SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13204
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Computer Vision-Based Intelligent Traffic Surveillance: Multi-Vehicle Tracking and Detection

Amir Mahmud Husein[1)*], Kevi Noflianhar Lubis[2)], Daniel Salim Sidabutar[3)], Yansan Yuanda[4)], Kevry[5)], Ashwini Waren[6)]
[1,2,3,4,5,6)]Universitas Prima Indonesia, Indonesia
[1)]amirmahmud@unprimdn.ac.id, [2)]kevinoflianharl04@gmail.com, [3)]danielsalim1996@gmail.com, [4)]yansan.yuanda@yahoo.com, [5)]tankevry11@gmail.com, [6)]asxxxwaren99@gmail.com

Submitted : Nov 28, 2023 | **Accepted** : Dec 19, 2023 | **Published** : Jan 1, 2024

Abstract: The application of vehicle detection in real-time traffic surveillance systems is one of the challenging research fields with different objectives. One of the problems is the detection of many vehicles simultaneously in a video sequence sourced from CCTV cameras. In many works, the focus is only on detecting vehicle classes such as motorcycles, buses, trucks, and cars or special vehicles such as ambulances and others. In this research, we propose to apply 13 classes of vehicle types and implement YOLOv4 in the traffic surveillance task. More specifically, all classes are labeled, and then the YOLOv4 model is trained on 800 images and tested on 23 videos from three intersections in Medan City, namely Juanda Katamso Intersection, Gatot Subroto Intersection, and Uniland Intersection. Based on the test results, YOLOv4 proves successful in detecting many vehicles in frame-by-frame sequence with various types of vehicles. All vehicle detection data will be stored in the file.

Keywords: Computer Vision; Deep Learning; Traffic Surveillance; Vehicle Detection; YOLOv4

INTRODUCTION

Advances in deep learning-based computer vision in the past decade have become the leading method for high-quality general object detection. Today's popular object detection methods consist of one-stage and two-stage approaches. R-CNN, Fast R-CNN, Faster R-CNN, and Spatial Pyramid Pooling Network (SSP-Net) are two-stage approach methods by optimizing Selective Search function to Regional Proposal Network (RPN) to generate candidate object boxes throughvarious algorithms, then classifying the objects with Convolutional Neural Network which results in high accuracy (Husein, Christopher, Gracia, Brandlee, & Hasibuan, 2020). However, the inference process can be too intensive mainly due to the generation of object bounding boxes which is computationally time-consuming and takes up unnecessary spaces (Yu, Shin, Kim, Roh, & Sohn, 2021).

In recent years, many object detection applications have been proposed as real-time traffic surveillance systems with different purposes, such as vehicle detection (Du, Zhang, Zhang, & Xu, 2021; Liang & Ji, 2022; Luo, Fang, Shao, Zhong, & Hua, 2021; Meng et al., 2020), vehicle counting (Lu et al., 2021; Nam Bui, Yi, & Cho, 2020; Yang et al., 2021), vehicle tracking (C. Liu et al., 2021; Sun, Zhang, Yang, & Fan, 2023), vehicle classification (Butt et al., 2021; Neupane, Horanont, & Aryal, 2022), and others. Various techniques have been proposed for intelligent traffic surveillance systems, but deep learning techniques are more accurate than traditional techniques (Li, Liu, Zhao, Zhang, & He, 2018) because they have the ability to learn image or video features for various classification, detection, or segmentation tasks (Ammar, Koubaa, Ahmed, Saad, & Benjdira, 2021). YOLO network architecture

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13204
e-ISSN : 2541-2019
p-ISSN : 2541-044X

is one of the models that attract the most interest of researchers to build traffic surveillance systems (Huang, Zheng, Sun, Yang, & Liu, 2020; Mahto, Garg, Seth, & Panda, 2020; Wang, Wang, Cao, & Wang, 2020). Several improvements on YOLO network have been proposed, such as YOLO-vocRV (Li et al., 2018), Gaussian-YOLOv3 (Choi, Chun, Kim, & Lee, 2019), MME-YOLO (Zhu et al., 2021), and SF- YOLO (Han, Lee, Lim, & Choi, 2020). Furthermore, Alexey Bochkovskiy recently released YOLOv4 (Bochkovskiy, Wang, & Liao, 2020) which is an improvement over YOLOv3 to detect objects in multiple detection areas quickly and accurately. YOLOv4 uses Cross-Stage-Partial-Connections (CSP) as a new backbone that can improve the learning capability of CNN and is faster than other single-stage approaches.

Various studies in this field have shown promising results, for example (Tian & Kim, 2023) proposed a method to improve the performance of Faster R-CNN in vehicle detection by classifying the weather condition in traffic images in advance. By using dark channel prior (DCP) and block-matching and 3D filtering (BM3D) algorithms, the image quality was improved and noise was reduced according to the weather condition in the image. The designed Faster R- CNN also utilized VGG16 and ResNet101 for better feature extraction and accuracy. Meanwhile, (Luo et al., 2021) performed vehicle detection by utilizing Neural Architecture Search (NAS) optimization for feature extraction in Faster R-CNN. Feature enrichment was also implemented so that vehicles that look small or are obscured by other objects in the image can be detected properly. In addition, the quality of the images used had been improved with the Retinex-based image adaptive correction (RIAC) algorithm.

Furthermore, (Kasper-Eulaers et al., 2021) adopted the YOLOv5 network architecture to detect trucks in real-time under dark winter condition. The truck detection also involved the classification of the main parts of the truck to deal with the problem of obscured truck object. Zhao *et al.* (J. Zhao et al., 2022) proposed the YOLOv4_AF model which is an extension of YOLOv4 for better vehicle classification and detection. YOLOv4_AF utilizes a convolutional block attention module (CBAM) so that the system can focus more on the most important features or parts of the image. In addition, the Feature Pyramid Network (FPN) in YOLOv4 was also modified to further improve detection capabilities. Testing on the BIT-Vehicle and UA-DETRAC datasets showed the superior performance of YOLOv4_AF over YOLOv4, Faster R- CNN, and EfficientDet in vehicle detection.

For faster vehicle detection, (Y. Liu & Zhang, 2021) proposed a redesign of the feature extraction network in SSD by replacing VGG16 with MobileNet to reduce the parameters used. Meanwhile, (M. Zhao, Zhong, Sun, & Chen, 2021) developed a strategy called feature pyramid enhancement strategy (FPES) to improve the performance of SSD in vehicle detection. SSD performance was balanced to have high detection accuracy and speed.

Computer vision with deep learning techniques plays an important role in surveillance systems traffic, various new models and techniques have been proposed by researchers with applications in different problems like vehicle counting, vehicle detection, congestion detection as well as vehicle tracking, all of which are very important to implement in Intelligent transportation System (ITS). However, some of the proposed models are difficult to determine best to adopt because the systems reported in many different studies on multiple stages, as well as adopting different data sets and testing conditions, then The problem formulation in this research is what is the computer vision framework with YOLOv4 can detect and track multiple vehicles on highway traffic videos?

In urban traffic surveillance systems, simultaneous multi-vehicle detection and different vehicle movement tracking are important but challenging task. Therefore, the main objective of this research is to implement and analyze the application of computer vision with YOLOv4 as the deep learning models in the construction of a traffic surveillance system based on CCTV videos of road traffic.

METHOD

The YOLOv4 framework is one of the most accurate methods in the field of computer vision with deep learning and will be applied to this research. All training experiments are conducted on Google Colaboratory, while the direct object detection testing on videos is conducted on a laptop with the specification of Intel Core i5-7300HQ quad-core 2.5GHz TurboBoost up to 3.5GHz, 8GB RAM, Nvidia GeForce GTX 1050 4 GB GPU, and Windows 10 operating system.

*name of corresponding author

DOI : https://doi.org/10.33395/sinkron.v9i1.13204
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Preparation

The dataset used in this research is sourced from Medan Department of Transportation. The dataset is obtained from CCTV recordings of traffic flow monitoring at Juanda Katamso Intersection, Gatot Subroto Intersection, and Uniland Intersection. The CCTV recordings are from April 21, 2020 to April 27, 2020 with a capacity of 247 GB and a total of 1,367 videos. Several cuts are made on each video to ignore red light conditions and to only take data with busy vehicle conditions. After that, the video data are divided daily for each intersection so that there are 21 videos which consist of 7 videos for each intersection. All videos have dimensions of 1280x720 with 25 frames per second (fps). The original dataset details are shown in Table 1.

Table 1
Original Dataset

Date	Number of Videos		
	Juanda Katamso	Gatot Subroto	Uniland
21/04/2020	45	38	94
22/04/2020	58	10	85
23/04/2020	58	45	86
24/04/2020	79	81	83
25/04/2020	79	79	79
26/04/2020	79	85	85
27/04/2020	43	38	38
Total	441	376	550

Class Labeling

For labeling process, we use Roboflow (https://roboflow.com/) which allows a maximum limit of 1000 images for each free account. Before labeling, we perform image extraction from videos with a maximum image frame of 250 for each intersection video. In addition, only two videos from each intersection are extracted. Because some of extracted images are quite similar, only 335 images of Juanda Katamso Intersection, 237 images of Gatot Subroto Intersection, and 302 images of Uniland Intersection are taken to be labeled. We use 13 vehicle classes to label vehicle objects in the extracted images. All 13 vehicle classes can be seen in Table 2. The labeled images are then divided into 70% as train data, 20% as validation data, and 10% as test data. The results of labeling from Roboflow can be seen in Fig. 1.

Table 2
Vehicle Type Classes

Class ID	Class Name
0	Big SUV
1	Bus
2	City Car
3	Low MPV
4	Medium Box
5	Medium MPV
6	Medium SUV
7	Mini Box
8	Minibus
9	Pick Up
10	Premium MPV
11	Sedan
12	Truk

*name of corresponding author

Volume 8, Number 1, January 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v9i1.13204 p-ISSN : 2541-044X

Fig. 1 Original and Labeled Images

Training Process

After labeling the images from the dataset, the training process of YOLOv4 will be carried out using Google Colab. The original file of YOLOv4 along with the dataset is imported to Google Colab. The yolov4-custom.cfg file and the yolov4.conv.137 file are also brought into Google Colab. Several configuration settings in the yolov4-custom.cfg file are adjusted, namely:

- batch size = 64 (number of images per iteration)
- subdivision = 24 (number of parts broken down into batches)
- max_batches = 32000
- steps = 25600.0,28800.0 (80% and 90% of max_batches value)
- filters = 64
- class = 12 (the index of the thirteenth class is 12 because the index starts from 0)

After the settings are made, the training is carried out within 1000 up to 8000 iterations. Additionally, the GPU runtime of Google Colab is activated in order to shorten the time needed for training process to be completed.

RESULT

The training results are mainly evaluated based on mAP (mean Average Precision) with iou_thresh of 0.5. The results are shown in Table 3.

Table 3. Training Results

Iteration	mAP@0.5	Precision	Recall	F1-score	Average IoU
1000	70.11	0.68	0.76	0.72	53.36
2000	71.22	0.68	0.79	0.74	53.05
3000	73.28	0.70	0.79	0.74	53.43
4000	71.36	0.69	0.78	0.74	53.65
5000	71.35	0.68	0.78	0.73	52.49
6000	70.58	0.69	0.77	0.73	53.97
7000	77.88	0.68	0.77	0.72	53.17
8000	68.89	0.69	0.76	0.72	54.34

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13204
e-ISSN : 2541-2019
p-ISSN : 2541-044X

As shown in Table 3, the highest mAP value is at the 7000th iteration and hence the YOLOv4 model at the 7000th iteration will be applied to directly detect vehicles in videos. For the testing process, we divide all videos into one video for each day. All testing videos are a combination of morning, afternoon, and evening times with the purpose to optimize the testing time. The total number of videos used for testing is 23 videos which comprise 21 videos from each intersection each day, 1 new video from recording vehicle activity on Manhattan road, and 1 new video of rainy night at Juanda Katamso. The YOLOv4 model detection results will be stored in the tracking.csv file and detec.csv file for each intersection. The tracking.csv file will store the tracking data of all vehicles per frame and the detec.csv file will store information on the number of vehicle classes detected in each frame. Vehicle type detection results at each intersection can be summarized as shown in Table 4, Table 5, and Table 6.

Table 4. Detection Results at Juanda Katamso Intersection

Date	Vehicle Class												
	0	1	2	3	4	5	6	7	8	9	10	11	12
21/04/2020	1,647	-	4,354	13,659	678	3,894	4,257	233	832	1,023	35	102	5
22/04/2020	2,064	-	3,429	17,057	378	7,480	3,055	587	974	2,515	212	216	221
23/04/2020	4,099	76	5,932	20,033	2,266	10,548	6,230	2,167	2,294	3,218	104	258	446
24/04/2020	1,572	37	3,626	16,354	446	3,100	3,485	1,086	701	2,198	34	166	327
25/04/2020	3,888	-	7,456	27,863	804	7,277	5,994	2,377	2,336	2,853	176	435	482
26/04/2020	3,041	-	4,219	14,949	144	5,451	2,648	255	585	1,003	92	114	-
27/04/2020	1,031	-	1,821	10,378	926	1,766	1,166	229	1,031	986	18	185	401
rainy night	4,422	-	2,081	8,784	1,579	8,959	2,386	1,969	1,013	2,701	33	11	911

Table 5. Detection Results at Gatot Subroto Intersection

Date	Vehicle Class												
	0	1	2	3	4	5	6	7	8	9	10	11	12
21/04/2020	2,611	-	279	2,892	1,140	361	2,678	949	214	1,670	196	-	88
22/04/2020	25	-	13	354	56	6	347	32	2	5	2	-	1
23/04/2020	2,244	-	344	5,029	378	786	1,988	1,160	193	3,861	37	3	764
24/04/2020	2,827	-	508	3,464	320	193	1,997	864	190	1,644	74	1	14
25/04/2020	4,544	-	1,183	6,370	734	719	4,138	594	530	1,519	255	-	1,024
26/04/2020	1,250	-	720	1,800	288	351	2,246	205	88	444	82	-	6
27/04/2020	4,977	-	1,203	10,773	2,886	2,153	12,512	2,427	424	1,980	196	99	318

Table 6. Detection Results at Uniland Intersection

Date	Vehicle Class												
	0	1	2	3	4	5	6	7	8	9	10	11	12
21/04/2020	1,647	115	1,269	4,708	858	3,088	907	1,067	688	857	191	44	3
22/04/2020	597	-	2,518	4,392	478	1,902	1,357	667	674	262	41	137	23
23/04/2020	2,261	1	1,218	3,625	313	2,966	1,298	928	654	1,492	47	182	14
24/04/2020	1,405	-	2,084	6,616	521	3,220	2,312	978	697	1,034	356	151	99
25/04/2020	3,482	-	3,641	9,965	496	3,696	3,198	896	412	347	648	53	2
26/04/2020	2,583	-	572	5,277	205	2,417	2,477	286	173	181	190	5	-
27/04/2020	946	-	1,033	5,267	164	1,025	692	241	741	83	619	16	41

DISCUSSIONS

Based on the results of testing vehicle type detection on 23 videos, the proposed YOLOv4 model can accurately detect vehicle types even though there are several processes where this model produces wrong class name and inaccurate object bounding box. Fig. 2 and Fig. 3 show some examples of false detection made by the YOLOv4 model.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v9i1.13204 p-ISSN : 2541-044X

Fig. 2 False Detection Results

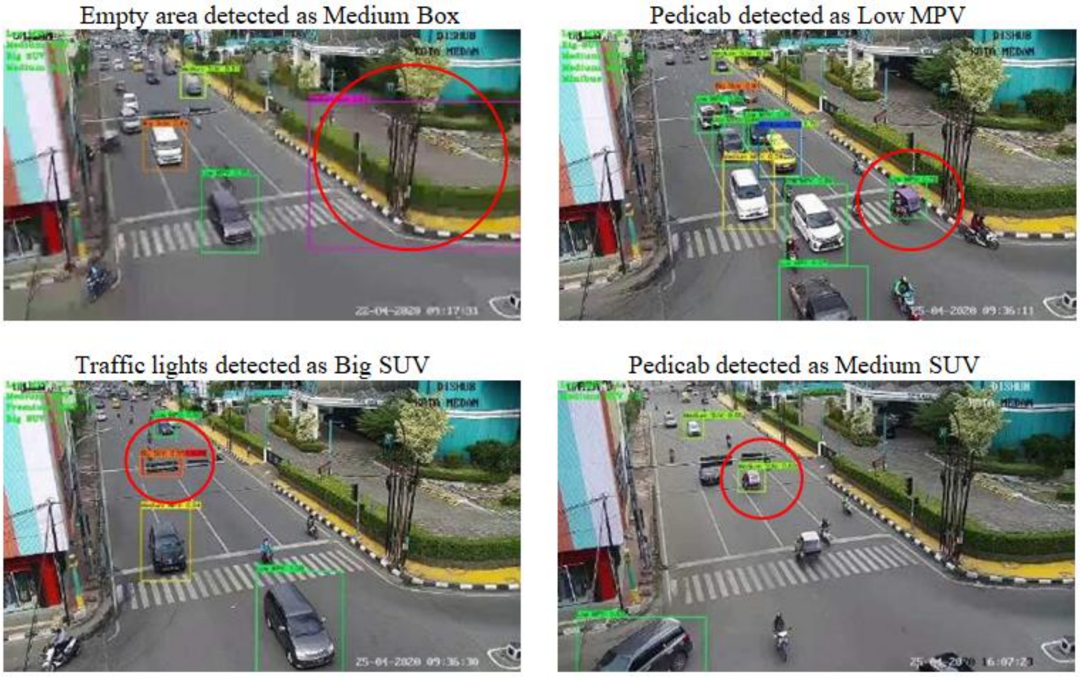

Fig. 3 False Detection Results

In Fig. 2 and Fig. 3, it can be seen that there are detection errors with different types of errors that most likely occur because of inconsistent data labeling process, especially in making bounding boxes to mark the vehicle objects. Besides that, the distance between the camera and the object being detected has an influence on determining vehicle class even though the detection accuracy value is quite high. These detection errors will be evaluated in future research.

The last experiment in this study is comparing the utilization of GPU, CPU 1 (Core i3 Gen 8 HDD), and CPU 2 (Core i3 Gen 7 SSD) in vehicle detection using YOLOv4. After testing, we find out that the inference time on GPU is 9 minutes 23 seconds, while the inference time on CPU 1 is 2 hours 4 minutes 30 seconds and the inference time on CPU 2 is 2 hours 31 minutes 1 second. From these results, we know that computer vision with deep learning technique should use GPU specification.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 1, January 2024
DOI : https://doi.org/10.33395/sinkron.v9i1.13204
e-ISSN : 2541-2019
p-ISSN : 2541-044X

CONCLUSION

Based on the results of testing on 23 CCTV videos of vehicle activities on the highway in Medan City, the proposed YOLOv4 model can directly and accurately detect the type of vehicle from 13 classes used. The YOLOv4 model produced mAP value of 77.88% on the training dataset at the 7000th iteration. The YOLOv4 model can directly detect multiple vehicle types in image files with ±98% accuracy. The developed application can record all detection result data in a video sequence for each frame and then save it to a file that is useful for predicting vehicle density level. There were still some detection errors even though the detection accuracy was quite high.

After experimenting in this study, we have some suggestions for improving the results in future research. The YOLOv4 model can be combined with tracking algorithms such as SORT, Deep SORT, or IOU to optimize detection time. Developing the application to be able to calculate the type of vehicle that passes so that it will be more optimal in predicting the level of vehicle density. Continuing this research for accident detection, vehicle license plate detection, and vehicle locator in video.

REFERENCES

Ammar, A., Koubaa, A., Ahmed, M., Saad, A., & Benjdira, B. (2021). Vehicle detection from aerial images using deep learning: A comparative study. *Electronics (Switzerland)*, *10*(7), 1–31. https://doi.org/10.3390/electronics10070820

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv*. Retrieved from http://arxiv.org/abs/2004.10934

Butt, M. A., Khattak, A. M., Shafique, S., Hayat, B., Abid, S., Kim, K. Il, … Adnan, A. (2021). Convolutional Neural Network Based Vehicle Classification in Adverse Illuminous Conditions for Intelligent Transportation Systems. *Complexity*, *2021*. https://doi.org/10.1155/2021/6644861

Choi, J., Chun, D., Kim, H., & Lee, H. J. (2019). Gaussian YOLOv3: An accurate and fast object detector using localization uncertainty for autonomous driving. *Proceedings of the IEEE International Conference on Computer Vision*, *2019-Octob*, 502–511. https://doi.org/10.1109/ICCV.2019.00059

Du, S., Zhang, P., Zhang, B., & Xu, H. (2021). Weak and Occluded Vehicle Detection in Complex Infrared Environment Based on Improved YOLOv4. *IEEE Access*, *9*, 25671–25680. https://doi.org/10.1109/ACCESS.2021.3057723

Han, B. G., Lee, J. G., Lim, K. T., & Choi, D. H. (2020). Design of a scalable and fast yolo for edge-computing devices. *Sensors (Switzerland)*, *20*(23), 1–15. https://doi.org/10.3390/s20236779

Huang, Y. Q., Zheng, J. C., Sun, S. D., Yang, C. F., & Liu, J. (2020). Optimized YOLOv3 algorithm and its application in traffic flow detections. *Applied Sciences (Switzerland)*, *10*(9), 3079. https://doi.org/10.3390/app10093079

Husein, A. M., Christopher, C., Gracia, A., Brandlee, R., & Hasibuan, M. H. (2020). Deep Neural Networks Approach for Monitoring Vehicles on the Highway. *SinkrOn*, *4*(2), 163. https://doi.org/10.33395/sinkron.v4i2.10553

Kasper-Eulaers, M., Hahn, N., Kummervold, P. E., Berger, S., Sebulonsen, T., & Myrland, Ø. (2021). Short communication: Detecting heavy goods vehicles in rest areas in winter conditions using YOLOv5. *Algorithms*, *14*(4). https://doi.org/10.3390/a14040114

Li, X., Liu, Y., Zhao, Z., Zhang, Y., & He, L. (2018). A deep learning approach of vehicle multitarget detection from traffic video. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/7075814

Liang, R., & Ji, G. (2022). Vehicle Detection Algorithm Based on Embedded Video Image Processing in the Background of Information Technology. *Journal of Electrical and Computer Engineering*, *2022*. https://doi.org/10.1155/2022/6917421

Liu, C., Zhang, Y., Luo, H., Tang, J., Chen, W., Xu, X., … Shen, Y. D. (2021). City-scale multi-camera vehicle tracking guided by crossroad zones. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, (1), 4124–4132. https://doi.org/10.1109/CVPRW53098.2021.00466

Liu, Y., & Zhang, G. (2021). Vehicle detection algorithm based on LW-SSD. *Journal of Physics: Conference Series*, *1748*(3). https://doi.org/10.1088/1742-6596/1748/3/032042

*name of corresponding author

Lu, J., Xia, M., Gao, X., Yang, X., Tao, T., Meng, H., … DIng, E. (2021). Robust and online vehicle counting at crowded intersections. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 3997–4003. https://doi.org/10.1109/CVPRW53098.2021.00451

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

Mahto, P., Garg, P., Seth, P., & Panda, J. (2020). Refining Yolov4 for vehicle detection. *International Journal of Advanced Research in Engineering and Technology*, *11*(5), 409–419. https://doi.org/10.34218/IJARET.11.5.2020.043

Meng, Q., Song, H., Zhang, Y., Zhang, X., Li, G., & Yang, Y. (2020). Video-Based Vehicle Counting for Expressway: A Novel Approach Based on Vehicle Detection and Correlation-Matched Tracking Using Image Data from PTZ Cameras. *Mathematical Problems in Engineering*, *2020*. https://doi.org/10.1155/2020/1969408

Nam Bui, K. H., Yi, H., & Cho, J. (2020). A multi-class multi-movement vehicle counting framework for traffic analysis in complex areas using CCTV systems. *Energies*, *13*(8). https://doi.org/10.3390/en13082036

Neupane, B., Horanont, T., & Aryal, J. (2022). Real-Time Vehicle Classification and Tracking Using a Transfer Learning-Improved Deep Learning Network. *Sensors*, *22*(10), 1–21. https://doi.org/10.3390/s22103813

Sun, L., Zhang, J., Yang, Z., & Fan, B. (2023). A Motion-Aware Siamese Framework for Unmanned Aerial Vehicle Tracking. *Drones*, *7*(3), 1–24. https://doi.org/10.3390/drones7030153

Tian, E., & Kim, J. (2023). Improved Vehicle Detection Using Weather Classification and Faster R-CNN with Dark Channel Prior. *Electronics (Switzerland)*, *12*(14). https://doi.org/10.3390/electronics12143022

Wang, X., Wang, S., Cao, J., & Wang, Y. (2020). Data-Driven Based Tiny-YOLOv3 Method for Front Vehicle Detection Inducing SPP-Net. *IEEE Access*, *8*, 110227–110236. https://doi.org/10.1109/ACCESS.2020.3001279

Yang, H., Zhang, Y., Zhang, Y., Meng, H., Li, S., & Dai, X. (2021). A Fast Vehicle Counting and Traffic Volume Estimation Method Based on Convolutional Neural Network. *IEEE Access*, *9*, 150522–150531. https://doi.org/10.1109/ACCESS.2021.3124675

Yu, B., Shin, J., Kim, G., Roh, S., & Sohn, K. (2021). Non-Anchor-Based Vehicle Detection for Traffic Surveillance Using Bounding Ellipses. *IEEE Access*, *9*, 123061–123074. https://doi.org/10.1109/ACCESS.2021.3109258

Zhao, J., Hao, S., Dai, C., Zhang, H., Zhao, L., Ji, Z., & Ganchev, I. (2022). Improved Vision-Based Vehicle Detection and Classification by Optimized YOLOv4. *IEEE Access*, *10*(i), 8590–8603. https://doi.org/10.1109/ACCESS.2022.3143365

Zhao, M., Zhong, Y., Sun, D., & Chen, Y. (2021). Accurate and efficient vehicle detection framework based on SSD algorithm. *IET Image Processing*, *15*(13), 3094–3104. https://doi.org/10.1049/ipr2.12297

Zhu, J., Li, X., Jin, P., Xu, Q., Sun, Z., & Song, X. (2021). MME-YOLO: Multi-sensor multi-level enhanced yolo for robust vehicle detection in traffic surveillance. *Sensors (Switzerland)*, *21*(1), 1–17. https://doi.org/10.3390/s21010027

*name of corresponding author