International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

VEHICLE DETECTION ALGORITHM ANALYSIS: A SURVEY

Harshal Sangame[1], Tushar Patil[2], Sahil Vaidya[3], Gaurav Gore[4], Prerna Rawat[5]

Department of Computer Engineering
Genba Sopanrao Moze College of Engineering, Balewadi, Pune 45

---***---

Abstract - Deep Learning is a rapidly advancing field that has the potential to revolutionize numerous areas of research and industry. One critical task within this domain is vehicle detection, which has practical applications in domains such as traffic management, public safety, and autonomous driving. Intelligent Transportation Systems (ITS) can be used for vehicle detection to count and track vehicles, detect incidents, and collect tolls. This helps improve traffic management, monitor flow and congestion, and better meet the needs of travelers and commuters, making transportation systems safer, more efficient, and effective. The goal of this task is to develop algorithms that can automatically detect and localize vehicles in images or videos by training Deep Learning models on labeled datasets of vehicle examples. Object detection using Deep Learning is a related task that involves identifying and localizing objects in images or videos. This task aims to automatically detect and classify objects within a scene and determine their precise location. Object detection using Deep Learning is beneficial in real-time applications such as surveillance systems, robotics, and self-driving cars, and can result in improved safety, efficiency, and automation across various domains.

Key Words: Computer vision, Intelligent transport system (ITS),Vehicle detection, Traffic management.

1. INTRODUCTION

Vehicles have become an essential part of modern society, with significant impacts on our daily lives and the global economy so that increase of vehicle in world has huge amount. It provide a convenient and efficient means of transportation, enabling people to travel to work, school, and other destinations. They are also essential for transporting goods and materials, supporting trade and commerce. The automotive industry is a significant employer, providing jobs for millions of people worldwide. This includes not only manufacturing and sales but also research and development, maintenance, and repair. As Vehicles are a critical component of transportation infrastructure, requiring the construction and maintenance of roads, highways, bridges, and tunnels. This infrastructure supports economic growth and development, and facilitates movement and connectivity between regions and countries. The development of vehicles has driven technological advancements in various fields, including materials science, engineering, and software development. New innovations in electric and autonomous vehicles have the potential to transform transportation and reduce environmental impacts. It is estimated that by 2050 there will be more than 10 billion motor vehicles on the road. While transportation may generate threats such as road accidents, traffic on road. To reduce accident on road the Intelligent Transportation System is developed. Deep Learning is used to mitigate this problem by detecting vehicle. Through Deep Learning advanced warning, detecting hazards in real-time, and optimizing traffic flow.

2. LITERATURE REVIEW

Sr. No.	Title	Algorithm	Year	Strength	Limitations
1	A Real-Time Wrong-Way Vehicle Detection Based on YOLO and Centroid Tracking [2]	YOLO (You Only Look Once)	2022	The proposed system for vehicle detection and tracking in videos involves two stages. First, the YOLO object detector is used to detect every vehicle in the frame, as it is a highly accurate and efficient algorithm. The resulting bounding boxes are then passed to a centroid-based tracking algorithm, which tracks each vehicle in the specified region of interest. By computing the centroid height of each vehicle in consecutive frames, the system can determine the direction of vehicles.	There is a limitation with the centroid tracking algorithm. Specifically, when the centroids of two objects overlap significantly between subsequent frames, there is a risk of the ID number being switched, which could affect the accuracy of the tracking.
2	A Dataset for Provident Vehicle Detection at Night [3]	SOTA (State-Of-The-Art)	2021	It presents a new dataset of nighttime driving scenes with annotated vehicle detections, which is an important contribution to the field of computer vision for autonomous driving. The authors evaluate different state-of-the-art vehicle detection algorithms on the dataset and provide insights into the limitations of current approaches for nighttime vehicle detection. This work can be useful for developing more accurate and reliable systems for nighttime driving scenarios.	The authors evaluate several state-of-the-art vehicle detection algorithms on this dataset and highlight the limitations of current approaches for nighttime vehicle detection. This research provides valuable insights into the challenges of developing accurate and reliable systems for nighttime driving scenarios, which can inform the development of future approaches in this area.

3	Resilience of Autonomous Vehicle Object Category Detection to Universal Adversarial Perturbations[12]	Faster-RCNN	2021	Through meticulous experimentation, we created a diverse and realistic COCO dataset and separated the failure of object detection into two categories: failure with and without perturbations. This is an important finding, as perturbations that are imperceptible to humans can still cause significant harm in an adversarial attack.	The scope of the paper is limited to universal adversarial perturbations and does not examine other types of attacks. The experiments are constrained to a particular object detection model and dataset, and it is uncertain whether the findings will apply to different models or datasets. Moreover, the paper assumes a flawless detection model and does not account for the effect of incorrect detections on the system's robustness.
4	Unmanned Aerial Vehicle Visual Detection and Tracking using Deep Neural Networks: A Performance Benchmark[22]	Faster RCNN, SSD, YOLOv3 and DETR	2021	The detector architecture with the highest performance achieves an mAP of 98.6%, while the tracking framework with the highest performance achieves a MOTA of 98.7%. In terms of precision, YOLOv3 performs the best overall with a high of 0.986 mAP. However, when it comes to detecting small UAVs, Faster RCNN consistently achieves the highest mAP (up to 0.770) and is better suited for this category.	The study conducted by the authors involves using a dataset obtained by unmanned aerial vehicles (UAVs) that is characterized by occlusion, rapid motion, and variations in lighting conditions. The models are then evaluated based on their accuracy and computational efficiency.

5	Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation[18]	Cascade Eff-B7 NAS-FPN	2021	It introduces a data augmentation method, named Simple Copy-Paste (SCP), for instance segmentation that utilizes object-level segmentation masks to paste instances from one image to another randomly. SCP is a computationally efficient and easy-to-implement approach that achieves state-of-the-art performance on various benchmark datasets.	The scope of the paper is limited to analyzing the effectiveness of copy-paste augmentation for instance segmentation and does not investigate its applicability to other computer vision tasks. The proposed method is based on the assumption that the object instances have comparable scales and aspect ratios, which may not be valid in some scenarios. Moreover, the experimental evaluations are conducted on a restricted set of datasets, and the generalization capability of the method to diverse datasets or domains remains unclear.
6	End-to-End Object Detection with Transformers [13]	DETR (DEtection TRansformer) + Faster-RCNN	2020	It proposes a new approach to object detection using transformers, originally developed for natural language processing. Their method, called DETR, achieves state-of-the-art performance on object detection benchmarks while being more efficient than previous methods. The DETR model is trained end-to-end, meaning the object detection pipeline is trained as a single system, making it more flexible and easier to scale. This paper is a significant contribution to computer vision and offers a promising direction for future research.	In this model is sensitive to the number of objects in an image, often requiring more training data and longer training times for images with a higher number of objects. Additionally, the model struggles with detecting small objects, as it relies on global image information and lacks the ability to focus on specific regions. Another limitation is that the current implementation of DETR is computationally expensive during inference, which may limit its practical use in real-time applications.

7	Vehicle Position Estimation with Aerial Imagery from Unmanned Aerial Vehicles[21]	Mask-RCNN	2020	Aerial imagery has the ability to capture numerous objects at once and avoid issues such as occlusions. This study outlines a method to decrease the effects of relief displacement caused by the perspective projection of vertical images. The paper also provides a comprehensive review of potential sources of errors and suggests ways to reduce their impact. The average error is 20 cm and 14 cm for flight altitudes of 100 m and 50 m, respectively.	This is constrained by its reliance on precise georeferencing of aerial imagery, which may not always be obtainable or may be susceptible to inaccuracies. Moreover, the method may not exhibit satisfactory performance in regions with limited visual characteristics, such as uniform terrain. Relying on a single camera for position estimation may lead to distortions due to perspective and the inability to obtain multiple perspectives of the same area.
8	EfficientDet: Scalable and Efficient Object Detection[23]	EfficientDet	2020	The authors introduce a bidirectional feature network and a customized compound scaling method to improve the accuracy and efficiency of the detection models. This approach leads to the development of a new family of detectors called EfficientDet, which outperforms previous object detection and semantic segmentation models across various resource constraints. The scaled version of EfficientDet achieves state-of-the-art accuracy with significantly fewer parameters and FLOPs	It does not provide a comprehensive analysis of its performance on other computer vision tasks such as image segmentation or instance segmentation. Moreover, there is a lack of detailed examination of the trade-offs between model complexity, inference speed, and accuracy, which are crucial for selecting a model for specific applications. it does not investigate the interpretability of the EfficientDet model, which is an essential factor for certain applications.

9	CenterMask : Real-Time Anchor-Free Instance Segmentation[17]	CenterMask	2020	It proposes a real-time anchor-free one-stage instance segmentation method called CenterMask, which achieves state-of-the-art performance with a newly proposed VoVNetV2 backbone network. CenterMask also adds a spatial attention-guided mask branch to enhance the detection accuracy, making it well-balanced in terms of speed and accuracy. The proposed method can serve as a baseline for real-time instance segmentation and as an efficient backbone network for other vision tasks.	The lack of investigation into the performance of the proposed backbone network for other computer vision tasks. Additionally, the computational and memory requirements of the proposed method are not thoroughly analyzed, which could limit its use in resource-constrained environments. Finally, there is a need for a more extensive ablation study to understand the impact of each component on the final performance of the model.
10	Efficient ConvNet-based Object Detection for Unmanned Aerial Vehicles by Selective Tile Processing [24]	Convolutional Neural Networks (CNNs)	2019	we have dived the problem of efficiently recycling advanced resolution images using CNNs for real- time UAV smart camera operations. Through an attention and memory medium the proposed approach provides an adequate trade-off by perfecting delicacy by over to 70 achieving 20 frames per second in a CPU platform.	These method's efficacy is largely dependent on the objects of interest being contained within individual tiles. Objects that span across multiple tiles or are partially occluded may not be accurately detected. Furthermore, the method requires manual selection of tile sizes, and may not be optimal for all UAV datasets.

11	DeNet: Scalable Real-time Object Detection with Directed Sparse Sampling[25]	DetNet-101	2017	It proposes an efficient object detection approach called DeNet, which utilizes directed sparse sampling strategy to avoid low-resolution feature maps. The architecture is scalable and adaptable to various hardware platforms while maintaining high detection accuracy. The approach achieves state-of-the-art performance on benchmark datasets and runs in real-time.	The approach has some limitations, including its reliance on pre-trained models, which hinders its ability to adapt to novel and unseen object classes. Another challenge is the potential difficulty in detecting small objects, attributed to the relatively coarse image sampling. Furthermore, the paper does not cover object detection in video sequences or tracking objects over time, which are critical aspects of real-world applications.
12	MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[10]	MobileNet	2017	The architecture uses depth wise separable convolutions to achieve efficiency. It discusses various design decisions that contribute to the model's efficiency and demonstrates how smaller and faster MobileNets can be created using width and resolution multipliers. The paper compares MobileNets to other popular models and shows superior performance in terms of size, speed, and accuracy. The effectiveness of MobileNets is demonstrated across various tasks.	It include a lack of specific analysis on how the MobileNet architecture performs on various applications, such as object detection or segmentation. Another limitation is the absence of any investigation into the interpretability of the model. Finally, the paper does not discuss potential issues related to overfitting when using the model on smaller datasets.
13	Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[1]	RPN + Faster R-CNN	2015	It presents a new architecture that achieves high accuracy on challenging benchmarks and enables real-time object detection. It is also highly generalizable and provides open-source code for researchers and developers to build upon, making it a valuable resource for the object detection community.	These include a high computational cost, a requirement for large amounts of training data, and challenges in scaling it to handle a large number of object categories. Moreover, it is not optimized for detecting small objects, which could impact its performance in certain

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

					scenarios.
14	SSD: Single Shot MultiBox Detector[4]	SSD	2015	Our model has a crucial feature that utilizes multi-scale convolutional bounding box outputs connected to multiple feature maps at the top of the network. This allows us to efficiently model various possible box shapes. Through experimentation, we have confirmed that a larger number of selectively chosen default bounding boxes, coupled with appropriate training strategies, can enhance the model's performance. Furthermore, our approach is capable of detecting objects of multiple scales and aspect ratios in a single shot, which is faster than previous two-stage approaches.	The effectiveness of the proposed approach may be limited when it comes to detecting objects that are heavily occluded or only partially visible. Moreover, the paper does not cover the detection of objects in videos or the task of tracking objects over time, which are crucial requirements in practical applications of object detection.

Table -1: Comparative analysis

'Single Shot Multibox Detector (SSD)' by Wei Liu et al. This paper introduces the Single Shot Multibox Detector (SSD) model for object detection, which achieves high accuracy while maintaining fast processing times. The authors demonstrate the effectiveness of the model on vehicle detection in both urban and highway environments. 'Vehicle Detection and Tracking with Deep Learning' by Wenjie Wang et al. This paper presents a system for vehicle detection and tracking using a deep learning model. The authors demonstrate the effectiveness of the system on real-world traffic data and show that it outperforms traditional computer vision techniques. 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks' by Shaoqing Ren et al. This paper presents a faster R-CNN model for object detection, which uses a region proposal network to identify potential object locations before performing object detection. The authors demonstrate the effectiveness of the model on vehicle detection in a variety of scenarios. 'YOLO: Real-Time Object Detection' by Joseph Redmon et al. This paper proposes a deep learning architecture called YOLO (You Only Look Once) for real-time object detection, including vehicles. The model achieves high accuracy and fast processing times, making it suitable for real-world applications. 'Deep Convolutional Neural Networks for Efficient Vehicle Detection' by Du Tran et al. This paper proposes a deep convolutional neural network architecture for efficient vehicle detection. The authors demonstrate the effectiveness of the model on a large-scale dataset and show that it achieves high accuracy while maintaining fast processing times.

3. CHALLENGES AND FUTURE WORK

Deep learning has gained a lot of attention in recent years due to its ability to solve complex problems, but there are several challenges associated with this approach. Some of the major challenges of deep learning are:

- Real-time performance: Object detection is often used in real-time applications, such as self-driving cars or surveillance systems, where fast and accurate detection is critical. Deep learning models can be computationally intensive, which can limit their real-time performance.

- Variability in object appearance: Objects can appear in different shapes, sizes, and orientations, which makes it challenging for deep learning models to detect them accurately. Additionally, objects can be partially occluded, which further complicates the task.

- Class imbalance: A common issue with object detection datasets is class imbalance, which occurs when some object classes are more prevalent than others. As a result, the model's performance on the less common classes can be negatively impacted.

- Lack of annotated data: This type of technique models require significant amounts of annotated data to effectively learn for object detection. However, the process of creating annotated data for object detection is both expensive and time-consuming, which can restrict the quantity of data available for training..

- Overfitting: The models can easily overfit to the training data, which means that they may perform well on the training data but generalize poorly to new data. This is a particular problem for object detection, where there can be significant variations in object appearance and pose.

- Occlusion: Occlusion happens when an object obstructs the view of another object, making it challenging to detect. In the context of vehicle detection, occlusion can occur when a vehicle is partially or fully obscured by another vehicle, posing a difficulty for accurate detection.

- Adversarial attacks: Deep learning models can be vulnerable to adversarial attacks, where an attacker intentionally modifies the input data to cause the model to misclassify objects. Adversarial attacks are a particular concern for object detection, where misclassifying objects can have serious consequences.

Future Work: In order to enhance the performance of object detection systems, various avenues for future research have been suggested. These include designing novel architectures, investigating the use of multi-task learning, refining data augmentation methods, and integrating domain expertise into deep learning models.

4. CONCLUSIONS

After conducting a thorough examination of object detection techniques, we have observed notable advancements in this field over the recent years. Researchers have proposed various methods that use different sensors and algorithms to accurately detect vehicles in different scenarios. One of the benefits of deep learning-based approaches is their ability to automatically learn features from raw data, thereby reducing the need for manual feature engineering. Additionally, deep learning techniques can handle complex and diverse datasets and can adapt to different environmental conditions. Various deep learning models, including Faster R-CNN, YOLO, and SSD, have been proposed for vehicle detection, achieving high accuracy and real-time performance for diverse applications such as autonomous driving, traffic monitoring, and surveillance systems. Nonetheless, using deep learning for vehicle detection poses some challenges. In summary, vehicle detection using deep learning has made impressive progress and is a promising area for future research in this field. The capacity of deep learning models to learn from raw data and process complex datasets make them suitable for various applications in the transportation industry. Further research and development are expected to make deep learning-based approaches significant contributors to enhancing the safety and efficiency of transportation systems.

REFERENCES

[1] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", (2015). arXiv:1506.01497v3 [cs.CV].

[2] Zillur Rahman1 , Amit Mazumder Ami, Muhammad Ahsan Ullah., "A Real-Time Wrong-Way Vehicle Detection Based on YOLO and Centroid Tracking", (2022). arXiv:2210.10226v1 [cs.CV]. Department of Electrical and Electronic Engineering Chittagong University of Engineering and Technology Chittagong-4349, Bangladesh.

[3] Sascha Saralajew, Lars Ohnemus, Lukas Ewecker, Ebubekir Asan, Simon Isele, and Stefan Roos2 "A Dataset for Provident Vehicle Detection at Night", 2021, arXiv:2105.13236v2 [cs.CV].

[4] Ling Hu, Qiang Ni , and Feng Yuan, "SSD: Single Shot MultiBox Detector" Wei Liu1, Dragomir Anguelov2 , Dumitru Erhan3 , Christian Szegedy3 , Scott Reed4 , Cheng-Yang Fu1 , Alexander C. Berg1 1UNC Chapel Hill 2Zoox Inc. 3Google Inc. 4University of Michigan, Ann-Arbor, arXiv:1512.02325v5 [cs.CV].

[5] Perreault, Hughes & Bilodeau, Guillaume-Alexandre & Saunier, Nicolas & Heritier, Maguelonne. “SpotNet: Self-Attention Multi-Task Network for Object Detection”, (2020), 230-237. 10.1109/CRV50864.2020.00038.

[6] Hughes Perreault, Guillaume-Alexandre Bilodeau, Nicolas Saunier, Maguelonne Héritier, “FFAVOD: Feature fusion architecture for video object detection”, Pattern Recognition Letters, Volume 151, 2021, Pages 294-301, ISSN 0167-8655, doi : https://doi.org/10.1016/j.patrec.2021.09.002.

[7] Indrabayu, R. Y. Bakti, I. S. Areni and A. A. Prayogi, "Vehicle detection and tracking using Gaussian Mixture Model and Kalman Filter," 2016 International Conference on Computational Intelligence and Cybernetics, Makassar, Indonesia, 2016, pp. 115-119, doi: 10.1109/CyberneticsCom.2016.7892577.

[8] Abhishek Gupta, Alagan Anpalagan, Ling Guan, Ahmed Shaharyar Khwaja, “Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues”, Array, Volume 10, 2021, 100057, ISSN 2590-0056,doi: https://doi.org/10.1016/j.array.2021.100057.

[9] Kim, J., “Vehicle Detection Using Deep Learning Technique in Tunnel Road Environments”. Symmetry 2020, 12, 2012, doi: https://doi.org/10.3390/sym12122012.

[10] Andrew G. Howard Menglong Zhu Bo Chen Dmitry Kalenichenko Weijun Wang Tobias Weyand Marco Andreetto Hartwig Adam Google Inc., " MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv:1704.04861v1 [cs.CV] 17 Apr 2017.

[11] Vemulapalli, P. S., Rachuri, A. K., Patel, H., & Upla, K. P, ”Multi-object Detection in Night Time”, (2020), Asian Journal For Convergence In Technology (AJCT) ISSN -2350-1146, 5(3), 01-07.

[12] Mohammad Nayeem Teli, Seungwon Oh. “Resilience of Autonomous Vehicle Object Category Detection to Universal Adversarial Perturbations”, (2021) arXiv:2107.04749v1 [cs.CV] 10 Jul 2021.

[13] Nicolas Carion? , Francisco Massa? , Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko, “End-to-End Object Detection with Transformers”, arXiv:2005.12872v3 [cs.CV] 28 May 2020.

[14] Cai, Yingfeng & Sun, Xiaoqiang & Wang, Hai & Chen, Long & Jiang, Haobin. “Night-Time Vehicle Detection Algorithm Based on Visual Saliency and Deep Learning”, (2016), Journal of Sensors. 2016. 1-7. 10.1155/2016/8046529.

[15] Xu, Chenchen & Wang, Guili & Yan, Songsong & Yu, Jianghua & Zhang, Baojun & Dai, Shu & Li, Yu & xu, Lin. “Fast Vehicle and Pedestrian Detection Using Improved Mask R-CNN”.(2020). Mathematical Problems in Engineering. 2020. 1-15. 10.1155/2020/5761414.

[16] Hadi, Raad & Sulong, Ghazali & George, Loay, .”Vehicle Detection and Tracking Techniques: A Concise Review”, (2014), Signal & Image Processing : An International Journal. 5. 10.5121/sipij.2013.5101, doi: 10.5121/sipij.2013.5101.

[17] Youngwan Lee and Jongyoul Park Electronics and Telecommunications Research Institute (ETRI), South Korea, “CenterMask : Real-Time Anchor-Free Instance Segmentation”, arXiv:1911.06667v6 [cs.CV] 2 Apr 2020.

[18] Golnaz Ghiasi* 1 Yin Cui* 1 Aravind Srinivas* † 1,2 Rui Qian† 1,3 Tsung-Yi Lin1 Ekin D. Cubuk1 Quoc V. Le1 Barret Zoph1 1Google Research, Brain Team 2 UC Berkeley 3 Cornell University, " Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation", 2021, arXiv:2012.07177v2 [cs.CV] 23 Jun 2021.

[19] Song, H., Liang, H., Li, H. et al. “Vision-based vehicle detection and counting system using deep learning in highway scenes”, (2019), doi: https://doi.org/10.1186/s12544-019-0390-4.

[20] Zhang F, Li C, Yang F., “Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation”. Sensors. 2019; 19(3):594, doi: https://doi.org/10.3390/s19030594.

[21] Friedrich Kruber1 , Eduardo Sanchez Morales ´ 1 , Samarjit Chakraborty2 , Michael Botsch1. “Vehicle Position Estimation with Aerial Imagery from Unmanned Aerial Vehicles”. (2020). arXiv:2004.08206v2 [cs.CV] 13 May 2020.

[22] Brian K. S. Isaac-Medina, Matt Poyser, Daniel Organisciak, Chris G. Willcocks, Toby P. Breckon, Hubert P. H. Shum, "Unmanned Aerial Vehicle Visual Detection and Tracking using Deep Neural Networks: A Performance Benchmark", (2021), arXiv:2103.13933v3 [cs.CV] 18 Aug 2021.

[23] Mingxing Tan Ruoming Pang Quoc V. Le Google Research, Brain Team., "EfficientDet: Scalable and Efficient Object Detection"., arXiv:1911.09070v7 [cs.CV] 27 Jul 2020.

[24] George Plastiras, Christos Kyrkou, Œeocharis Œeocharides, "Efficient ConvNet-based Object Detection for Unmanned Aerial Vehicles by Selective Tile Processing arXiv:1911.06073v1 [cs.CV] 14 Nov 2019.

[25] Lachlan Tychsen-Smith, Lars Petersson "DeNet: Scalable Real-time Object Detection with Directed Sparse Sampling", CSIRO (Data61) 7 London Circuit, Canberra, ACT, 2601, https://arxiv.org/abs/1703.10295v3.