such as differences in common climate conditions and road complexity, there are differences in signal distribution between training data and actual application scenes, resulting in poor generalization ability of the model. Third, privacy issues. The automatic driving signal contains sensitive information such as vehicle location and surrounding environment, which is easy to cause leakage risk if it is directly shared.

Secondly, due to the untraceable decision-making process and the lack of quantifiable basis, the "black box" characteristics caused by in-depth learning contradict the safety requirements of automatic driving. The decision-making process of the model relies on complex internal parameter operations and lacks an intuitive logical chain. Once there is a deviation in signal processing, it is difficult to locate the specific reason. At the same time, the real-time performance of the model is difficult to meet the requirements of automatic driving, and the complex structure of the deep neural network leads to a huge amount of calculation. Even the fusion model designed for multi-sensor signals, the running speed on the vehicle terminal equipment is often lower than the real-time standard required by the actual scene, affecting the response efficiency of the system to the dynamic environment.

Third, at present, the computing power level of on-board computing hardware is limited, and there is a significant gap with server level computing equipment, which is difficult to support the deployment of complex deep learning models, limiting the practical application of multi-source signal fusion, high-precision environment perception and other technologies; At the same time, the high cost of high-performance sensors and dedicated computing hardware restricts the large-scale popularization of automatic driving technology to a certain extent. Automatic driving signal processing needs to meet the two major standards of functional safety and expected functional safety. The functional safety requires that when the model or hardware fails, it has the ability of redundant error correction to avoid safety accidents, but the randomness of the parameter optimization of the deep learning model is difficult to ensure the stability of the failover process; The expected functional safety requires that there is no signal processing deviation when the model operates normally. However, the current problem of insufficient data collection in extreme scenarios is likely to lead to "unexpected failure" of the model, which can not accurately process the signals in unconventional scenarios, and then generate potential safety hazards [10].

To address these challenges, collaborative efforts between academia and industry are ongoing. Self-supervised learning reduces reliance on labeled data, while federated learning enables collaborative model optimization without sharing raw data. The cooperative practice between Baidu Apollo and several auto companies shows that the multi-sensor fusion model using federal learning and training improves the accuracy by 12% compared with single enterprise data training while protecting data privacy [11]. In addition, digital twin technology generates extreme scenario data through virtual simulations, compensating for real-world data limitations. Model optimization focuses on lightweighting, interpretability, and robustness. Techniques like knowledge distillation, pruning, and quantization reduce model size and computation while maintaining accuracy,

DOI: 10.54254/2755-2721/2025.LD28516

improving real-time performance and robustness in complex environments. These innovations aim to promote the large-scale application of deep learning in autonomous driving signal processing.

5. Conclusion

Through adaptive feature learning ability, deep learning significantly improves the accuracy of automatic driving signal processing. Deep learning shows high accuracy and adaptability in vehicle environment perception and obstacle recognition. It also realizes the design of decision-making mechanism and the construction of path prediction model in complex environment, which significantly improves the intelligent and accurate effect of path planning. At the signal processing level, multi type network models solve the processing problems of image, time sequence and generated signals respectively. However, the current fusion technology still faces many challenges, which need to be solved through interdisciplinary collaboration and industrial collaboration.

With the continuous improvement of algorithm efficiency and the optimization and integration of computing resources, deep learning technology will be more widely used in more complex traffic scenes to provide a safer and more efficient path for driverless vehicles. The future development needs to focus on several major directions: the data end breaks through the data bottleneck through self-monitoring learning, federal learning and digital twins; At the model end, the practicality is improved through lightweight, interpretability and robustness optimization; The hardware side provides computing power support through special chips and multi chips. With the breakthrough of these technologies, the integration of in-depth learning and automatic driving signal processing will move from "technical feasibility" to "industrial availability", promote the leap of automatic driving from L3 to L4, and finally realize the future travel vision of "safety, efficiency and convenience".

Overall, the limitation of this article is that it focuses on the integration of deep learning and signal processing, and does not involve the comparison of the hybrid architecture of traditional methods and deep learning. In the future, we can further explore the hybrid signal processing architecture of "traditional method+deep learning", and strengthen the research on the technical practice of domestic autonomous driving enterprises, so as to provide a more comprehensive reference for industrial development.

References

[1] Qiu, Y. M., Gu, J. C., & Pan, S. Y. (2023). Intelligent vehicle target detection and visual navigation based on deep learning. Light Industry Science and Technology, 39(04), 107-110.

[2] Huang, T. N. (2025). Application of Deep Learning Algorithms in Driverless Path Planning. China Computer & Communication, 37(11), 10-12.

[3] Pu, Z. Z., & Luo, N. N. (2023). Construction of an automatic language verification model integrating long short-term memory and recurrent neural networks. Automation & Instrumentation, (12), 23-27. https: //doi.org/10.14016/j.cnki.1001-9227.2023.12.023

[4] Chen, W., Tu, J. L., Tan, Y. X., & Jin, F. J. (2021). Wind power scenario transfer generation using deep convolution generative adversarial network. Techniques of Automation and Applications, 1-6

[5] Jia, Y. C., Yang, Z. J., Peng, G. L., & Han, Y. N. (2025). Optimization algorithm for traffic sign recognition based on LeNet-5 network. Internet of Things Technologies, 15(15), 18-22. https: //doi.org/10.16667/j.issn.2095-1302.2025.15.004

[6] Zhou, X. X., Wang, K. Q., Zhou, X., Zhang, Q., & Xue, G. Q. (2025). DSA-YOLOv8 algorithm for traffic sign detection, 46(08), 2320-2327. https: //doi.org/10.16208/j.issn1000-7024.2025.08.024

[7] Wu, L. (2023). Research on Autonomous Vehicle Behavior Decision Making and Trajectory Planning. Chang'an University. https: //doi.org/10.26976/d.cnki.gchau.2023.001072

[8] Wang, X. J. (2024). Research on decision-making method of lane keeping with deep reinforcement learning in urban environments (Master's thesis). Xi'an University of Technology. https:

DOI: 10.54254/2755-2721/2025.LD28516

//doi.org/10.27398/d.cnki.gxalu.2024.000507
[9] Han, S. M., Xiao, F., & Cheng, W. S. (2023). Overview of the Application of Deep Reinforcement Learning in Autonomous Driving Systems (Natural Science Edition), 42(4), 25-31.
[10] Wu, Y. Q., & Cai, J. Q. (2022). Deep Learning-Based 3D Object Detection for Autonomous Driving：A Comprehensive Review. CAAI Transactions on Intelligent Systems, 1-23.
[11] Baidu Apollo. (2022). Apollo navigated assisted driving. https: //www.apollo.auto/anp