Computing, Performance and Communication Systems (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/cpcs.2023.070107
ISSN 2371-8889 Vol. 7 Num. 1

Design of Vehicle Roof-mounted Multi-functional Video Perception and Warning System

Jiawei Zhao[1,a,*], Ruikang Wang[2,b], Yiming Hao[1,c], Hongxiang Yin[3,d], Zhongjia Ma[2,e]

[1]*School of Transportation, Jilin University, Changchun, Jilin, China*
[2]*School of Automotive Engineering, Jilin University, Changchun, Jilin, China*
[3]*School of Materials Science and Engineering, Jilin University, Changchun, Jilin, China*
[a]*m15547852619@163.com*, [b]*cheb2000@163.com*, [c]*haoym1720@jlu.edu.cn*, [d]*yhx306170@163.com*, [e]*mazhongjia2002@gmail.com*
[*]*corresponding author*

Keywords: Roof-mounted Camera, Multi-Functional Video Perception, Route Suggestion, Driving Safety, Edge Computing, Lightweight, YOLOv8, Ultra Fast Deep Lane

Abstract: A multi-functional video perception and warning system based on vehicle roof-mounted cameras can analyze images captured by the cameras in real time, achieving information fusion, route suggestion, and visual enhancement functions, improving driving safety and experience. The system is developed based on edge computing devices, with no requirements for whole vehicle configuration, low threshold, and easy to promote. The system innovatively increases the height of traffic information acquisition, obtaining source advantages, and making long-distance result analysis possible. Combining these features, more system functions can be developed, such as route selection, live streaming, and visual enhancement. The project solves key issues such as designing and perfecting the mechanical structure for controlling the stable lifting and rotation of the roof-mounted camera, automating the control of the robotic arm in combination with real-world scenarios to achieve long-distance route selection, real-time operation of target detection and route recommendation algorithms on low-power devices, adjusting adaptively to different weather and road conditions, and designing software for automatic and manual control interaction between the roof-mounted camera.

1. Introduction

With the continuous development of transportation and the accelerated progress of urbanization, road traffic congestion is becoming more and more serious, and traffic safety is also receiving increasing attention [1][2]. In this context, the intelligent, automated, and networked development of vehicles has become a hotspot in current research. In this background, vehicle-mounted video perception and warning systems have become widely used technical means, and roof-mounted multi-functional video perception and warning systems, as an important type, have also attracted attention.

At present, there have been some research achievements in vehicle-mounted video perception and warning systems both domestically and internationally [3]. For example, the domestic "Vehicle-Road Cooperative System" and the international "Intelligent Transportation System" are representative research results in this field. However, most of these systems have only implemented single functions, such as detection of road driving status, vehicle follow-up control, etc., and have not truly achieved multi-functional integration [4-5].

Therefore, this project aims to develop a roof-mounted multi-functional video perception and warning system, integrating multiple functions such as road driving condition detection, lane departure warning, intelligent route selection, real-time live streaming, and visual enhancement. Through the application of intelligent recognition algorithms, the system can analyze road conditions in real time, provide timely driving suggestions, improve driving safety and experience, and provide drivers with comprehensive vehicle perception and real-time monitoring of road conditions.

2. Optimization Principle

2.1. Optimization Ideas

This system is mainly composed of mechanical structures and algorithm software, and is equipped with an accompanying app to enhance the user experience.

2.1.1. Optimization of the Roof-mounted Camera's Mechanical Part

To optimize the mechanical part of the roof-mounted camera, this scheme adopts optimization design methods and finite element analysis to improve compactness, stability, and reliability. The drive mechanism is improved using suitable motors and transmission devices, along with adaptive control technology for real-time adjustment. High-strength materials are selected to enhance component strength and stiffness. Automatic cleaning technology is introduced to reduce pollution impact and manual maintenance costs.

2.1.2. Algorithm Software and Edge Computing Device Optimization

To improve model computation efficiency, this scheme uses Jetson edge computing devices and optimizes them using PyTorch-GPU/OpenCV/Python3 environment.

The latest YOLOv8 (Ultralytics) is chosen as the end-to-end detection model to enhance detection efficiency and accuracy.

The traditional image edge filter is combined with the Ultra Fast Deep Lane model to achieve fast and accurate lane line detection.

A multi-lane counting scheme based on anchor points is developed to achieve multi-lane target detection result summarization.

2.1.3. Optimization of Multi-functional Video Perception and Warning App

UX optimization is achieved through user research and testing to identify possible issues and improve the user interface and interaction methods.

Functional upgrade optimization involves continuously improving the app's features to increase practicality and convenience.

Data optimization improves data accuracy, real-time performance, and reliability, enhancing the app's efficiency and accuracy.

Personalized recommendation optimization provides customized services based on user behavior and preferences, enhancing user satisfaction.

Security optimization strengthens data security protection, and network optimization improves network connection and transmission efficiency while ensuring network adaptability and robustness.

2.2. Research Method

2.2.1. Design of Vehicle Top-mounted Camera Mechanical Structure

We have written a complete process for implementing this mechanical structure in the form of an invention patent. The invention is a vehicle top-mounted camera installation mechanism, as shown in Figure 1. Traditional vehicle-mounted video capture devices have fixed installation positions, which can only increase the number of cameras for real-time side views of the vehicle, increasing costs and higher hardware requirements. The vehicle top-mounted camera installation mechanism can be installed on the top of the vehicle, achieving 360-degree rotation shooting, effectively solving the problem of limited driver visibility and improving driving safety. The mechanism includes components such as a screw frame, screw, drive assembly, support, connection plate, telescopic assembly, rotating assembly, and collection camera, which can achieve synchronous co-directional rotation, movement along the length direction of the connection plate, axial rotation, and other functions. The rotating assembly can also rotate 360 degrees, meeting different shooting angle requirements, as shown in Figure 2 and 3.

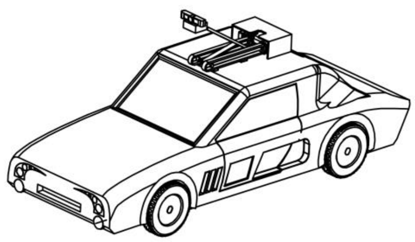

Figure 1: Schematic diagram of the overall structure.

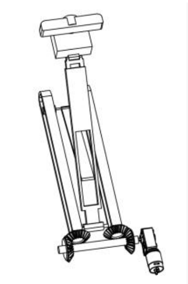

Figure 2: Partial view (1).

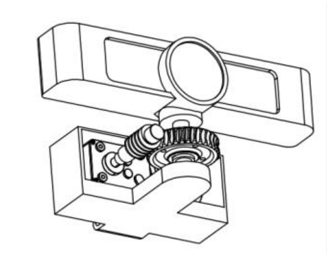

Figure 3: Partial view (2).

2.2.2. Multi-lane Vehicle Recognition, Judgment, and Decision-making Algorithm Model and Related Hardware

In the early stage of the research, the project team selected Jetson edge computing hardware produced by NVIDIA. Considering the balance between cost and performance, the Jetson Nano platform was finally chosen [6].

The platform is equipped with the Linux-based JetPack system and provides the same CUDA-X software and tools used by professionals worldwide for each Jetson developer. The development process is simplified due to the support of cloud-native technology. Developers can take advantage of GPU acceleration libraries and SDKs (such as NVIDIA DeepStream for intelligent video analysis) for deeper development. The benefits of the project team choosing this hardware are obvious. With lower cost, the Nano has its own onboard GPU, which most edge computing devices do not possess. Please refer to Table 1 for specific details.

Table 1: Jetson Nano 4GB Specification

Items	Parameters	
CPU	4Cores ARM Cortex-A57 MPCore Processor	
GPU	NVIDIA Maxwell Architecture 128 NVIDIA CUDA Cores	
VRAM	4 GB 64bit LPDDR4	
VRAM Bandwidth	1600 MHz – 25.6 GB/s	
Video Decoding	500MP/s	1x 4K @ 60 (HEVC)
		2x 4K @ 30 (HEVC)
		4x 1080p @ 60 (HEVC)
		8x 1080p @ 30 (HEVC)
Module Size	69.6mm×45mm	
Power Mode	5W/MaxN20W	

On this basis, NVIDIA has developed the TensorRT high-performance inference engine for its GPUs. Compared with PyTorch and OpenCV, it provides a faster and more efficient solution for deploying deep learning models. TensorRT supports trained models from mainstream deep learning frameworks such as Caffe, TensorFlow, and ONNX, and achieves efficient model execution through techniques such as pruning, optimization, and high parallelization. Jetson development boards support various interfaces such as GPIO and USB, allowing for the reasonable addition of hardware modules (such as CSI cameras) according to project requirements. After the project is completed, there is no need to use the development board; only the smaller, lower-power Jetson Nano Module embedded computing module is needed for operation. This hardware can be directly booted from an SD card, making it easy to replicate and deploy in batches, as shown in Figure 4.

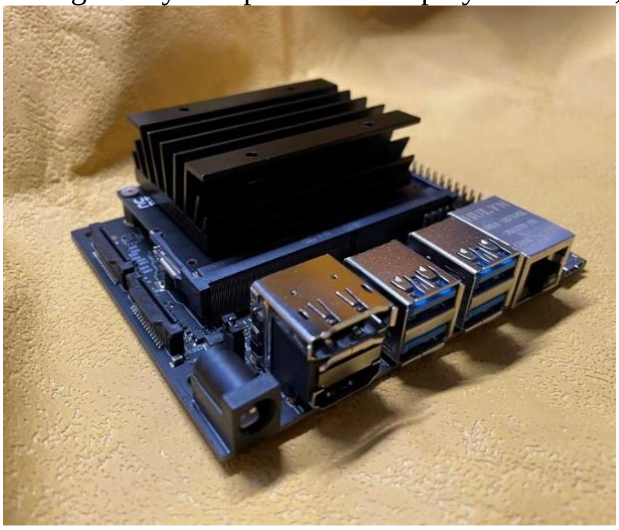

Figure 4: Jetson Nano Developer Kit

As shown in Figure 5, during the development process, the project team added expansion devices such as cooling fans, acrylic cases, wireless network cards, and CSI cameras to meet development needs.

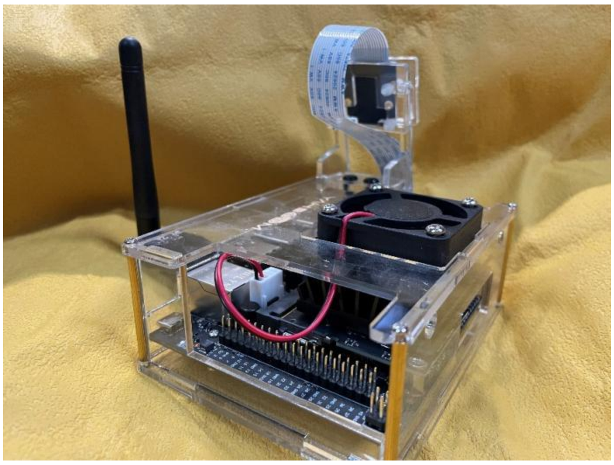

Figure 5: Jetson Nano in action

YouOnlyLookOncev8 (YOLOv8), developed by the Ultralytics team, is a cutting-edge, state-of-the-art YOLO model that can quickly perform end-to-end object detection, image segmentation, and classification. Building on the success of previous YOLO versions, it introduces new features and improvements to further enhance performance and flexibility. YOLOv8 is designed to be fast, accurate, and easy to use, making it an excellent choice for a wide range of object detection, image segmentation, and image classification tasks [7-8], as shown in Figure 6.

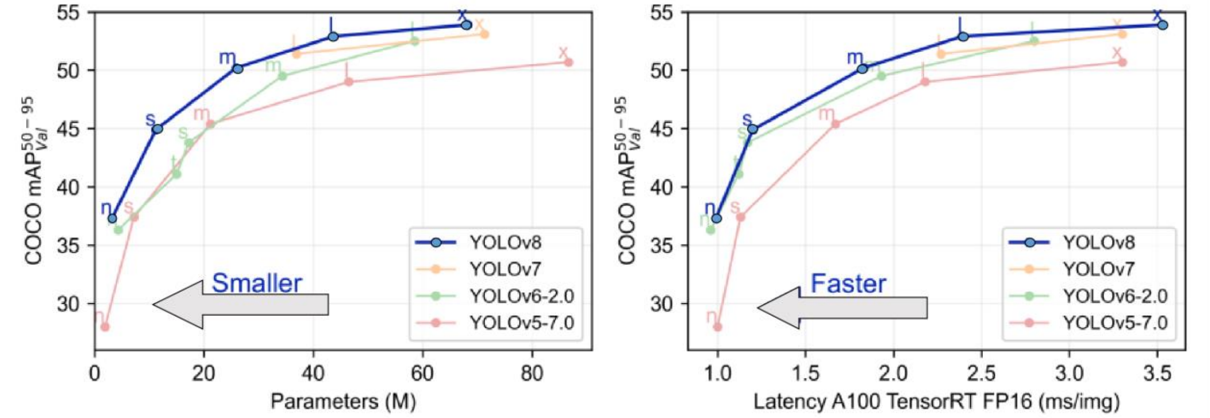

Figure 6: Comparison of YOLOv8 with other YOLO models

In this project, the team trained the model on a 3070 140W PC platform using the PyTorch framework, as shown in Figure 7. First, they needed to obtain the .pt pre-trained weight files generated by Ultralytics based on the COCO dataset. On top of that, the team further trained the YOLOv8 model using the UA-DETRAC traffic dataset and then tested it for accuracy on the test set. When the average loss value no longer decreases, and the detection mean average precision (mAPval) is relatively stable, the training can be considered complete. The team then exported the weight files generated during training to the ONNX format for easy portability and deployment on different platforms. Next, on the Jetson platform, they used the APIs provided by TensorRT to compile the ONNX file into an engine file, set the data input format (internal storage or camera), and computation accuracy to perform real-time object detection tasks, as shown in Figure 8 and Table 2.

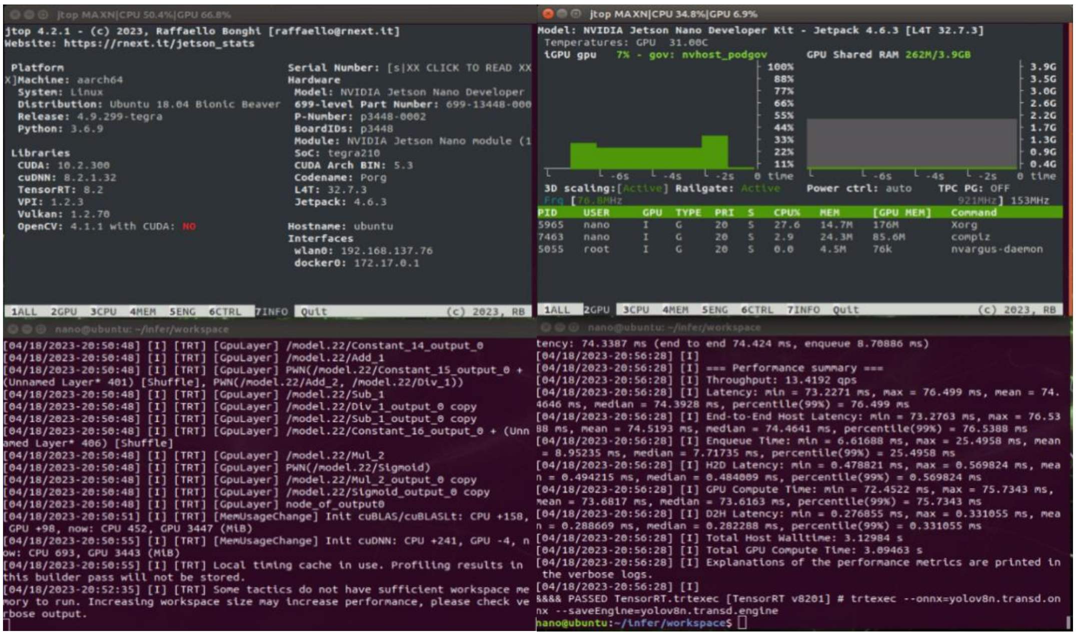

Figure 7: Generate the .engine file for TensorRT using static batch mode

Figure 8: YOLOv8 recognition results using default pre-trained weights

Table 2: YOLOv8 object detection performance indicators

Model	size	mAP^{val}	Speed (fps)	Speed (fps)	Speed (fps)
	(pixels)	0.5-0.95	Amazon EC2 P4d CPU ONNX	Jetson Nano PyTorch	Jetson Nano TensorRT
YOLOv8n	640	37.3	12.4	6.2	30
YOLOv8s	640	44.9	7.8	3.0	17
YOLOv8m	640	50.2	4.3	1.4	7.1
YOLOv8l	640	53.2	2.7	0.76	3.9
YOLOv8x	640	54.3	2.1	0.49	2.6

The mean average precision (mAP) values based on the COCO val2017 dataset using a single-model and single-scale approach.

Speed averaged over COCO val2017 images using Jetson Nano 4GB.

Due to the Ubuntu desktop performance consumption, the data might exhibit fluctuations, as shown in Figure 9.

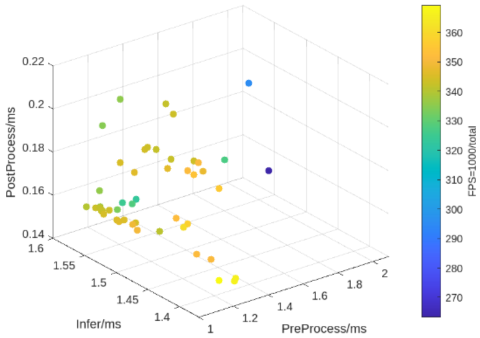

Figure 9: YOLOv8 RTX3070-140W TensorRT inference performance (48 experiments)

The schematic of a more mature lane-specific counting scheme in academia includes the following characteristics: static lane lines and monitoring perspective. Lane lines can be simply divided using a threshold, and then a specific position in each lane can be calibrated. When the vehicle's detection box passes this position, the statistical results are accumulated. In conjunction with DeepSort and other object tracking methods, vehicle trajectory lines are generated to assist in judgment. However, for the application scenario of this project, the position of the lanes is constantly changing, and the position of other vehicles relative to the vehicle equipped with this system is also changing. The results of lane recognition using traditional threshold, filtering, and edge extraction methods are very unstable [9].

Different from the above methods, this project uses the newly proposed Ultra Fast Deep Lane [10], which is an ultra-fast deep lane detection model that treats lane line detection as a classification problem and uses fully connected layers to enhance global perception capabilities, as shown in Figure 10.

Figure 10: UFDL's lane anchor point recognition results for 4K resolution video sources

Drawing on existing experimental data [11], this study compared the results of training on three datasets - TuSimple, CULane, and Mvirgo. The analysis revealed that the Mvirgo dataset exhibited the highest level of accuracy and optimal performance compared to the other two datasets [12]. In contrast, the testing outcomes from the CULane dataset were the poorest, while those from the TuSimple dataset were also suboptimal. The underperformance of both datasets stemmed primarily from weak low-level features such as dim images and indistinct lane markings, as shown in Figure 11. Moreover, combining Tusimple and Mvirgo datasets for training produced similar results as using the Mvirgo set alone but with comparatively lower performance in detecting dashed lines. In light of these findings, the Mvirgo dataset was determined to be the most suitable and was ultimately selected for training the lane detection neural network under investigation [13].

This model has a hybrid anchor point mechanism, using different anchor points for different lanes, which solves the lane detection performance issue [14]. While ensuring speed, it effectively handles situations such as occlusion and low light, achieving good practical performance, as shown in Figure 12.

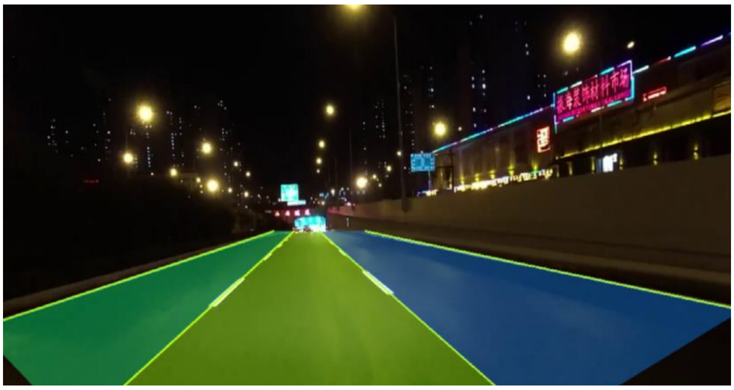

Figure 11: Lane anchor point fitting effect under low light conditions (nighttime)

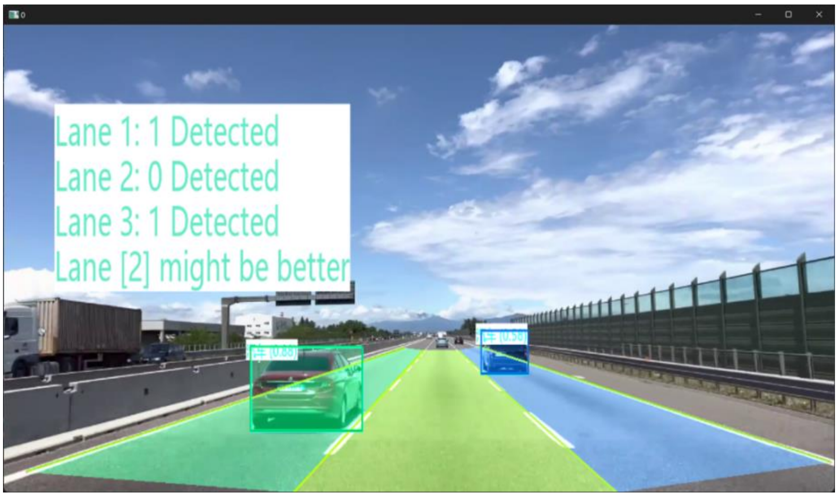

Figure 12: Inductive results of YOLOv8 combined with UFDL (real-time inference)

2.2.3. Design of a multifunctional video perception and warning APP

To enhance the user experience, we have developed a companion app:

"Happy Travel": road judgment and decision-making, voice guidance, road warnings, and visual enhancement;

"Happy Life": car owner forums, landscape live streaming, film and television shooting, and unique perspectives on the world; efficient third-party services to meet all needs along the way.

3. Innovation Features

Firstly, this system provides multi-lane real-time judgment functionality, offering optimal route planning suggestions based on real-time data, enabling drivers to avoid congested roads or choose the fastest routes. This feature is relatively rare in the market's navigation systems, greatly improving drivers' travel efficiency and experience.

Secondly, this system can flexibly adapt to various driving conditions and traffic scenarios, meeting the transportation needs of different cities and regions. Moreover, the system can perform vehicle object detection and multi-lane counting, providing more accurate traffic information for a better driving experience.

Thirdly, this system is based on the Jetson edge computing platform, enabling flexible deployment and real-time learning and optimization of neural network algorithms to improve computational efficiency and accuracy, thus achieving long-range result analysis. Additionally, the system can continuously improve and optimize algorithms based on real-time feedback and data learning from drivers.

Finally, this system has multiple extended features, with the top-mounted camera covering various functionalities such as driving assistance, entertainment, and information acquisition. For example, the system can provide high-definition panoramic video, driver status monitoring, and intelligent audio, offering a more comprehensive travel experience for drivers.

4. Application Prospects

This system delves into the intelligent transportation field, emphasizing safety and convenience, and can be applied in the automotive industry, transportation industry, and urban management industry. The top-mounted camera enhances the height and range of traffic information sources; the low-power edge computing hardware has a small size and simple power supply, is not dependent on the overall vehicle configuration, has a low threshold for use, and is easy for the general public to accept and promote. According to statistics, the global intelligent transportation market is continuously expanding, expected to reach 110 billion US dollars by 2025, with the Chinese market becoming one of the largest growth points in the global intelligent transportation market. In China, the intelligent transportation application field is vast, particularly with the rapid development of artificial intelligence and the Internet of Things technologies, the demand and market potential for intelligent transportation applications are further expanded, indicating that this project has excellent market prospects. Globally, this project can improve driving safety and comfort, optimize route selection, reduce vehicle congestion, enhance transportation efficiency, and subsequently lower accident rates and pollutant emissions, bringing significant economic and social benefits to society.

References

[1] Shi Xiaoran, Wan Zimo, Wu Chenxi. Application of Automation Technology in Mechanical Manufacturing. Southern Agricultural Machinery, 2016, 47(12), pp. 126.

[2] Zhang Lihong, Hao Junke. Exploration of Problems and Improvement Methods in Mechanical Automation Design and Manufacturing. Science & Technology Wind, 2020, No. 405(01), pp. 155.

[3] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0), pp. 186, 221.

[4] Zequn Qin; Pengyi Zhang; Xi Li. Ultra Fast Deep Lane Detection with Hybrid Anchor Driven Ordinal Classification.IEEE transactions on pattern analysis and machine intelligence, 2022: pp.1-14

[5] Du Zhongqiang, Tang Linbo, Han Yuqi. Lane Line Detection Method for Embedded Platforms. Infrared and Laser Engineering, 2022, 51(07): pp. 483-490.

[6] J. B. Kim, "Vehicle detection using deep learning technique in tunnel road environments," Symmetry (Basel), vol. 12, (12), pp. 1-12, 2020.

[7] M. Umair et al, "Efficient video-based vehicle queue length estimation using computer vision and deep learning for an urban traffic scenario," Processes, vol. 9, (10), pp. 1786, 2021.

[8] Park, J. et al. (2020). A High-Accuracy Roof-Mounted Perception System for Autonomous Driving. In Proceedings of the 2020 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 1-6.

[9] Li, C., Chen, M., Cao, J., & Sun, Z. (2020). Vehicle detection and tracking using multi-sensor fusion for intelligent transportation system. Measurement, 157, 107680, pp. 125-127.

[10] Kim, Y. J., Hong, S., & Kim, J. H. (2021). Efficient and Accurate Object Detection with Multi-Resolution Feature Fusion and Cascaded Region Proposal Networks. IEEE Transactions on Circuits and Systems for Video Technology, 31(1), pp. 83-94.
[11] Debnath, B. et al. (2021). Vehicle Detection and Tracking Using Roof-Mounted Cameras. In Proceedings of the 2021 International Conference on Electronics, Communications and Computers, pp. 132-137.
[12] Ryu, H. et al. (2019). Development of a Roof-Mounted Multi-Functional Video Perception System for Autonomous Vehicles. IEEE Access, 7, 51157-51168. pp. 235-236.
[13] Kim, S. et al. (2019). A Low-Cost Multi-Functional Roof-Mounted Perception System for Autonomous Vehicles. IEEE Access, 7, 96345-96355, pp. 3, 5.
[14] Cho, J. et al. (2019). A Roof-Mounted Multi-Functional Perception System for Autonomous Vehicles. In Proceedings of the 2019 IEEE International Conference on Robotics and Automation (ICRA), pp. 743-748.