Research Article

Implementation of the Convolutional Neural Network Method in Highway Traffic Monitoring Systems

[1]Atthariq, [2,*]Azhar, [3]Hendarawaty, [4]Jumadi Mabe Parenreng

[1,2,3] Department of Information Computer Technology, Politeknik Negeri Lhokseumawe, Lhokseumawe, 24301, Indonesia
[4] Department of Computer Engineering, Universitas Negeri Makassar, South Sulawesi, Indonesia

* Corresponding Author: azhar.tik@pnl.ac.id

Abstract: Traffic Flow Calculation is one of the first steps in urban planning and road infrastructure management, for monitoring traffic flow on a road is very important. To do traffic planning, the Department of Transportation must count every passing vehicle where later the data will be used as material for analysis. Currently, the Department of Transportation calculates vehicles that pass on a road by calculating it with manual tools, so it requires large operational costs and takes a long time. Based on the problems faced, the researcher offers a solution for an intelligent traffic flow monitoring system that can count the number of vehicles that pass by using a closed circuit television camera (CCTV) installed at every city traffic light. Here we propose a high-performance algorithm model You Only Look Once (YOLO), which is based on the TensorFlow framework, to improve real-time vehicle monitoring. From the results of testing the system was built using the Python programming language using the YOLOv4 method, the Tensorflow library, and the PyQT5 library. The accuracy of reading the number of passing vehicles is 97%

Keywords: Artificial Intelligence, YOLO, vehicles, monitoring system, intelligent traffic

check for updates

Citation: Atthariq, A., Azhar, Hendarawaty, H., & Mabe Parenreng, J. (2025). Implementation of the Convolutional Neural Network Method in Highway Traffic Monitoring Systems. Iota, 5(1). ISSN 2774-4353. https://doi.org/10.31763/iota.v5i1.862

Academic Editor: Adi, P.D.P

Received: November 02, 2024

Accepted: November 22, 2024

Published: February 01, 2025

1. Introduction

The higher level of urbanization affects the number of vehicles on the urban road network. This also causes traffic jams and the problem of accident incidents that result in substantial economic losses and disrupt people's daily lives. According to statistical data released by the Indonesian Central Statistics Agency (BPS), there has been a significant increase. Table 1 presents data on the number of traffic accidents from 2014 to 2019.

Therefore, traffic emergencies must be handled in an intelligent transportation system to ensure safe, responsive, and efficient transportation for everyone (Peppa et al. 2018; Fedorov et al. 2019). Understanding road traffic behavior is a key component of a traffic emergency response plan. Traffic flow estimation is the first step to identifying road traffic patterns, contributing to the traffic modeling, urban planning, and design process for all aspects of the road network (Fedorov et al. 2019). This is also related to the planning for the construction of a road network in urban areas by the transportation agency.

Therefore, to improve traffic services, the Department of Transportation is responsible for this. To carry out traffic development planning, the Department of Transportation requires data on vehicles that pass at the intersection to be developed in real-time, for now, the Department of Transportation performs a manual calculation of vehicles that pass, resulting in time-consuming and large costs. Overcoming this problem is by developing real-time traffic flow monitoring technology. The purpose of developing this technology is to facilitate the process of calculating passing vehicles to improve the performance of the Department of Transportation. Traffic Flow Monitoring was developed by utilizing cameras which will later be placed on traffic lights. One approach is to utilize digital image processing. Digital image processing is the science of computer programming to process and ultimately understand images and videos or make computers view them (I. Culjak et. Al 2012).

Table 1. Number of Accidents, Deaths, Serious Injury, Minor Injury, and Material Losses 2014-2019

Accident	Number of Accidents, Deaths, Serious Injury, Minor Injury, and Material Losses					
	2014	**2015**	**2016**	**2017**	**2018**	**2019**
Number of Accidents	95906.00	96233.00	106644.00	104327.00	109215.00	116411.00
Dead Victim (Person)	28297.00	24275.00	31262.00	30694.00	29472.00	25671.00
Serious Injury (Person)	26840.00	22454.00	20075.00	14559.00	13315.00	12475.00
Minor Injury (Person)	109741.00	107743.00	120532.00	121575.00	130571.00	137342.00
Material Losses (Million Rupiah)	250021.00	215892.00	229137.00	217031.00	213866.00	254779.00

Fully automated traffic flow monitoring must go through several stages, such as detection and classification (J. Redmon et al. 2018). Detection of objects in digital image processing is a process used to determine the existence of certain objects in a digital image. This detection process serves to distinguish the types of vehicles that pass in an image, to increase accuracy and speed in the detection process, a modern method is used, namely You Only Look Once (YOLO) (J. Xu, F. Pan, et.al, 2024), (R. Hamzah, et.al, 2024), (H. Liu, 2024), (Y. Qi and H. Sun, 2024).

The You Only Look Once (YOLO) method is one of the fastest and most accurate methods for object detection and is even capable of exceeding 2 times the capabilities of other methods. This method is very suitable to be implemented in the case of object detection in real-time because this method uses the GPU to increase the speed and accuracy of object detection to produce fps of 40-90 per second while object detection using CPU produces 1-2 fps (J. Redmon and A. Farhadi. 2018).

The method that will be applied in this study is the latest version of the YOLO method (Y. Huang, Q. Chai and W. Wang, 2024), (A. Liu, Y. Liu and S. Kifah, 2024), and (Y. Ji et al, 2024), namely 4 by using this method it is hoped that vehicle objects can be detected properly and can classify vehicles based on their type with a maximum fps speed to produce a system that can be used directly at the Lhokseumawe City Transportation Office.

2. Research Methodology

2.1 Monitoring System

This study aims to monitor traffic flow from low-quality CCTV video data and will classify vehicles. To achieve this goal, researchers used the new YOLOv4 algorithm using data collected from CCTV cameras to detect five classes of objects namely, cars, buses, trucks, vans, and bicycles. After that, the trained weights are used to calculate and classify the types of vehicles. Vehicle detection draws a bounding box around the vehicle object to locate it within the frame. In this study, we will discuss a bit of the CNN method as a comparison to the method that will be used in this study.

Figure 1. Block diagram for detecting vehicles in real

2.2 OverFeat

One of the first architectures to use deep learning to detect objects was OverFeat. Published in 2013 by researchers from New York University, OverFeat utilizes a multi-scale sliding window algorithm and CNN (Fedorov et al. 2019; Oltean et al. 2019):

2.2.1 R-CNN (Regional Convolutional Neural Network)

The problem with using the sliding window technique is the large number of image pieces that must be processed by CNN. Each piece will go through a convolution process, to then be classified into a background or object. This means that, with so many image locations and the size of the sliding window used, the computation of the entire process will be very heavy(Fedorov et al. 2019; Oltean et al. 2019). Moreover, The R-CNN process itself consists of 3 stages;

- look for regions or parts of the image that may be an object, with the proposal region method. One example of a region proposal technique is selective Search.
- Each region is then used as input for CNN as feature extractors from each region.
- Each feature generated, then becomes input for SVM (which will generate classes from the region) and linear regressor (which will produce a bounding box). (Fedorov et al. 2019; Oltean et al. 2019).

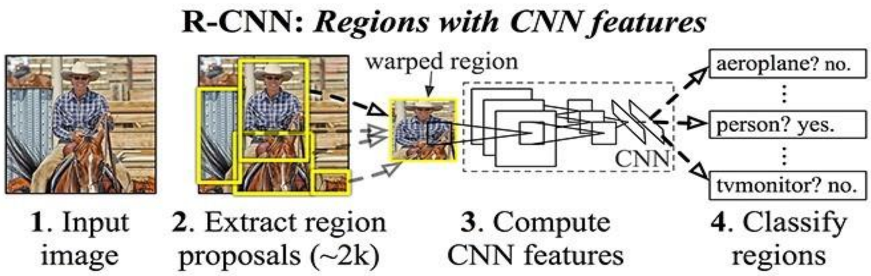

Figure 2. Two-stage detection vs single-stage detection R-CNN architecture

This process will cut a large image into smaller parts as in OverFeat and R-CNN will reduce the amount that has to be processed into multiple regions. That way, R-CNN managed to increase the performance of OverFeat by almost 50% (Fedorov et al. 2019; Oltean et al. 2019).

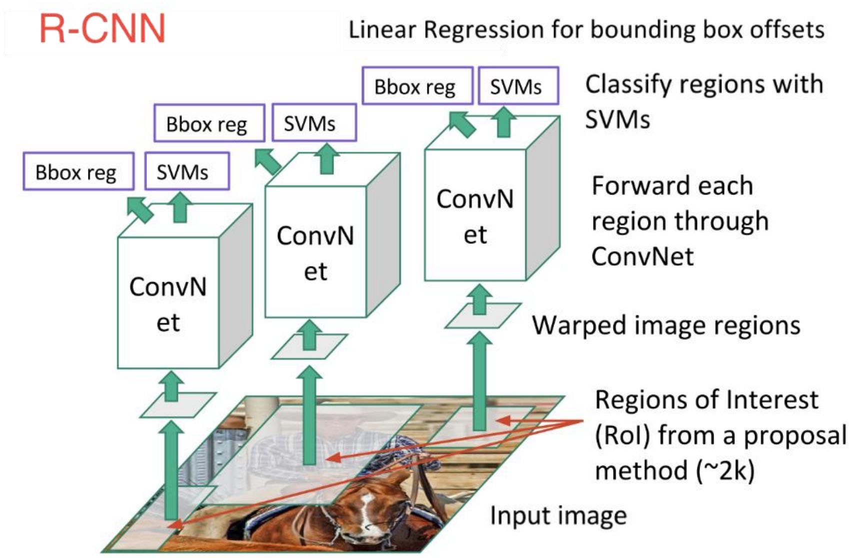

Figure 3. Architecture R-CNN in details

Naturally, techniques or algorithms that use separate processes are becoming obsolete because they generally increase training time and decrease accuracy. The development trend of deep learning architecture for object detection is starting to move towards an architecture that can detect objects more quickly and more efficiently.

2.3 *You Only Look Once (YOLO)*

You Only Look Once or what is usually called YOLO is a method that can detect and recognize various objects in an image in real time. The YOLO method uses a Convolutional Neural Network (CNN) to predict various class probabilities and bounding boxes simultaneously (B. Gong et al. 2020). Based on developments to date, the YOLO method has developed up to the 4th version. But for the 4th version itself, the method is no longer developed by the original inventors, namely Joseph Redmon and Ali Farhadi due to several reasons.

The fourth version of YOLO was developed by a Russian programmer, Alexey Bochkovskiy, and the official journal was issued by two researchers from Taiwan, Chien Yao Wang and Hong Yuan Mark Liao. They developed a CNN method that can operate in real-time using a conventional GPU to detect objects with high speed and accuracy (J. Redmon and A. Farhadi. 2018). The architecture of the method can be seen in Figure 1.

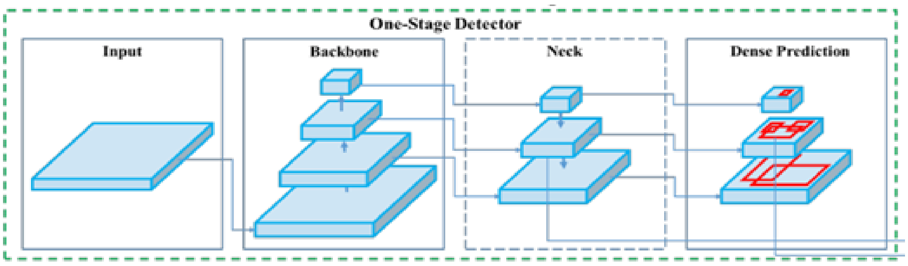

Figure 4. One Stage Detector details

The Yolov4 Backbone section implements CSPDarknet53 (J. Redmon and A. Farhadi. 2018). The CSPDarknet53 network is a CSPNet (Cross Stage Partial Network) network that was added to the base of Darknet53. Darknet53 draws on the idea of ResNet, i.e. residual connection, to ensure that the network has depth as well as reduce the missing gradient problem. CSPNet can improve CNN's learning ability while reducing the number of computations and memory usage. A good detector should have a larger receptive field. The Yolov4 neck uses two networks, SPP and PAN (J. Redmon and A. Farhadi. 2018). The SPP network applied to the neck can effectively improve field receptivity and help separate contextual features. PANet (Path Aggregation Network) plays a role in shortening the path that connects low-level and high-level information, as well as convergent parameters at different levels. The Yolov4 network head inherits the Yolov3 head structure.

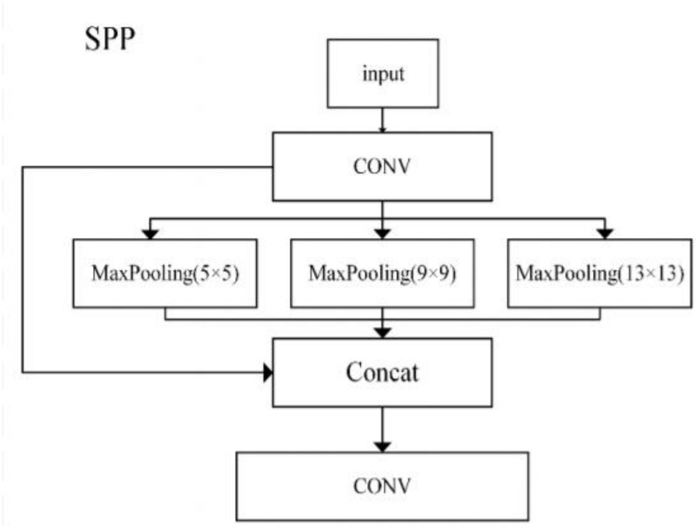

Figure 5. Spatial pyramid network (SPP)

Head is used to predict the bounding box of the object, and displays the coordinates of the center, width, and height, namely {xcenter, ycenter, w, h} (B. Gong et al. 2020). Then the predicted bounding box expression is as follows equation 1, 2, 3, and 4.

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w . e^{t_w} \quad (3)$$

$$b_h = p_h . e^{t_h} \quad (4)$$

Description: p_w and p_h represent the width and height of the previous bounding box, respectively. c_x and c_y are the coordinates of the top left corner of the image. Moreover, Figure 4 describes the process carried out by SPP where the input image results will be convoluted first to speed up the object detection process. Figure 5 describes the bounding box on an object that is detected as a result of the conversion between the actual bounding box and the predicted bounding box.

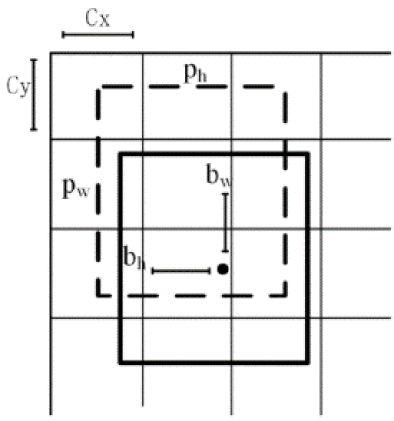

Figure 5. Conversion of the Real Barrier Box and the Predicted Result Barrier Box

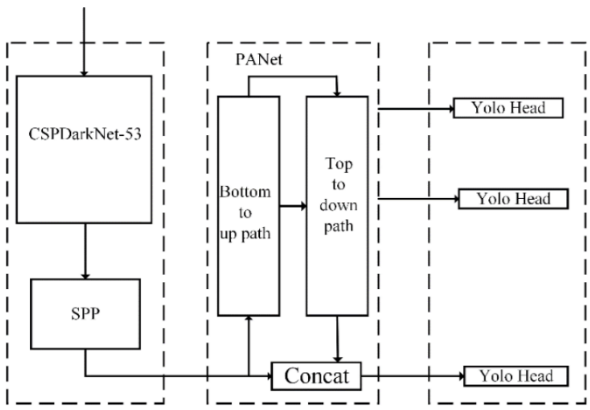

Figure 6. YOLO v4 structure

Figure 6 describes the overall structure of YOLOv4 where the input image will be processed by CSP53 Darknet then SPP, PANet, and YOLO Head to perform the detection process on an image.

3. Result and Discussion

Researchers used a CCTV image and video dataset from the Transportation Service of Lhokseumawe City, Aceh - Indonesia. As a case study in this study, the researcher chose a busy road, namely "Merdeka Barat Road –Lhokseumawe City". This image data is used to train YOLOv4 while the recorded data set is used to validate the traffic flow calculation algorithm in real-time. CCTV cameras in the location are selected to be set at a frequency resolution of 10 frames per second (fps) and a resolution of 1280x720. Table 2 and Table 3 summarize the details of the image and video datasets respectively.

Table 2. Details of the image dataset used to train YOLOv4

Vehicle Class	Total Instances
Car	827
Bus	145
Van	565
Truck	650
Bike	1520

Table 3. Details of the analyzed CCTV videos

Description	Start Time	Finish Time	Duration
Video 01	08:00:00 (UTC + 7:00)	12:00:15 (UTC + 7:00)	4 hr, 00 mins & 15 secs
Video 02	13:00:11 (UTC + 7:00)	18:06:00 (UTC + 7:00)	4hr, 06 mins & 11 secs
Video 03	19:00:49 (UTC + 12:00)	22:00:45 (UTC + 7:00)	3 hr, 00 mins & 45 secs

3.1 *User Interface*

The user interface aims to make it easier for users to interact with the system. The simpler and less complicated the facilities used, the more helpful the user is in understanding the processes carried out by the system. The following is the result of the implementation of the system user interface design. Figure 7 is a splash page that serves to greet the user by conveying some information, namely what system is being run and information on the system maker as well as preparation of the main page. Figure 8 is the main page that functions as a place for users to carry out their activities, inputting video, viewing detection results, viewing graphic visualizations, and saving calculation results into a CSV file.

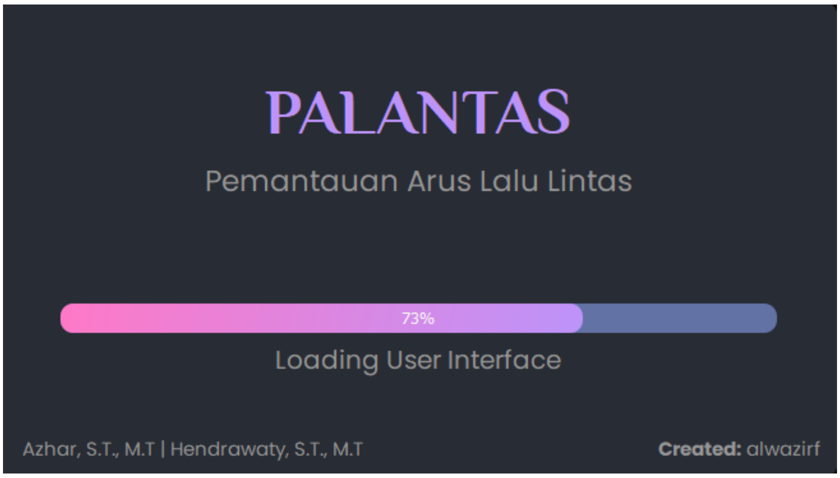

Figure 7. Splash Screen Design Results

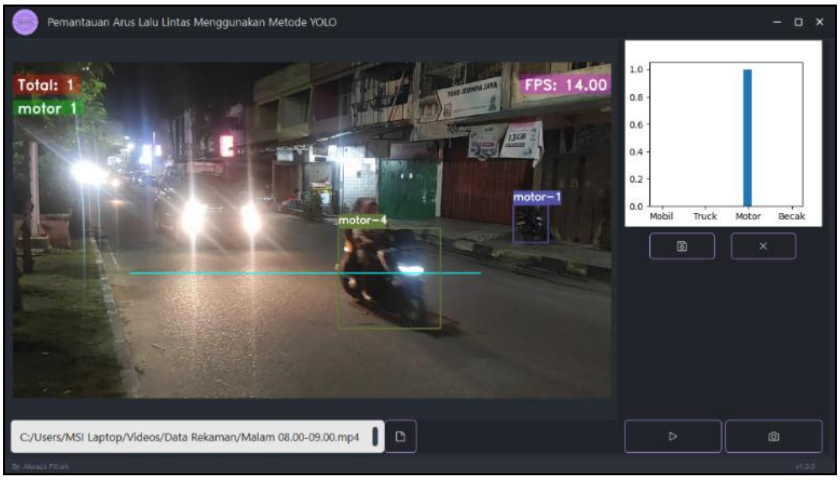

Figure 8. Test results at the Crossroads

3.2 Test Results

The average value of the YOLOv4 model vehicle detection accuracy rate as a whole is 95.1%, and the performance in each class follows the dataset, the results can be seen in Table 4.

Table 4. Average level of vehicle detection accuracy

Class	Average Precision
Car	96.34%
Bus	95.44%
Van	96.25%
Truck	90.15%
Bike	97.32%

4. Conclusion

After conducting research and discussion on Traffic Flow Monitoring Using the YOLO Method Case Study of one of the roads in Lhokseumawe City, it is concluded that the YOLO method works very well at detecting an object so that the system created can be used directly by the Department of Transportation (DISHUB) of Lhokseumawe City. With a high percentage of accuracy, namely the accuracy of object classification based on vehicle type of 95.1%.

In the study, researchers focused on real-time vehicles using low-quality CCTV footage by choosing one of the roads in the city of Lhokseumawe. By training the YOLOv4 model to detect cars, buses, vans, trucks, and bicycles. The test results from this study obtained high accuracy but due to our unbalanced training image dataset, we obtained a lower accuracy score for the truck class. Therefore, using a better data set to be trained in YOLOv4 with a larger number of vehicles for low accuracy classes.

Acknowledgments: Our thanks go to colleagues who have helped in collecting data and suggestions that are very useful in this research. We also don't forget to thank the Lhokseumawe City Transportation Office for helping us to get data and to students who support our work in the field.

Author contributions: The authors were responsible for building Conceptualization, Methodology, analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision of project administration, funding acquisition, and have read and agreed to the published version of the manuscript.

Funding: The study was conducted without any financial support from external sources.

Availability of data and Materials: All data are available from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Additional Information: No Additional Information from the authors.

References

1. Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. Retrieved from http://arxiv.org/abs/2004.10934.
2. Algiriyage, N., Prasanna, R., Doyle, E. E., Stock, K., & Johnston, D. (2020). Traffic flow estimation based on deep learning for emergency traffic management using CCTV images. In: 17th International Conference on Information Systems for Crisis Response and Management, pp. 100–109.
3. Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and real-time tracking. In 2016 IEEE International Conference on Image Processing (ICIP), pp. 3464–3468.
4. Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. In arXiv preprint arXiv:2004.10934.
5. Bui, K.-H. N., Yi, H., Jung, H., & Cho, J. (2020). Video-based traffic flow analysis for turning volume estimation at signalized intersections. In Asian Conference on Intelligent Information and Database Systems. Springer, pp. 152–162.
6. Gong, B., Ergu, D., Cai, Y., & Ma, B. (2020). A method for wheat head detection based on Yolov4. Res. Sq., pp. 1–16.
7. Culjak, I., Abram, D., Pribanic, T., Dzapo, H., & Cifrek, M. (2012). A brief introduction to OpenCV. In MIPRO 2012 - 35th International Convention on Information and Communication Technology, Electronics, and Microelectronics - Proceedings, pp. 1725–1730.
8. Im, H., Hong, B., Jeon, S., & Hong, J. (2016). Big data analytics on CCTV images for collecting traffic information. In 2016 International Conference on Big Data and Smart Computing (BigComp). IEEE, pp. 525–528.
9. Indira, K., Mohan, K., & Nikhilashwary, T. (2019). Automatic license plate recognition. In Recent Trends in Signal and Image Processing. Springer, pp. 67–77.
10. Redmon, J., & Farhadi, A. (2018). YOLO v.3. Tech Report, pp. 1–6. Retrieved from https://pjreddie.com/media/files/papers/YOLOv3.pdf.
11. Ke, R., Li, Z., Tang, J., Pan, Z., & Wang, Y. (2018). Real-time traffic flow parameter estimation from UAV video based on ensemble classifier and optical flow. IEEE Transactions on Intelligent Transportation Systems, 20(1), pp. 54–64.
12. Laroca, R., Severo, E., Zanlorensi, L. A., Oliveira, L. S., Gonçalves, G. R., Schwartz, W. R., & Menotti, D. (2018). A robust real-time automatic license plate recognition based on the YOLO detector. In 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, pp. 1–10.
13. Peppa, M., Bell, D., Komar, T., & Xiao, W. (2018). Urban traffic flow analysis based on deep learning car detection from CCTV image series. In International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(4).
14. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788.
15. J. Xu, F. Pan, X. Han, L. Wang, Y. Wang, and W. Li, "EdgeTrim-YOLO: Improved Trim YOLO Framework Tailored for Deployment on Edge Devices," 2024 4th International Conference on Computer Communication and Artificial Intelligence (CCAI), Xi'an, China, 2024, pp. 113-118, doi: 10.1109/CCAI61966.2024.10602964.

16. R. Hamzah, L. Ang, R. Roslan, N. H. Ibrahim Teo, K. Abdul Samad and K. A. Fariza Abu Samah, "Comparing modified Yolo V5 and Faster Regional Convolutional Neural Network performance for Recycle Waste Classification," 2024 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), Shah Alam, Malaysia, 2024, pp. 415-419, doi: 10.1109/I2CACIS61270.2024.10649835.
17. H. Liu, "Fine-Grained Classification and Anomaly Detection System of Motion Injury Images Based on Improved YOLO Algorithm," 2024 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 2024, pp. 1-5, doi: 10.1109/ICDSNS62112.2024.10690861.
18. Y. Qi and H. Sun, "Defect Detection of Insulator Based on YOLO Network," 2024 9th International Conference on Electronic Technology and Information Science (ICETIS), Hangzhou, China, 2024, pp. 232-235, doi: 10.1109/ICETIS61828.2024.10593675.
19. Y. Huang, Q. Chai, and W. Wang, "Modified YOLO network for symptom detection in panoramic oral roentgenogram," 2024 43rd Chinese Control Conference (CCC), Kunming, China, 2024, pp. 7848-7853, doi: 10.23919/CCC63176.2024.10661828.
20. A. Liu, Y. Liu, and S. Kifah, "Deep Convolutional Neural Network for Enhancing Traffic Sign Recognition Developed on Yolo V5," 2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2024, pp. 1-6, doi: 10.1109/ASSIC60049.2024.10508025.
21. Y. Ji et al., "Lightweight Road Damage Detection Algorithm based on the Improved YOLO Model," 2024 5th International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Shenzhen, China, 2024, pp. 832-835, doi: 10.1109/AIEA62095.2024.10692408.