Algerian Journal of Research and Technology
https://doi.org/10.58681/ajrt.25090102
A.J.R.T Volume 9 (N°1) (2025) 12 – 22

A Concise Overview of Vehicle Detection Techniques

Zentar Mohamed Dhia El Hak, Bencheriet Chemesse ennahar

Laboratoire d'automatique et Informatique de Guelma, Université 8 Mai 1945 de Guelma Guelma, Algeria

*Corresponding author. *mohameddhiaelhak.zentar842@gmail.com*; *cbencheriet@yahoo.fr*

Received. June 17, 2024. Accepted. December 06, 2024. Published. December 20, 2024.

DOI: https://doi.org/10.58681/ajrt.25090102

Abstract. Vehicle detection, a specialized subset of object detection, has gained significant importance in recent years, particularly in the realms of autonomous and assisted driving technologies. This field, while promising, grapples with several challenges including occlusion, scalability issues, and the complexity of real-world backgrounds. This paper sets out to provide a summary of recent state-of-the-art advancements in vehicle detection technology. First, it organizes vehicle detection approaches into three primary categories: classical methods, deep learning techniques, and hybrid approaches that combine elements of both. Within the deep learning category, the paper further distinguishes three subcategories: anchor-based methods, anchor-free methods, and attention-based techniques. Each of these approaches offers unique advantages and addresses different aspects of the vehicle detection challenge. Secondly, it provides a literature review of different papers on vehicle detection.

Keywords. Vehicle detection, Deep neural networks, Traffic surveillance, Object detection.

INTRODUCTION

Recent years have seen remarkable advancements in artificial intelligence, with significant impacts on computer vision and, notably, vehicle detection technologies. The ability to accurately identify vehicles is crucial across various domains, including intelligent transportation systems, self-driving vehicles, and driver assistance platforms. For practical applications like driver assistance, vehicle detection systems must deliver both precision and speed to safeguard all road users. One persistent challenge in this field is dealing with occlusions, which are particularly prevalent in busy urban environments.

This review aims to offer a comprehensive look at cutting edge methods currently employed in vehicle detection. By examining these state-of-the-art techniques, we seek to provide

insight into the current landscape of detection technologies and their capabilities in addressing real-world challenges.
Generally speaking, vehicle detection system architecture consists of: A training phase and a testing phase (Fig. 1).
Usually, in the training phase, the inputs are a collection of images of vehicles and their labels. These images undergo a pre-processing stage which includes operations like resizing, normalization, noise reduction...etc. This step creates uniform input data, establishing a consistent format that enables effective learning in subsequent stages.

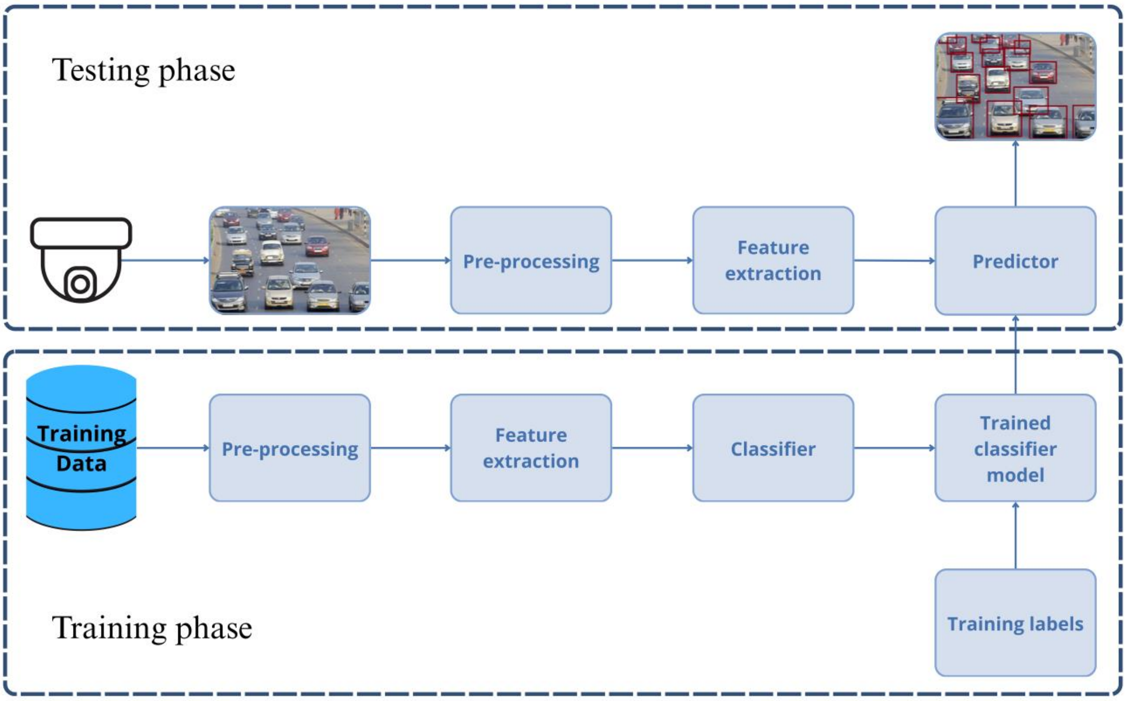

Fig. 1. General Vehicle Detection Framework.

After pre-processing, comes feature extraction in order to identify significant visual elements from the images. These elements typically include edges, shapes, and textures and the characteristics that distinguish vehicles from their surroundings. The extracted features are then fed into a classifier, which is trained using the labelled data. During this phase, the classifier learns patterns and associations between the features and the corresponding labels. As a result, a trained classifier model is developed, capable of accurately identifying vehicles in new, unseen data.
In the testing phase, the trained model is deployed to analyse real-world inputs. Camera-based sensor systems capture images or video containing vehicles, which are then passed through the same pre-processing pipeline to ensure consistency with the training data. The standardized images are subjected to feature extraction, where relevant attributes are isolated. These features are then passed to a predictor, which utilizes the trained classifier model to analyse the data. The predictor determines whether vehicles are present in the input and, if so, localizes them by drawing bounding boxes around each detected vehicle in the scene.
The primary goal of data collection in the system mentioned is to enhance safety. However, some individuals may perceive it as an invasion of privacy. Bloom et al. (2017) revealed that a significant number of people expressed discomfort with the idea of data being collected. In modern systems, data processing often involves sharing information with third parties, such as traffic management systems, which raises further privacy concerns.

Fortunately, techniques like Transport Layer Security (TLS), Differential Privacy (DP), and K-anonymity are already being used to address this issue (Allen and Dierks , 1999 ; Dwork and Roth, 2013; Wang et al., 2020). Another concern is data retention.
According to the European Data Protection Supervisor (2019a), there have been cases where users of car rental services were able to access and control vehicle systems, underscoring the importance of limiting data storage. Data should not be retained indefinitely; the European Data Protection Supervisor (2019b) recommends establishing a defined retention period for stored information. Similarly, the European Data Protection Board (EDPB) advocates for mechanisms that allow users to delete their personal data (EDPB, 2021). A further solution involves adopting a privacy-by-design approach, which emphasizes data minimization—particularly in applications such as pedestrian detection—by collecting only what is strictly necessary (EDPB, 2021).
From an ethical perspective, companies must ensure transparency with all stakeholders regarding the types of data collected, the purposes for collection, how the data is processed, and who has access to it (European Data Protection Supervisor, 2019b). Clear, visible notifications—such as stickers on vehicles to inform pedestrians that they are being recorded—can also contribute to greater awareness and accountability (Krontiris et al., 2020). Nonetheless, despite these efforts, current legislation remains insufficient and fails to address the full range of potential scenarios (IEEE Spectrum, 2024).

CLASSIFICATION OF PEDESTRIAN DETECTION METHODS

Vehicle detection techniques can be categorized into three main branches: Classical Methods, Deep Learning Methods, and Hybrid Methods (Fig. 2).

Classical Methods

This category includes traditional computer vision methods that use hand-crafted features and conventional algorithms, which were common before deep learning became prominent.
Techniques like Histogram of Oriented Gradients (HOG), Support Vector Machines (SVM), and other feature-based approaches are part of this group. While these methods established the foundation for vehicle detection, they often struggle with complex scenarios compared to modern techniques.

Deep learning Methods

Deep learning techniques have revolutionized vehicle detection by enabling the development of more advanced and robust models. These approaches can generally be categorized into three main types:

1) **Anchor-Based Methods**: These techniques rely on predefined anchor boxes to predict bounding boxes around vehicles. They are typically divided into:

- **One-stage detectors**, such as *You* Only Look Once (YOLO) (Redmon et al., 2016) and *Single Shot Detector* (SSD) (Liu et al., 2016; Chen et al., 2022; Cao et al., 2020), which perform detection in a single step. These models offer faster inference but may compromise on accuracy.
- **Two-stage detectors**, such as Fast Region-based Convolutional Neural Network (Fast R-CNN) (Girshick, 2015; Arora et al., 2022), which first generate region proposals and then classify them. This approach typically achieves higher accuracy at the cost of increased computational complexity.

2) **Anchor-Free Methods**: These models eliminate the need for predefined anchor boxes by directly predicting object locations from the image. Notable examples include *CenterNet* (13; 34) and Fully Convolutional One-Stage Object Detection (FCOS)

(Tian et al., 2019; IEEE Xplore, 2025a). Anchor-free methods are often simpler and can achieve faster inference with competitive performance.

3) **Attention-Based Methods**: These methods incorporate attention mechanisms to enhance detection performance by focusing on the most informative regions of the image. Attention improves feature representation and is particularly effective in complex scenes. Prominent examples include the Detection Transformer (DETR) (Carion et al., 2020; S. P. & Mohandas, 2023) and Vision Transformers (ViT) (Dosovitskiy et al., 2020).

Hybrid Methods

These methods aim to build robust systems by integrating classical and deep learning techniques. They strive to balance traditional feature extraction with modern deep learning models, leveraging the strengths of both approaches while addressing their limitations.

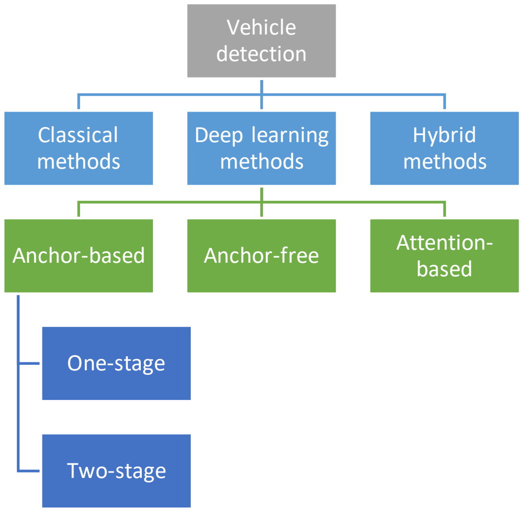

Fig. 2. Categorisation of different vehicle detection techniques.

Here is a summary of some state-of-the-art papers:

Li et al. (2023)**:** This paper addresses the issue of low detection accuracy in vehicle and pedestrian detection models by incorporating a Convolutional Block Attention Module (CBAM) into the cross-stage partial Darknet-53 (CSPDarknet53)-tiny module to enhance feature extraction capabilities and mitigate the limitations of using a single attention module. Additionally, the original simple convolutional module is replaced with a Cross Stage Partial Dense Block Layer (CSP-DBL) to better preserve high-resolution features and improve detection accuracy. The public BDD100K dataset (BDD100K, 2024) is used for evaluation, employing average precision (AP), mean average precision (mAP), and recall as performance metrics. The proposed model achieves the highest precision and an mAP of 88.74%. The results also show improvements in recall by 0.73% for cars and 0.01% for people. However, the model records the lowest speed, with only 63 FPS.

Xiong et al. (2023)**:** This paper presents an improved lightweight YOLOX real-time vehicle detection algorithm that enhances both detection speed and accuracy while reducing parameter count. It introduces a lightweight backbone feature extraction network and a new α-Complete Intersection-over-Union (α-CIoU) loss function to improve regression accuracy and convergence. Inspired by GhostNet, two new modules are proposed:

- The Cross Stage Partial Ghost Module (CSPGhost Module or CSPGM)**,** which splits input feature maps into two parts—one passes through stacked Ghost modules and the other merges via a cross-stage pathway.

- The CGM structure, which also splits input features into two branches, adjusts the output channel on one side and uses a derived feature map on the other before combining them.

The α-CIoU loss function retains CIoU's benefits but emphasizes high-IoU targets, optimizing detection frame regression without increasing inference time. The modified model is evaluated on the BIT-Vehicle dataset (Dong et al., 2015), which includes 9,850 images across categories like Bus, Microbus, Minivan, Sedan, SUV, and Truck. Metrics used include parameter count, model size (MB), FPS, and mAP@0.5. The results show a 0.99% improvement in mAP, a 41.2% reduction in parameters, and a 12.7% increase in FPS compared to the original YOLOX-S.

S. P. and Mohandas (2023)**:** The authors address DETR's limitations in detecting small features and its slow training convergence. To enhance performance, they reduce the number of layers to 40, use a shortcut layer instead of max-pooling, and introduce a Spatial Pyramid Pooling (SPP) block. The model employs a modified ResNet-50 for feature extraction, followed by a transformer with a multi-head self-attention encoder-decoder and a feedforward network (FFN) for end-to-end detection. It is trained on the MS COCO 2017 dataset (Papers with Code, 2024a) and evaluated using a custom video dataset from KELTRON (KELTRON, 2024) for vehicle detection. Evaluation metrics include precision, recall, mAP, FLOPS, and FPS. The model achieves a mAP of 51.31% on MS COCO 2017, outperforming SSD, YOLOv3-tiny, and the baseline DETR, but records the lowest FPS at 53, highlighting the need for speed optimization in real-time applications. Additionally, it shows a significant mAP improvement of 0.03 in the Wilcoxon test.

Hong et al. (2020)**:** This study tackles the problem of detecting multi-scale vehicle targets—particularly small ones—in traffic surveillance videos. It proposes a codec-based vehicle detection algorithm built upon YOLOv3. The method introduces a new multi-level feature pyramid that integrates a codec module to better detect vehicles across various scales. Multi-level features from the backbone are stitched into basic features, passed through the codec module, and merged with equivalent-scale decoder features for final detection. This enhances YOLOv3's capability in vehicle detection, especially for small targets.

Table 1. State Of The Art

Paper	**Architecture**	**Dataset**	**Results**
Li et al. (2023)	YOLOv4 tiny*	BDD100K (2024)	- Achieves an 88.74% mAP. A lower speed of 63 FPS.
Xiong et al. (2023)	YOLOX-based	BIT-Vehicle (Dong et al. , 2015)	- 0.99% increase in mAP. 41.2% reduction in parameters. - 12.7% increase in FPS compared to the original YOLOX-S.
S P and Mohandas (2023)	DETR-SPP	MS COCO 2017 (Papers with Code , 2024a) KELTRON (2024)	- Achieved 51.31% mAP, outperforming SSD, YOLO V3 tiny and baseline DETR.
Hong et al. (2020)	YOLOv3-based	KITTI (Geiger et al. 2013)	- On KITTI, the algorithm achieved average precisions of 95.04%, 92.39%, and 87.51% for

		-UA-DETRAC (Papers with Code, 2024b)	easy, moderate, and hard subsets. - On UA-DETRAC, the algorithm significantly improved over YOLOv3 across all detection conditions.
Hao et al. (2023)]	CPAM Network	-MS COCO (Lin et al. , 2014) -UA-DETRAC (Wen et al., 2020)	- Outperformed most of the other detectors on both datasets.

In traffic surveillance videos. The key innovations of this article can be sited as follow:

- A YOLOv3 Integration: The algorithm introduces the YOLOv3 algorithm to multi-scale vehicle detection in traffic videos and improves upon it.
- Feature Encoding and Decoding Structure Module: A module is proposed to generate high-order multi-scale feature maps through a simple U-shaped structure.
- Attention Mechanism: A special diagnosis module integrated with an attention mechanism enhances the model's expression ability.

The evaluation of the proposed method is conducted using two main datasets: the KITTI dataset (Geiger et al., 2013), which was captured while driving in the rural areas surrounding Karlsruhe—including the city itself and nearby highways—and the UA-DETRAC dataset (Geiger et al., 2013), a real-world benchmark for multi-object detection and tracking. UA-DETRAC comprises 10 hours of video footage recorded with a Canon EOS 550D camera across 24 distinct locations in Beijing and Tianjin, China. The videos were recorded at 25 frames per second (fps) with a resolution of 960 × 540 pixels.

As for evaluation metrics, average precision (AP) is used to assess the performance of the proposed architecture.

- On the KITTI dataset, the results show that the proposed algorithm achieves an AP of 95.04%**,** 92.39%, and 87.51% for the easy, moderate, and hard subsets, respectively—outperforming YOLOv3 by 2.49%**,** 3.68%**,** and 9.73% in each subset while maintaining competitive speed relative to other models.
- On the UA-DETRAC dataset, the proposed algorithm also demonstrates significant improvements over YOLOv3 across all detection conditions: easy, medium, hard, full, sunny, rainy, night, and cloudy.

Hao et al. (2023)**:** This paper addresses the challenge of multi-target vehicle detection in intelligent transportation systems (ITS), with a specific focus on detecting small and distant vehicles. To that end, the authors propose a network called **Corner Pooling with Attention** Mechanism (CPAM), which enables anchorless detection. The main contributions of the CPAM network include:

- **Hourglass with Coordinate Attention (Hourglass-CA)** as the backbone: Based on Hourglass-104 but reduced to 54 layers to lighten the network and increase speed. Three collaborative attention mechanisms are introduced into the decoder to extract key information at three feature scales: 384, 384, and 256.
- **Multi-Level Attention Network (MLA)**: Designed to enhance feature maps generated by the backbone using attention mechanisms at multiple scales, this module improves detection accuracy for vehicles of various sizes—especially smaller ones.

- **Multi-Level Attention Loss Function**: Calculates the discrepancy between predicted and ground-truth attention maps, allowing the network to prioritize relevant features and correct deviations during the upsampling process.

The model is evaluated on two major datasets: MS COCO (Lin et al., 2014), which includes three vehicle types (car, bus, truck), and UA-DETRAC (Wen et al., 2020), which emphasizes detection under occlusion.

Evaluation results:

- On UA-DETRAC, the architecture achieves a mAP of 70.64%, outperforming detectors such as Faster R-CNN, YOLOv3, CenterNet, and CornerNet. The model scores 90.72%, 74.12%, and 52.94% in the easy, medium, and hard subsets, respectively. It also performs well across weather conditions, with a mAP of 76.16% (cloudy), 78.62% (sunny), and 59.37% (rainy).
- On MS COCO, it achieves an AP of 43.3%**,** AP50 of 59.2%**,** AP75 of 46.9%**,** APs of 24.4%**,** APm of 44.8%, and APl of 57.5%**.**

CRITICAL ANALYSIS

Current research demonstrates a clear shift toward combining multi-scale fusion, attention mechanisms, and lightweight backbones to improve both accuracy and efficiency. Traditional HOG-based approaches continue to benefit from modern refinements. For instance, standard HOG achieves 93% accuracy but suffers from errors due to inaccurate hypothesis generation (Cheon et al., 2012); enhanced HOG reaches 97% accuracy with near real-time performance, though it remains somewhat computationally demanding (Niknejad et al., 2012). Region-driven HOG (RDHOG) further boosts accuracy to over 99% on traffic footage but proves ineffective in complex scenes with heavy occlusions (Wu et al., 2014).

Meanwhile, Haar-like cascades achieve sub-5ms detection per window but fail in occluded or high-density traffic scenarios (IEEE Xplore, 2025b). Even deformable part models (DPM) improve through PCA-based filter compression and FFT-accelerated convolutions—reducing parameters by 30% and accelerating matching—yet continue to struggle with overlapping objects (Ma and Xue, 2024). Hybrid models such as Dense-ResNet architectures, which combine residual and dense connections, outperform YOLOv3 by more than 5 AP on small and medium vehicles but require two to three times more memory and approximately 50% longer training time (Sun et al., 2019).

Two-stage frameworks like the Improved Region-based Convolutional Neural Network for Vehicle Detection (IRCNN-VD) eliminate background pixels using SIFT, incorporate hard negative mining, and leverage evolutionary hyperparameter tuning. This approach achieves 0.85 mAP on the BOXY dataset (Behrendt, 2019) in under 1 ms—twice the speed of Faster R-CNN—though it incurs high computational cost and lacks robustness in adverse weather conditions (Djenouri et al., 2022). Similarly, Retinex preprocessing combined with a NAS-optimized ResNet101 backbone and IoU-guided anchors boosts UA-DETRAC mAP from 62.13% to 68.25%, and small-vehicle AP from 14.16% to 43.64%, but operates at less than 2 FPS (Luo et al., 2021).

Attention-enhanced architectures show further promise. As demonstrated by Li et al. (2023), incorporating CBAM into CSPDarknet53-tiny along with CSP-DBL yields 88.74% mAP on BDD100K, with improved recall for cars and pedestrians, while maintaining 63 FPS. In another study, a streamlined YOLOX-S variant reduces parameters by 41.2%, increases mAP by 0.99%, and improves FPS by 12.7% on the BIT-Vehicle dataset (Xiong et al., 2023). Transformer hybrids also deliver strong results: DETR-SPP, which integrates spatial pyramid pooling and a reduced ResNet-50 backbone, improves MS COCO mAP to 51.31%, outperforming SSD-YOLOv3 tiny, albeit at a lower speed of 53 FPS (S P and Mohandas, 2023).

In more challenging visibility scenarios, a Swin Transformer adaptation for hazy conditions achieves 91% AP on a custom Haze-Car dataset and 82.3% on the Real Haze-100 dataset, though at a modest cost to speed (Sun et al., 2022). A codec-based, multi-level feature pyramid combined with attention mechanisms and integrated into YOLOv3 delivers 95.04% AP on the KITTI dataset and significant improvements on UA-DETRAC under easy, medium, and hard settings (Hong et al., 2020). Finally, the anchor-free CPAM network—featuring an Hourglass-CA backbone, multi-level attention modules, and a specialized attention loss function—achieves 70.64% mAP on UA-DETRAC and maintains strong AP scores under various weather conditions, along with competitive results on MS COCO (Hao et al., 2023).

Based on these findings, it is clear that traditional approaches offer low complexity and real-time performance but struggle in occlusion-rich or high-density environments. In contrast, deep learning-based methods significantly improve accuracy and small-object detection but introduce higher computational demands and reduced robustness under adverse conditions. Furthermore, attention-enhanced and transformer-based models advance detection under occlusion, multi-scale, and low-visibility scenarios, but require greater processing time and more powerful hardware.

This highlights the importance of developing context-specific solutions rather than aiming for a universal vehicle detection model.

FEATURE DIRECTIONS

As previously discussed, occlusion remains one of the primary challenges that vehicle detection models must address. Real-world environments are highly dynamic, ranging from sunny to rainy, foggy, and other adverse weather conditions that often obscure vehicle visibility. These environmental variations significantly affect a model’s ability to accurately detect and recognize vehicles. At present, most algorithms are optimized for specific settings, and no universal solution has been developed that can adapt effectively across diverse conditions. This limitation underscores the importance of designing a unified framework that integrates multiple detection strategies to ensure robustness in varying weather scenarios (Berwo et al., 2023).

Beyond environmental variability, achieving a balance between detection speed and accuracy also presents a significant design challenge. In most cases, improvements in one metric tend to come at the expense of the other, which can diminish the system’s overall robustness in real-world applications (Berwo et al., 2023).

Given that the backbone of a deep learning model plays a crucial role in its performance, much of the current research is directed toward designing more advanced and efficient backbone architectures. These efforts aim to facilitate the delicate balance between speed and accuracy, which remains central to the development of high-performance vehicle detection systems (Berwo et al., 2023).

CONCLUSION

Vehicle detection is a vital area in computer vision, playing a key role in enhancing technologies like driver assistance systems. Despite challenges such as occlusion, varying object scales, and complex backgrounds, researchers are still continuing to try to develop more advanced and robust detection methods. While relatively significant progress has been made, the field still offers many opportunities for further research and innovation.

REFERENCES

Allen, C., & Dierks, T. (1999). The TLS Protocol Version 1.0 (RFC 2246). Internet Engineering Task Force. https://doi.org/10.17487/RFC2246

Arora, N., Kumar, Y., Karkra, R., & Kumar, M. (2022). Automatic vehicle detection system in different environment conditions using fast R-CNN. Multimedia Tools and Applications, 81(13), 18715–18735. https://doi.org/10.1007/s11042-022-12347-8

Behrendt, K. (2019). Boxy Vehicle Detection in Large Images. In 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) (pp. 840–846). https://doi.org/10.1109/ICCVW.2019.00112

Berwo, M. A., et al. (2023). Deep Learning Techniques for Vehicle Detection and Classification from Images/Videos: A Survey. Sensors, 23(10), 4832. https://doi.org/10.3390/s23104832

Bloom, C., Tan, J., Ramjohn, J., & Bauer, L. Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles.

Cao, J., et al. (2020). Front Vehicle Detection Algorithm for Smart Car Based on Improved SSD Model. Sensors, 20(16). https://doi.org/10.3390/s20164646

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. In Computer Vision – ECCV 2020 (Vol. 12346, pp. 213–229). Springer International Publishing. https://doi.org/10.1007/978-3-030-58452-8_13

Chen, Z., et al. (2022). Fast vehicle detection algorithm in traffic scene based on improved SSD. Measurement, 201, 111655. https://doi.org/10.1016/j.measurement.2022.111655

Data Download — BDD100K documentation. (2024). Retrieved July 27, 2024, from https://doc.bdd100k.com/download.html

Djenouri, Y., Belhadi, A., Srivastava, G., Djenouri, D., & Chun-Wei Lin, J. (2022). Vehicle detection using improved region convolution neural network for accident prevention in smart roads. Pattern Recognition Letters, 158, 42–47. https://doi.org/10.1016/j.patrec.2022.04.012

Dong, Z., Wu, Y., Pei, M., & Jia, Y. (2015). Vehicle Type Classification Using a Semisupervised Convolutional Neural Network. IEEE Transactions on Intelligent Transportation Systems, 16(4), 2247–2256. https://doi.org/10.1109/TITS.2015.2402438

Double FCOS: A Two-Stage Model Utilizing FCOS for Vehicle Detection in Various Remote Sensing Scenes | IEEE Journals & Magazine | IEEE Xplore. (2025). Retrieved April 10, 2025, from https://ieeexplore-ieee-org.sndl1.arn.dz/document/9793845

Dosovitskiy, A., et al. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. https://doi.org/10.48550/ARXIV.2010.11929

Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., & Tian, Q. (2019). CenterNet: Keypoint Triplets for Object Detection. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 6568–6577). https://doi.org/10.1109/ICCV.2019.00667

Dwork, C., & Roth, A. (2013). The Algorithmic Foundations of Differential Privacy. Foundations and Trends® in Theoretical Computer Science, 9(3–4), 211–407. https://doi.org/10.1561/0400000042

European Data Protection Board. (2024). EDPB Guidelines on Connected Vehicles. Retrieved August 15, 2024, from https://www.edpb.europa.eu/system/files/2021-03/edpb_guidelines_202001_connected_vehicles_v2.0_adopted_en.pdf

European Data Protection Supervisor. (2019). TechDispatch: connected cars. Issue 3. Publications Office of the European Union. https://data.europa.eu/doi/10.2804/70098

European Data Protection Supervisor. (2024). Resolution on Data Protection in Automated and Connected Vehicles. Retrieved August 15, 2024, from https://www.edps.europa.eu/sites/default/files/publication/resolution-on-data-protection-in-automated-and-connected-vehicles_en_1.pdf

Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research, 32(11), 1231–1237. https://doi.org/10.1177/0278364913491297

Girshick, R. (2015). Fast R-CNN. In 2015 IEEE International Conference on Computer Vision (ICCV) (pp. 1440–1448). https://doi.org/10.1109/ICCV.2015.169

Hao, L.-Y., Yang, J.-R., Zhang, Y., & Zhang, J. (2023). Multi-target vehicle detection based on corner pooling with attention mechanism. Applied Intelligence, 53(23), 29128–29139. https://doi.org/10.1007/s10489-023-05084-4

Hong, F., Lu, C.-H., Liu, C., Liu, R.-R., & Wei, J. (2020). A Traffic Surveillance Multi-Scale Vehicle Detection Object Method Base on Encoder-Decoder. IEEE Access, 8, 47664–47674. https://doi.org/10.1109/ACCESS.2020.2979260

IEEE Spectrum. (2024). Self-Driving Cars Will Be Ready Before Our Laws Are. Retrieved August 14, 2024, from https://spectrum.ieee.org/selfdriving-cars-will-be-ready-before-our-laws-are

Improvement and Application of Deformable Parts Model in Vehicle Detection-【VIP Journal Official Website】- Chinese Journal Service Platform. (2025). Retrieved May 21, 2025, from https://cstj.cqvip.com/Qikan/Article/Detail?id=7000043723

Jobson, D. J. (2004). Retinex processing for automatic image enhancement. Journal of Electronic Imaging, 13(1), 100. https://doi.org/10.1117/1.1636183

Keltron dataset. (2024). KELTRON. Retrieved June 10, 2024, from https://www.keltron.org/index.php/intelligent-traffic-systems

Krontiris, I., et al. (2020). Autonomous Vehicles: Data Protection and Ethical Considerations. In Computer Science in Cars Symposium (pp. 1–10). ACM. https://doi.org/10.1145/3385958.3430481

Li, J., Xu, Z., & Xu, L. (2023). Vehicle and pedestrian detection method based on improved YOLOv4-tiny. Optoelectronics Letters, 19(10), 623–628. https://doi.org/10.1007/s11801-023-3078-x

Lin, T.-Y., et al. (2014). Microsoft COCO: Common Objects in Context. arXiv. https://doi.org/10.48550/ARXIV.1405.0312

Liu, W., et al. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), Computer Vision– ECCV 2016 (pp. 21–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

Luo, J., Fang, H., Shao, F., Zhong, Y., & Hua, X. (2021). Multi-scale traffic vehicle detection based on faster R–CNN with NAS optimization and feature enrichment. Defence Technology, 17(4), 1542–1554. https://doi.org/10.1016/j.dt.2020.10.006

Ma, C., & Xue, F. (2024). A review of vehicle detection methods based on computer vision. Journal of Intelligent and Connected Vehicles, 7(1), 1–18. https://doi.org/10.26599/JICV.2023.9210019

Papers with Code - MS COCO Dataset. (2024). Retrieved July 27, 2024, from https://paperswithcode.com/dataset/coco

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91

Sun, X., Huang, Q., Li, Y., & Huang, Y. (2019). An Improved Vehicle Detection Algorithm Based on YOLOV3. In 2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom) (pp. 1445–1450). https://doi.org/10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00208

Sun, Z., Liu, C., Qu, H., & Xie, G. (2022). A Novel Effective Vehicle Detection Method Based on Swin Transformer in Hazy Scenes. Mathematics, 10(13), 2199. https://doi.org/10.3390/math10132199

Tehrani Niknejad, H., Takeuchi, A., Mita, S., & McAllester, D. (2012). On-Road Multivehicle Tracking Using Deformable Object Model and Particle Filter With Improved Likelihood Estimation. IEEE Transactions on Intelligent Transportation Systems, 13(2), 748–758. https://doi.org/10.1109/TITS.2012.2187894

Tian, Z., Shen, C., Chen, H., & He, T. (2019). FCOS: Fully Convolutional One-Stage Object Detection. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 9626–9635). https://doi.org/10.1109/ICCV.2019.00972

UA-DETRAC Dataset | Papers With Code. (2024). Retrieved July 27, 2024, from https://paperswithcode.com/dataset/ua-detrac

Vision-based scale-adaptive vehicle detection and tracking for intelligent traffic monitoring | IEEE Conference Publication | IEEE Xplore. (2025). Retrieved May 21, 2025, from https://ieeexplore-ieee-org.sndl1.arn.dz/document/7090470

Wang, J., Cai, Z., & Yu, J. (2020). Achieving Personalized k-Anonymity-Based Content Privacy for Autonomous Vehicles in CPS. IEEE Transactions on Industrial Informatics, 16(6), 4242–4251. https://doi.org/10.1109/TII.2019.2950057

Wen, L., et al. (2020). UA-DETRAC: A new benchmark and protocol for multiobject detection and tracking. Computer Vision and Image Understanding, 193, 102907. https://doi.org/10.1016/j.cviu.2020.102907

Wu, B.-F., Kao, C.-C., Jen, C.-L., Li, Y.-F., Chen, Y.-H., & Juang, J.-H. (2014). A Relative-Discriminative-Histogram-of-Oriented-Gradients-Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking. IEEE Transactions on Industrial Electronics, 61(8), 4228–4237. https://doi.org/10.1109/TIE.2013.2284131

Xiong, C., Yu, A., Yuan, S., & Gao, X. (2023). Vehicle detection algorithm based on lightweight YOLOX. Signal, Image and Video Processing, 17(5), 1793–1800. https://doi.org/10.1007/s11760-022-02390-1

Zoph, B., & Le, Q. V. (2016). Neural Architecture Search with Reinforcement Learning. arXiv. https://doi.org/10.48550/ARXIV.1611.01578