Viraj Raghunath Sonawane, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Neural Network for Autonomous Vehicle Perception: A Review of Traffic Sign Recognition, Lane Detection, and Deployment Challenges

Research Scholar Viraj Raghunath Sonawane, Professor Balasaheb Agarkar, Associate Professor Sachin Chaudhari

Department of Electronics and Telecommunication Engineering, SRES's Sanjivani College of Engineering Kopargaon, India

Abstract- This review provides a comprehensive study of neural network-based perception pipelines for autonomous vehicles, concentrating on the accuracy-complexity trade-offs affecting Traffic Sign Recognition (TSR) and lane detecting systems. The survey outlines the progression from conventional vision algorithms to modern deep learning frameworks, emphasizing how Convolutional Neural Networks (CNNs) and analogous architectures have revolutionized robustness, feature discrimination, and real-time inference capabilities. In TSR, high-capacity architectures such as ResNet-101 and attention-augmented models demonstrate improved resilience to small-scale, occluded, and luminance-variable sign categories, while single-stage detection frameworks, exemplified by the YOLO family, offer latency-constrained inference suitable for embedded applications. Conversely, two-stage frameworks such as Faster R-CNN provide enhanced representational accuracy but entail much higher computational expenses, underscoring the intrinsic precision–efficiency trade-off characteristic of contemporary vision systems. This study broadens the inquiry into lane detection, where deep learning techniques encounter challenges related to environmental unpredictability, low-contrast features, occlusions, and sensor noise. Techniques employing structured ROI extraction significantly diminish unnecessary computation while preserving accuracy. The study combines algorithmic performance, computational constraints, and deployment viability to emphasize the importance of ongoing innovation in neural network topologies, as well as the provision of comprehensive and carefully annotated datasets for the unified approach of lane and traffic sign detection. Future research should focus on closing the gap between performance and the real time requirement of driver assistant systems, aiming for a fine mix of accuracy, efficiency, and reliability in dynamic real-world contexts.

Keywords— traffic sign detection, classification, lane detection, neural network, autonomous driving

I. INTRODUCTION

In the current era of advanced transportation infrastructure and increasing road safety concerns, driver-assistance technology have become extremely essential. Traffic sign and lane recognition systems, particularly those based on neural networks, are crucial for vehicle autonomy and the reduction of accidents. Conventional approaches, although crucial, occasionally struggle to address the dynamic and unexpected real-world driving conditions, including varying illumination, inclement weather, and obstacles[1].

The capability of neural networks to learn is employed in computer vision to interpret high-dimensional input, such as pictures, which is crucial for self-driving cars to precisely identify and classify things in real time. Traffic sign recognition and lane finding are also possible using neural network architectures like VGG16, AlexNet, and ResNet. Deep learning algorithms handle data collected from sensors in real time to improve autonomous driving safety and efficiency [2]. Due to more powerful graphics processing units and better deep learning models.

Deep learning for image recognition, real-time object tracking, and environmental perception improves autonomous driving systems by reframing object identification as a one-pass regression problem. You Only Look Once improved real-time application speed and accuracy. As proved by YOLOv2 and YOLOv3, deep learning can recognize objects in real time, which is critical for advanced driver assistance systems. Lane detection and traffic sign recognition depend on vision data evaluation speed and accuracy, which improve vehicle safety and navigation[3]. Architectural improvements, feature extraction improvements, and training methodologies have allowed the YOLO series, including YOLOv5, YOLOv7, and YOLOv8, to improve real-time object detection, increasing their applicability in dynamic autonomous driving contexts. Architectural improvements in YOLOv11 improve detection speed, precision, and robustness in complex situations, especially for vehicle detection, and edge device support is an essential advantage. Current research seeks to create more resilient algorithms capable of functioning effectively in diverse environmental conditions and employing multi-modal sensor fusion for enhanced perception to mitigate these limitations.

This study presents a detailed analysis of neural network-based perception pipelines necessary for autonomous vehicles, primarily focused on Traffic Sign Recognition and lane detecting systems. It carefully follows the evolution from traditional vision algorithms to advanced deep learning frameworks, demonstrating the revolutionary influence of Convolutional Neural Networks on real-time inference, resilience, and feature discrimination in various applications. A significant contribution of this work is a comprehensive review of accuracy-complexity trade-offs inherent in various models, from high-capacity architectures like ResNet-101 to efficient single-stage detectors like the YOLO family, alongside an in-depth study of both classification and detection performance across benchmark datasets. Furthermore, the paper delves into deployment challenges, such as hardware considerations and model optimization strategies, providing insights and recommended approaches for striking the right equilibrium between precision, computational efficiency, and reliability in evolving real-world autonomous driving scenarios.

Need

Maintaining lane position is one of the most common driving tasks, sometimes overlooked due to its regular nature. Following basic training, most can do this if not distracted, inebriated, or otherwise impaired. A computer has a much harder time "maintaining the vehicle within the lane markings" than a person. However, people can lose focus due to activities like adjusting the radio, talking to a passenger, exhaustion, or substance use, but computers cannot. Given its superior sustained focus, a computer could replace a driver if we could teach it to recognize lane lines. At least one car departing from its lane causes many collisions. Detecting and fixing lane deviations early could prevent crashes. Lane and traffic sign detection is vital when thinking about driverless vehicles, however in real life, the road is filled with roadside objects, the lane pattern changes, and the lane marking color changes, making it difficult to determine the lane using solely image processing algorithms.

II. LITERATURE REVIEW

With their ability to extract spatial features from images, convolutional neural networks are the preferred traffic sign identification architecture. Since developments have not yet made these systems resilient to occlusions or new signs, research into more resilient designs and training methods is necessary[1]. Deep architectures and advanced neural network designs can extract complex features from small or difficult traffic signals, improving detection and identification. Real-time implementation in constrained automotive applications requires balancing processing overhead and accuracy [1].

Strong lane recognition algorithms, particularly those that use deep learning, have enabled autonomous driving systems to progress significantly[2], [3]. Monocular lane recognition algorithms based on deep learning are a promising research field since they are both successful and can provide valuable 3D lane data for future applications.

1. Traffic Sign Classification

Traffic signs guide drivers about driving rules, regulations, and recommendations. For vehicles planning to assist or drive autonomously using a vision-based system in real life, it is crucial to locate and categorize traffic signs in the image to interpret road authorities' guidelines to complete the task safely. Unlike humans, computer-based visual systems struggle to multitask. Some of these issues include computational complexity and battery-operated processing capability, output generation time, ability to analyze partial data like broken sign boards or sunray illumination, etc. Researchers working on autonomous driving, driver assistant systems, and computer vision experts study traffic sign detection and classification due to its complexity and problems. CNN multi-task architectures can refine ROI and classify traffic indicators faster than sequential algorithms. Minimal computational and memory requirements make lightweight models like MobileNet and ShuffleNet viable for autonomous vehicle edge devices for real-time processing due to their low computational and memory needs. Embedded systems can use these small networks for fast, accurate inference. Models like MicronNet and SqueezeNet are also suited for resource-limited scenarios because they reduce parameter quantity and processing demands while maintaining recognition accuracy [4].

Colors classify traffic signs. Color eases and boosts image processing and traffic sign identification search space reduction. RGB, HSI, HSL, and HSV pixel-wise color threshold segmentation processes traffic sign colors. Establish traffic sign color ranges. Color probability models can increase traffic sign red, blue, or yellow while decreasing backdrop color s. Increases contrast for traffic signal detection [4], [5]. ResNet-101 CNN architectures extract complex features from small or problematic traffic signs. Recent research uses transformer-based designs and attention mechanisms to increase feature learning and contextual comprehension for small, obscured signals. The designs MobileNet, ShuffleNet, MicronNet, and SqueezeNet minimize parameter quantity and computing complexity while maintaining recognition precision.

Benchmark Dataset

Train and test traffic sign categorization models using datasets. These datasets differ in size, categorization, and image capturing configurations. The German Traffic Sign Recognition Benchmark (GTSRB) dataset, which contains 51,839 traffic sign images from 43 classes, is commonly used by researchers to standardize and compare network accuracy and computational complexity [4]. The Belgium Traffic Sign Recognition dataset is difficult because to its closely connected classes and shallow due to its limited number of training images. A reference table of traffic sign classification datasets from various scholars is presented in Table 1.

Table 1: Traffic Sign Classification Dataset

Dataset	Country	Classes	Total Images	Image Sizes (pixel)
(GTSRB) [4]	Germany	43	51,839	15×15 to 250×250
(TT-100K) [5]	China	211	16,000	Various
LISA TS	USA	49	7855	6 x 6 to 167 x 168
Malaysian TSR	Malaysia	68	10,000	Various
Indian TS	India	24	8,000	Various
Belgium TSD [4]	Belgium	62	7095	48 X 48

These datasets provide different traffic sign scenarios and categories for training, validating, and assessing recognition techniques.

Performance Comparison

Classification task of traffic sign is majorly validated over GTSRB dataset for the parameter likes Accuracy, FLOP's. Accuracy refers to the ability of a neural network to correctly classify traffic signs in a test dataset that the model has not encountered during training. Methods adopted can broadly be categorised into CNN model, Light weight CNN, Transformer based Model. CNN model focus of accuracy and generally not focused on utilisation of computational resources used, they were trained and executed on GPU based platform.

Table 2: Accuracy On Testing With Gtsrb Dataset With Details

Model	Accuracy (%)	Parameters(M)	FLOPs(M)
Improved CNN	98.9	--	--
Advance CNN	94.76	--	--
HOG-SVM	91.25	--	--
MicronNet	98.9	0.510	~10
MicronNet-BF	99.383	~0.46	~10
EfficientNet-B0	98.64	~5.3	~390
MobileNetV2	98.00–98.85	~3.4	~300
VGG 16	99.30	138	15.3
AlexNet	98.2 %	61	724
ViT	99.7 %	--	--

Few examples of such model are AlexNet, VGG16, MobileNet, EfficientNet[6]. Contrary to this Light weight CNN focuses on reduction of computational complexity while achieving the state of art accuracy, prime goal of these models were to be executable on edge devices. MicronNet, MobileNetV2, are Example of Light weight CNN. Transformer based model uses a Vision Transformers (ViT) kind of architecture as backbone and develop a remote system for classification, utilizes cloud infrastructure for the task. Comparison of these models were showcased in Table 2.

2. Traffic Sign Detection

Traffic sign detection is the task of locating traffic signs in road images. Localizing and anticipating traffic sign boundaries in an image is traffic sign detection. The You Only Look Once series (e.g., YOLOv1 through YOLOv11) are prime examples of single-stage detectors that balance speed and precision by treating detection as a one-pass regression problem. Researchers offer several YOLO-based methods. Architectural improvements and refined feature extraction make YOLOv5, YOLOv7, and YOLOv8 stand out for their speed and precision. Anchor-free detection and a C2F module for multi-scale object identification are in YOLOv8. This model is popular because it can identify traffic signs in real time faster than two-stage detectors with comparable accuracy. Using additional steps to refine proposals, two-stage detector approaches like Faster R-CNN improve accuracy. Often more precise, especially with small or occluded traffic signs, they nevertheless have higher computational complexity and may fail to meet autonomous driving system real-time needs due to their slower processing speed. Region-CNN versions refine computationally intensive. The fast and accurate YOLOv1-YOLOv8 approach is popular for real-time traffic sign identification [7].

Benchmark Dataset

Datasets are crucial for traffic sign detection training, validation, and evaluation. Since traffic signs differ widely among countries, the dataset chosen relies on the research subject and deployment region. Table 3 shows a few datasets.

Table 3 : Traffic Sign Detection Dataset

Dataset	Country	Classes	Total Images	Image Sizes
(GTSRB) [4]	Germany	4	900	1360 × 800
(TT-100K) [5]	China	211	10,000	Varied
Belgium TSD[4]	Belgium	62	1,506	640 × 480
LISA TS	USA	49	6610	640 × 480

Each traffic sign in these files is carefully identified as a rectangle detection region.

Performance Comparison

Recently, traffic-sign detection has focused on attention-enhanced architectures, small-object–oriented feature fusion, and lightweight inference pipelines spanning the YOLO family. Sign-YOLO, PSFE-YOLO, and STC-YOLO use squeeze-and-excitation (SE) blocks, pixel-wise spatial enhancement, and multi-head attention to improve feature discrimination and contextual encoding, achieving top performance on GTSDB.

Parallel research on small-object detection introduces hybrid CNN-transformer backbones, enhanced pyramid fusion strategies (e.g., BiFPN, SPP-Boost), and dilated or SPD-based down-sampling to preserve fine-grained spatial cues, resulting in consistent performance gains of up to 10% over baseline YOLO variants. The existence of lightweight, deployment-oriented models like YOLOv5S-A2, CEL-YOLO, improved YOLOv4-Tiny, and Improved YOLOv5 derivatives shows that parameter-efficient designs, shared detection heads, and depth-wise or multi-branch modules can provide real-time inference without compromising accuracy. Table 4 compares traffic sign detection performance across datasets.

Table 4: Traffic Sign Detection Performance On Various Dataset

Innovations	Model Name	Dataset	mAP / Accuracy (%)
Attention Mechanisms	Sign-YOLOv7	GTSDB	99.10 (mAP)
	Improved YOLOv7-Tiny	GTSDB	93.47 (mAP)
Enhanced Small Object Detection	STC-YOLOv8	GTSDB	96.1%
	BiFPN-YOLOv5	GTSDB	81%
	YOLOF-F	GTSDB	74.57 (AP)
	YOLO-SG	GTSDB	Not specified
	ETSR-YOLOv5s	TT100K, CCTSDB2021	Not specified
	SPP-Boosted YOLOv3	BTSD	99.28%
Lightweight and Real-Time	Improved YOLOv4	GTSDB	91.74 %
	Improved YOLOv5	GTSDB	92.5 (mAP)
	CEL-YOLOv8	CCTSDB2021	97.1
Feature Extraction	SSD	BTSD	83.5 %
	RetinaNet	BTSD	86.9 %

3. Traffic Sign Detection and Classification

Traffic sign detection and categorization must be regarded as a combined problem rather than as discrete issues if we aim to help drivers through a vision-based system, necessitating the clarity of the challenges posed by both aspects.

Traffic Sign Detection and Classification Model Comparison

Recent developments in traffic sign recognition (TSR) exhibit remarkable accuracy improvements, with numerous models over 96% on benchmark datasets such as GTSRB, BTSD, and CCTSDB. YOLOv8 establishes a premier benchmark with 98.0% accuracy on general TSR tasks, highlighting its effectiveness in high-performance contexts. Hybrid methodologies demonstrate superior performance: MobileNetV2, in conjunction with YOLOv5, achieves a 98.0% mean Average Precision (mAP) on the GTSRB and BTSD datasets, designed for real-time advanced driver-assistance systems (ADAS) according to a 2025 study. Additionally, color and shape preprocessing enhances MobileNetV2's accuracy to 99.5% on GTSRB by minimizing false positives using conventional Region of Interest (ROI) extraction. Likewise, color and shape preprocessing utilizing ResNet-18 achieves an accuracy of 97.5% on CCTSDB, effectively balancing computational economy and performance. Improving PANet in YOLO greatly increases mAP for small images [8].

Table 5: Traffic Sign Classification And Detection Model Comparison

Model Name	Dataset	mAP / Accuracy (%)
Sparse R-CNN	BTSD	96.40 %
MobileNetV2 + YOLOv5	GTSRB, BTSD	98.0 %
Color-proc + MobileNetV2	GTSRB	99.5 % (Acc)
Color Pre-proc + ResNet-18	CCTSDB	97.5 % (Acc)

Sparse R-CNN achieves 96.40% mAP on BTSD, improving detection speed and robustness, although encounters difficulties with small objects. The results indicate a transition towards lightweight, preprocessing-enhanced CNNs for effective, real-time TSR in autonomous vehicles [8] as shown in table 5.

Traffic Sign Detection and Classification Specific Challenges

Vegetation and vandalism can obscure traffic signs, making identification harder[9]. Small traffic signs, especially distant or low-resolution ones, are hard to detect. Model training may discard instances below a certain pixel count [10]. Warped signs and odd angles hinder recognition algorithms. Confused backdrops make traffic signs hard to see, Unbalanced datasets, Traffic Sign Variability affect training

4. Lane Detection

While reflective lanes may have different colors, lane markings are usually yellow and white. Light circumstances including sunlight, shadows, and low light can impact RGB color values. Color-detecting algorithms can be confused by inconsistent lighting. RGB values are commonly normalized or transformed to HSV or Ohta space to improve illumination resilience [11].

Various color processing approaches tackle these issues. Pixel colors can be converted to YCbCr from highly correlated RGB. This is useful since the human visual system is less sensitive to color but retains important visual information in the Y (luminance) component.

Region of Interest

Regions of Interest are examined by lane detection systems. Avoids thorough image examination. Lane-lined regions benefit from this strategy. The lane line data is typically in the lower image because the road is usually ahead of the car. Some ROI calculations use the bottom 25% of the image with trapezoid mask for simplicity and efficiency.

Lane Detection Specific Challenges

Guardrails, pavement markings, road dividers, road cracks, and vehicle lines may be misconstrued as lane markers, causing false positives. minimal contrast: Lane markers and road surfaces can have minimal contrast if the lane is degraded, colored differently, or the road illumination matches the lane color. Limitations of sensors Wipers, droplets, and windshield reflections can distort views in heavy rain or snow.

Lane Detection Dataset

Table 6 summarizes prominent lane detection datasets, including lane color, classes, number of photos, image size, and other criteria. Lane markings are usually white and yellow, however datasets may not specify them.

TABLE 6 : LANE DETECTION DATASET

Dataset	Frames	Resolution	Road Types
TuSimple	6,408	1280 x 720	Highway
CULane	1,33,235	1640 x 590	Urban, rural, highway
BDD100K	1,00,000	1280 x 720	

Performance Comparison of Lane Detection

Recent advancements in lane recognition reveal distinct trade-offs among real-time speed, accuracy, and durability. LaneNet emphasizes reduced computational resources through a two-stage pipeline comprising pixel-wise lane edge proposals and lane line localization [12]. It possesses a processing speed of 62.6 FPS and a success rate of 96.38% on TuSimple. It functions effectively across diverse road conditions without presuming parallel lanes [13]. Recently presented LC-YOLOv8, a comprehensive detection-segmentation system enhanced by Large Separable Kernel Attention and Coordinate Attention. It attained 98.30% accuracy and 97.9% mean Average Precision (mAP) on TuSimple, with a mAP of 69.4%, and was engineered for real-time applications. This indicates the increasing trend towards lightweight and precise end-to-end lane detecting systems tabulated in table 7.

Table 7 Comparative Of Lane Detection Results

Model Name	Dataset	mAP / Accuracy	FPS
Ultra-Fast Structure-aware CNN	TuSimple, CULane	95.87%	322.5
LC-YOLOv8	Custom Dataset	90.6	117
ENet-21	TuSimple	N/A	N/A

LaneNet	TuSimple	96.38	250 FPS

5. Simultaneous Lane and Traffic sign detection

Processing traffic sign detection, classification, and lane detection presents several significant challenges like the inherent complexity of visual data, and the computational demands of advanced algorithms [14]. Computational complexity and real-time performance Real-time autonomous vehicle applications must balance speed and accuracy. Due to slower processing speeds, two-stage detection may not meet real-time needs despite higher accuracy. Durability and Generality The development of algorithms that work under various environmental, sign, and road circumstances is ongoing.

6. Hardware Platform

Traffic sign classification systems in autonomous vehicles require careful hardware platform selection for deep neural network computing and real-time performance. Deep learning, especially Convolutional Neural Networks, takes a lot of processing power and memory, which limits real-time behavior in resource-limited automotive applications. GPUs are poor for on-board vehicle use because to their power and heat consumption. High-performance GPUs increase processing speed but also computational expenses. Speed and precision must balance for real-time autonomous vehicles. Recognition accuracy is maintained via architectures that reduce parameter count and computational complexity. Cataloguing examples include MobileNet, ShuffleNet, MicronNet, and SqueezeNet. MobileNetV2 and a plug-and-play edge device are used in some systems to recognize traffic signs immediately. Quantization, which compresses neural networks by decreasing the bit representation of weights and activations, may impair accuracy but reduce computational and memory needs. Pruning decreases superfluous connections or redundancy in convolutional layers, reducing model size and inference speed by reducing parameters. Convolution Factorization breaks large filters (5x5, 7x7) into smaller convolutions (3x3, 1x1) to minimize processing and model size. Continuous development attempts to balance precision, computational efficiency, and memory consumption for real-time applications in varied operational settings [4].

III. DISCUSSION

1. Challenges Driving Hardware Selection

Hardware for real-time car traffic-sign and lane-detection systems must balance accuracy, latency, power, size, and cost. ADAS and autonomous driving require low-latency, high-throughput inference for safety-critical decisions. Implementable on-board systems must be energy-efficient, compact, thermally controlled, and cost-effective . GPUs like the NVIDIA Tesla and TITAN series speed up model training and testing, but their size, power consumption, and cost make them unsuitable for in-vehicle deployment.For on-board inference, Jetson-class modules, Google Coral accelerators, and mobile SoCs offer high throughput, energy efficiency, and form factor. Low-power, high-efficiency FPGAs, DSPs, and ASPs are specialized accelerators for small network models. In automobiles, offloading or hybrid on-vehicle and proximate edge processing reduces computing strain but increases latency and system complexity .

2. Model and Implementation Optimizations

Optimize performance by limiting parameters and operations with lightweight, edge-ready networks like YOLOv5-Lite[19]. Quantization, pruning, convolution factorization, and knowledge distillation reduce model size and speed inference without compromising accuracy. When possible, combine detection and classification in a single-stage network to reduce latency and computation. Match model widths, configurations, and memory architectures to the target accelerator for throughput and energy efficiency.

3. Simultaneous Detection and Classification

To reduce latency and bandwidth consumption, use single-pass networks that create boundary geometry and class labels simultaneously. The unified model should be optimized and structured to meet real-time, power, and memory requirements of the chosen embedded accelerator. Use specialized accelerators (e.g., TPUs, Jetson modules, FPGAs) for energy-efficient parallelism in integrated tasks [15].

4. Recommended approaches
Simplify sensors and computers with accurate, reliable single-camera systems. Make a GPU-accelerated profile and prototype. To improve frame rates and accuracy, quantize models for embedded targets such automotive SoCs, Jetson modules, and FPGAs . Train and profile with strong GPUs and compact accelerators (Jetson, Coral, automotive SoCs, FPGAs). Use hardware-aware architectures with quantization, pruning, factorization to meet latency, power, and memory goals. Unified detection/classification networks tuned to the deployment accelerator improve latency [15].

IV. CONCLUSION

Literature research indicates that neural network methodologies, especially those utilizing Convolutional Neural Networks, have markedly improved traffic sign recognition and lane detection in autonomous vehicles. It delineates several datasets essential for training and evaluation purposes, including GTSRB, GTSDB, Tsinghua–Tencent 100K, TuSimple, and CULane, for traffic sign categorization, detection, and lane detection. Models such as VGG 16 attain an accuracy of 99.30%, whereas MicronNet-BF achieves 99.383% on the GTSRB dataset for the classification task, exhibiting significant variation in the number of parameters. YOLO-based models aim for an optimal equilibrium between speed and accuracy, surpassing benchmarks such as SSD and RetinaNet, particularly in addressing class imbalance and detecting minor traffic sign occurrences. Conversely, Enhanced SCNN-based models demonstrate a high accuracy of 98.21% in tackling issues such as damaged marks and occlusions. Literature also discusses the challenges posed by occlusions, minor indicators, fluctuating brightness.

Consolidating detection and classification into single-stage networks reduces end-to-end latency and redundant processing, making them perfect for embedded accelerators. Multi-modal sensor fusion to improve perception and overcome constraints. In resource-constrained automotive applications, real-time deployment requires ongoing innovation to balance processing overhead and accuracy.

REFERENCES

1. L. Mussone and O. el Hassan, "An analysis of pedestrian crossings through deep learning models and crash data," Transp Res Interdiscip Perspect, vol. 31, May 2025, doi: 10.1016/J.TRIP.2025.101449.
2. K. Ishihara, A. Kanervisto, J. Miura, and V. Hautamäki, "Multi-task Learning with Attention for End-to-end Autonomous Driving," arXiv preprint arXiv:2104.10753, Jan. 2021, doi: 10.48550/arxiv.2104.10753.
3. W. Choi, M. Shin, H. Lee, J. H. Cho, J. H. Park, and S. Im, "Multi-task Learning for Real-time Autonomous Driving Leveraging Task-adaptive Attention Generator," arXiv preprint arXiv:2403.03468, Mar. 2024, doi: 10.48550/arxiv.2403.03468.
4. R. Ventura, S. Roussou, A. Ziakopoulos, B. Barabino, and G. Yannis, "Using computer vision and street-level videos for pedestrian-vehicle tracking and behaviour analysis," Transp Res Interdiscip Perspect, vol. 30, Mar. 2025, doi: 10.1016/J.TRIP.2025.101366.
5. J. Duan, H. Ye, H. Zhao, and Z. Li , "Deep Cascade AdaBoost with Unsupervised Clustering in Autonomous Vehicles," Electronics (Switzerland), vol. 12, no. 1, Jan. 2023, doi: 10.3390/ELECTRONICS12010044.
6. M. Tan and Q. V Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proceedings of the 36th International Conference on Machine Learning (ICML), 2019, pp. 6105–6114. [Online]. Available: https://arxiv.org/abs/1905.11946
7. S. Saxena, S. Dey, M. Shah, and S. Gupta, "Traffic sign detection in unconstrained environment using improved YOLOv4," Expert Syst Appl, vol. 238, Mar. 2024, doi: 10.1016/J.ESWA.2023.121836.
8. [35] R. Zhang, K. Zheng, P. Shi, Y. Mei, H. Li, and T. Qiu, "Traffic Sign Detection Based on the Improved YOLOv5," Applied Sciences

(Switzerland), vol. 13, no. 17, Sep. 2023, doi: 10.3390/APP13179748.
9. T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection," IEEE Trans Pattern Anal Mach Intell, vol. 42, no. 2, pp. 318–327, Aug. 2017, doi: 10.1109/TPAMI.2018.2858826.
10. T. Liang, H. Bao, W. Pan, and F. Pan, "Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles," J Adv Transp, vol. 2022, 2022, doi: 10.1155/2022/3825532.
11. Y. Yang, H. Luo, H. Xu, and F. Wu, "Towards Real-Time Traffic Sign Detection and Classification," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 7, pp. 2022–2031, Jul. 2016, doi: 10.1109/TITS.2015.2482461.
12. W. J. Yang, Y. T. Cheng, and P. C. Chung, "Improved Lane Detection with Multilevel Features in Branch Convolutional Neural Networks," IEEE Access, vol. 7, pp. 173148–173156, 2019, doi: 10.1109/ACCESS.2019.2957053.
13. S. Sultana, B. Ahmed, M. Paul, M. R. Islam, and S. Ahmad, "Vision-Based Robust Lane Detection and Tracking in Challenging Conditions," IEEE Access, vol. 11, pp. 67938–67955, 2023, doi: 10.1109/ACCESS.2023.3292128.
14. X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial CNN for traffic scene understanding," 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, pp. 7276–7283, 2018, doi: 10.1609/AAAI.V32I1.12301.
15. Z. Qin, P. Zhang, and X. Li, "Ultra Fast Deep Lane Detection with Hybrid Anchor Driven Ordinal Classification," arXiv preprint arXiv:2206.07389, Jan. 2022, doi: 10.48550/arxiv.2206.07389.

1.