Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 6
ISET 2024

Multi-scenario Vehicle Detection Based on the YOLOv5 Algorithm

Liang Zhou*
College of Information and Technology, Northwest University, Xi'an 710127, China
*Corresponding Author: 2022117530@stumail.nwu.edu.cn

Abstract. This paper addresses the issues of increasing traffic volume leading to frequent traffic congestion and accidents by studying the effectiveness of occluded vehicle detection based on the YOLOv5 algorithm. The aim is to explore the performance of this algorithm in various occlusion scenarios. Firstly, we introduce the occluded vehicle images used for testing and describe different types of occlusions in detail. Then, the basic principles and construction process of the YOLOv5 algorithm are explained. Subsequently, we propose a vehicle detection and recognition model based on YOLOv5 and validate the model's effectiveness through convergence analysis of the training loss curve and accuracy analysis of the test set. Finally, experimental results demonstrate that the YOLOv5 algorithm can efficiently and accurately detect vehicles in complex scenarios and partially occluded conditions, showcasing its application potential in intelligent transportation systems.

Keywords: YOLOv5, vehicle detection, intelligent transportation system, real-time detection.

1. Introduction

With the continuous rise in economic output, more and more people choose to travel by car, leading to various traffic problems such as congestion and an increase in annual traffic accidents. Therefore, in the field of intelligent transportation systems, vehicle detection, as one of the key technologies, is of great significance for traffic management, road safety, and applications like autonomous driving[1]. With the accelerating urbanization process and increasing traffic flow, accurately and quickly detecting and identifying vehicles has become an urgent issue to address. However, in real traffic scenarios, vehicle occlusion frequently occurs, posing significant challenges to detection algorithms[2]. Detecting occluded vehicles not only requires identifying partially occluded vehicles but also accurately locating their boundaries, which demands higher robustness and precision from the algorithm[3].

In recent years, researchers have proposed practical application operations by introducing new operators based on a thorough analysis of modern traditional motion target recognition algorithms, aiming to improve traditional algorithms and address difficulties in practical applications. For instance, the ViBe algorithm[4] proposed by Barnich and Van Droogenbroeck enhances the handling of dynamic backgrounds and partial occlusions through mechanisms like random sampling and spatial consistency. The PBAS algorithm[5] proposed by Hofmann et al. improves performance in complex backgrounds through adaptive background modeling. Zach et al.'s TV-L1 optical flow method[6] combines total variation regularization and L1 norm data fidelity term, enhancing robustness against occlusion and noise issues.

Meanwhile, the rapid development of deep learning technology has brought revolutionary changes to the fields of image processing and computer vision. Especially, detection algorithms based on Convolutional Neural Networks[7] already have demonstrated excellent performance in various complex scenarios. For example, Girshick's Fast R-CNN[8] and He's Mask R-CNN achieve high-precision detection in complex scenes and partial occlusions through the combination of Region Proposal Networks and fully convolutional networks. Moreover, methods like ensemble learning and transfer learning have been introduced into vehicle detection to further improve the algorithm's generalization ability and detection accuracy.

Against this background, the YOLO (You Only Look Once) series of algorithms, as representatives in the field of object detection, achieve efficient and real-time object detection through their single-stage detection architecture, performing particularly well in complex scenarios and occlusions. YOLOv5, as the latest version of this series, significantly improves detection accuracy and speed by enhancing the feature extraction network and optimizing the loss function.

This paper aims to study the effectiveness of occluded vehicle detection based on the YOLOv5 algorithm, exploring whether this algorithm can accurately detect vehicles under various occlusion conditions. Firstly, the prepared test images will be introduced, and the various occlusion methods in the test images will be detailed. Secondly, the basic principles and construction of the YOLOv5 algorithm will be presented. Then, a vehicle detection and recognition model based on the YOLOv5 algorithm will be proposed, followed by experiments on the convergence analysis of the training loss curve and the overall accuracy analysis of the test set. Finally, the detection results under different occlusion scales will be used to verify the effectiveness of the YOLOv5 algorithm in detecting occluded vehicles.

2. Test pictures and methods

2.1. Image acquisition[9]

The experimental images were collected from the overpass on the middle section of South Second Ring Road in Xi'an. A professional camera was used to capture a continuous photo dataset. The dataset has an approximate FPS (Frames Per Second) of 6-8 fps, with each image having a resolution of 2288 x 1712 pixels and a size of 1.08 MB. The specific situation is shown in Figure 1.

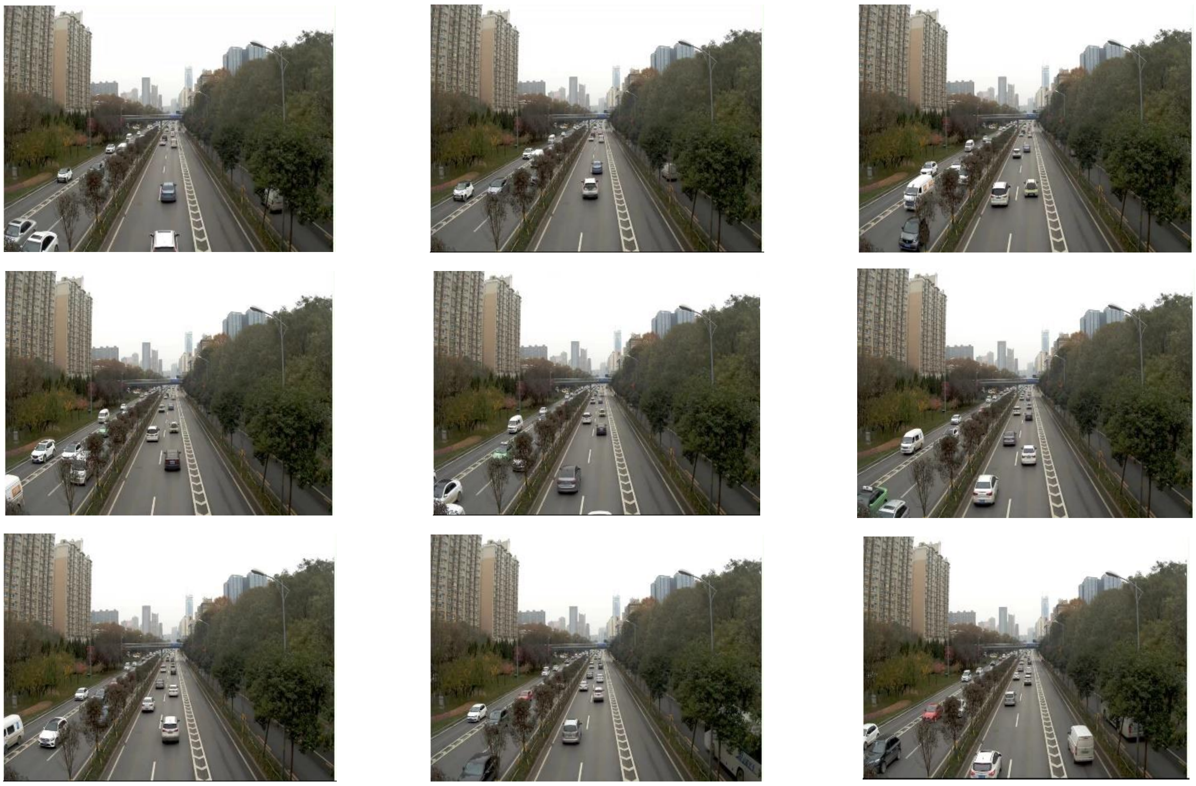


Figure 1. Test images

2.2. Sample data set and test method

The collected vehicle dataset was processed using Labelimg, categorizing the vehicles into five classes: car, taxi, minivan, truck, and motorcycle, to facilitate better target recognition. Following the labelimg process, the dataset was split into training, validation, and test sets according to the requirements for training and recognition. The training and validation sets together accounted for 80%

of the dataset, while the test set represented the remaining 20%. Within the combined training and validation sets, 80% was designated for training and 20% for validation. Once the dataset was prepared, it was used to train the YOLOv5 object detector. After completing the model training, 20% of the total dataset was used for testing. Finally, this paper compares and analyzes the existing vehicle target detection methods.

The training results, visualized using a visualization tool, are shown in Figure 2.

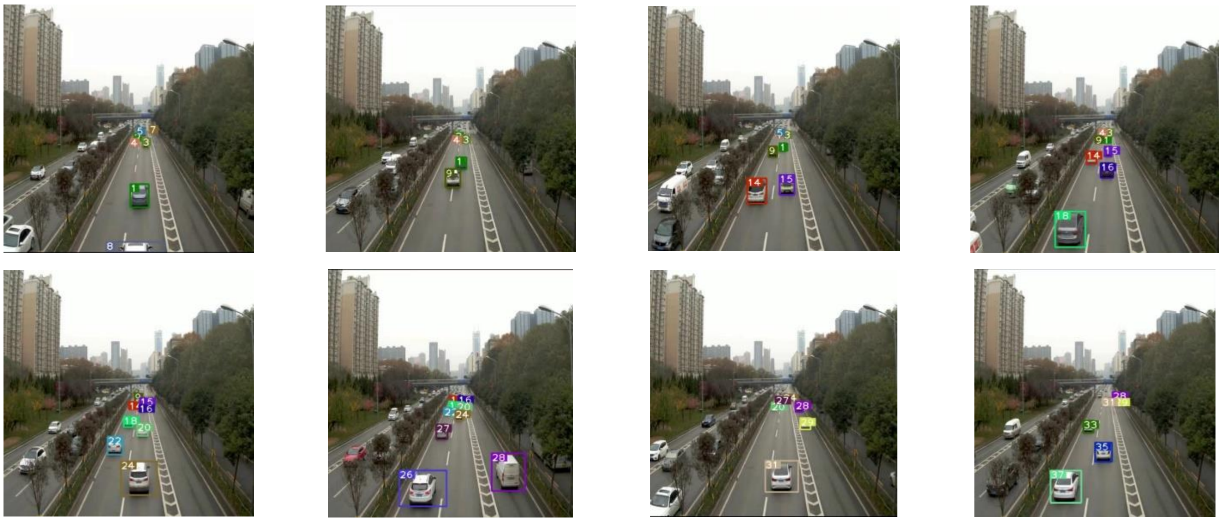


Figure 2. Training renderings

2.3. Construction of YOLOv5 network

YOLO, a single-stage object detection algorithm, has implemented various new enhancements based on YOLOv4, leading to notable improvements in both speed and accuracy.

YOLO processes the entire image simultaneously, passing it through a 20-layer GoogleNet. This is followed by four convolutional layers with activation functions(ReLU) and two fully connected layers. The dimensions are then reshaped to 7*7*30, as shown in Figure 3.

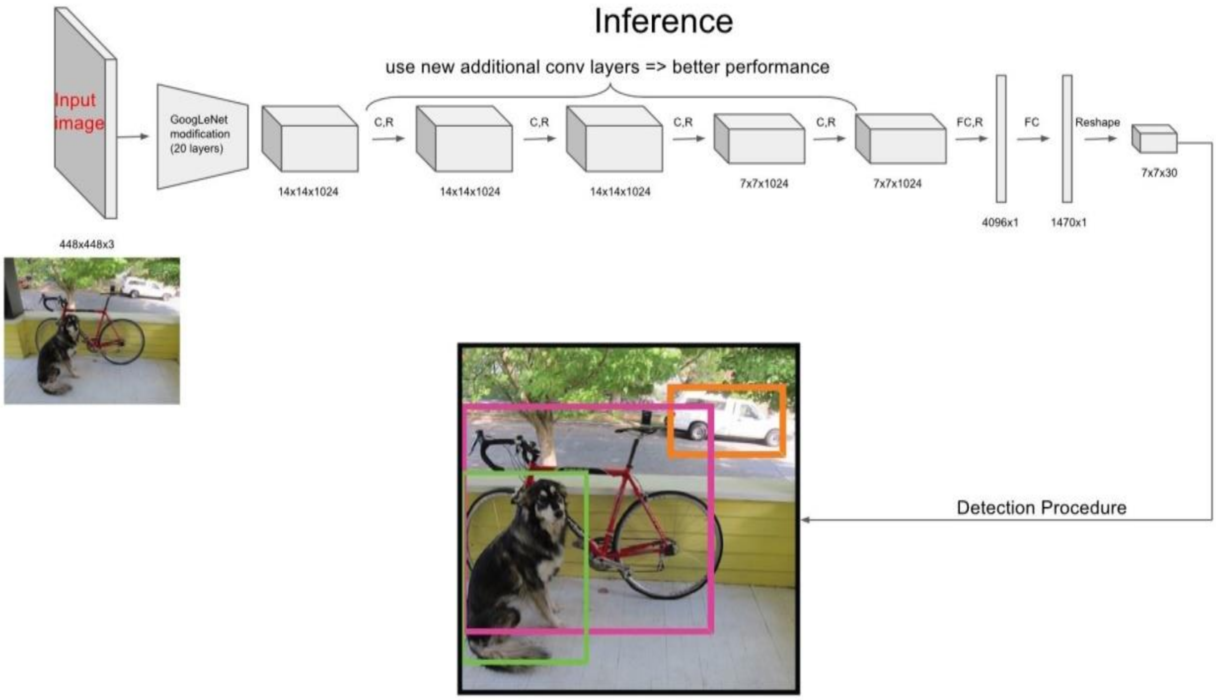


Figure 3. YOLOv5 model structure[10]

The input is partitioned into a grid of S = 7 cells (S×S). Each grid cell corresponds to two bounding boxes of varying dimensions and shapes. Each bounding box is characterized by five parameters: the x and y coordinates of the center point, the width (w), the height (h), and the confidence score. Each grid cell also has a set of conditional class probabilities equal to the number of classes C. When a grid cell contains an object, it can be encoded as a tensor:

$$S*S(B*5+C) \tag{1}$$

S×S: Total number of grid cells

B: Count of bounding boxes per grid cell

C: Quantity of classes

The term B×5 is due to each bounding box having 5 parameters, which contain the object's position information. The term C includes the object's classification information.

In formal terms, the algorithm defines confidence as

$$\Pr(\text{Object})*IOU_{pred}^{truth} \tag{2}$$

The IOU (Intersection over Union) is calculated as the ratio of the overlap area between the predicted and ground truth results to their combined area.

If a target center has no node position within this central network, Pr(Object) = 0 can be obtained, meaning the confidence score for this target is zero. Otherwise, if there exists a node within the central network, during the network training process, Confidence should approach infinity to equal the intersection (IoU) between the predicted box and the true value. Therefore, Confidence = IoU.

The loss function for YOLOv5 is described in Equation 2-3.

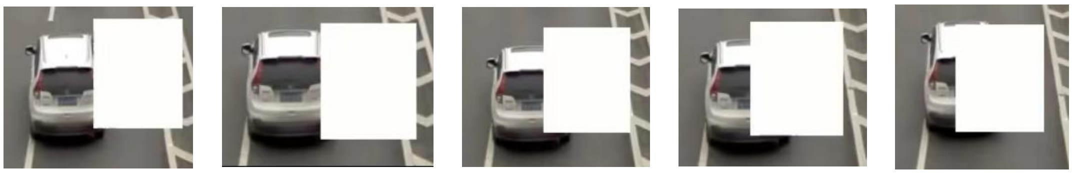
$$\begin{gathered}
\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B}\mathrm{II}_{ij}^{obj}[(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2]+ \\
\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B}\mathrm{II}_{ij}^{obj}[(\sqrt{w_i}-\sqrt{\widehat{w}_i})^2+(\sqrt{h_i}-\sqrt{\widehat{h}_i})^2]+ \\
\textstyle\sum_{i=0}^{S^2}\sum_{j=0}^{B}\mathrm{II}_{ij}^{obj}[(C_i-\hat{C}_i)^2]+\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B}\mathrm{II}_{ij}^{obj}[(C_i-\hat{C}_i)^2]+ \\
\sum_{i=0}^{S^2}\mathrm{II}_{ij}^{obj}\sum_{c\in classes}(p_i(c)-\hat{p}_i(c))^2
\end{gathered} \tag{3}$$

Based on the above, YOLO, based on the GoogleNet architecture, reads the entire image at once and utilizes grid cells and bounding boxes to simultaneously extract positional and classification features.

3. Experiment

3.1. The influence of different occlusion on algorithm effect

Firstly, preprocess the dataset as depicted in Figure 4, using images with varying degrees of white rectangular overlays as occlusions. The same vehicle is occluded to different extents from left to right to assess the impact of occlusion severity on algorithm accuracy.



(a)5% (b)10% (c)35% (d)50% (e)80%

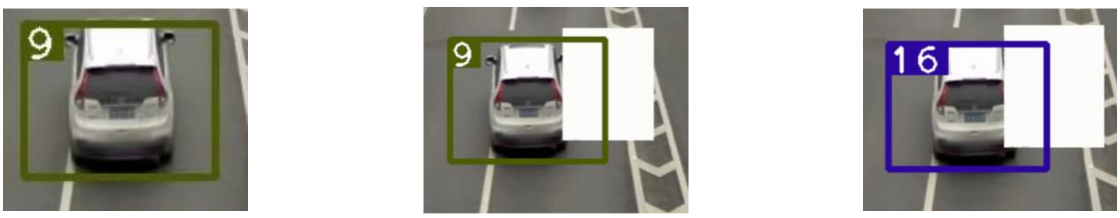
Figure 4. The same car covered to different degrees

Using visualization tools, the impact of varying degrees of occlusion on algorithm accuracy can be observed. Successful identification is indicated by accurately marking the vehicle's location with a rectangular box. The success rate for vehicle identification is calculated by dividing the count of correctly identified vehicles by the total amount of images in the test set, which includes varying levels of occlusion. The identification results are illustrated in Table 1, and the visualization of the recognition performance is depicted in Figure 5.

Table 1 Recognition results of vehicles

Parameter	Vehicle occlusion is 5%	Vehicle occlusion is 10%	Vehicle occlusion is 35%	Vehicle occlusion is 50%	Vehicle occlusion is 80%
Recognized vehicles number	200	199	190	195	186
Vehicles number	200	200	200	200	200
Recognition rates/%	100	99.5	95	97.5	93

According to Table 1, as the degree of vehicle occlusion increases, there is a slight decrease in the recognition accuracy of the YOLOv5 detection algorithm. However, overall accuracy remains above 93%, indicating excellent performance of the trained YOLOv5 algorithm in occlusion experiments.



(a) 0% (b) 5% (c) 10%

(d) 35% (e) 50% (f) 80%

Figure 5. The influence of different occlusion degree on algorithm accuracy

3.2. The influence of rainy day on algorithm effect

Firstly, preprocess the dataset by applying blur effects and adding raindrop effects to simulate rainy weather conditions, as illustrated in Figure 6. Use images captured in rainy conditions to assess the impact on algorithm accuracy.

(a) (b) (c) (d)

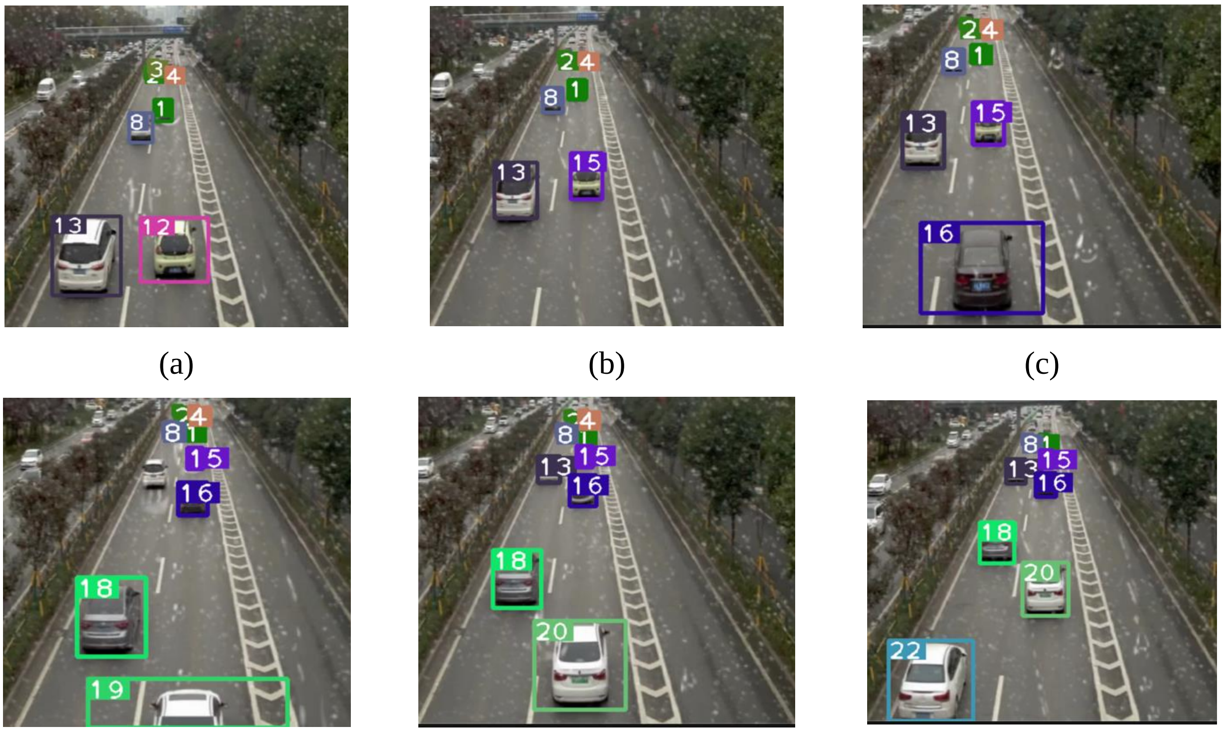
Figure 6. The influence of different occlusion degree on algorithm accuracy

With the aid of visualization tools, the impact of rainy weather on algorithm accuracy can be observed. Successful identification is indicated by accurately marking the vehicle's location with a rectangular box. The success rate for recognizing vehicles is determined by dividing the number of correctly identified vehicles by the total amount of vehicles in images. The identification results are illustrated in Table 2, and the visualization of the recognition performance is depicted in Figure 7.

Table 2 Recognition results of vehicles

Picture	Vehicles number	Recognized vehicles number	Recognition rates/%
(a)	6	6	100
(b)	6	6	100
(c)	7	7	100
(d)	9	8	88.8
(e)	9	9	100
(f)	8	8	100

According to Table 2, in severe rainy weather conditions, the average success rate of the YOLOv5 detection algorithm reaches 98.1%, indicating excellent performance of the trained YOLOv5 algorithm in adverse weather occlusion experiments.



(a) (b) (c)

(d) (e) (f)

Figure 7. The influence of rainy day on algorithm accuracy

4. Conclusion

This study investigates the effectiveness of the YOLOv5 algorithm for detecting occluded vehicles and explores its capability to accurately detect vehicles under varying degrees of occlusion and adverse weather conditions (rain). We begin by introducing the prepared test images and provide a detailed explanation of the various methods of vehicle occlusion present in these images. Subsequently, we outline the fundamental principles and architecture of the YOLOv5 algorithm. We then propose a vehicle detection and recognition model based on YOLOv5 and conduct experiments that include analyses of training loss curve convergence and overall accuracy on the test set. Finally, we validate the effectiveness of the YOLOv5 algorithm in occluded vehicle detection by observing changes in detection results under different occlusion scales.

The experimental results demonstrate that YOLOv5 achieves high detection accuracy and speed in complex scenarios and under partial occlusion, effectively identifying and locating partially occluded vehicles. This achievement provides reliable technical support for vehicle detection in intelligent transportation systems, contributing to the improvement of traffic management and road safety. Furthermore, this study serves as a reference for further optimization and enhancement of occluded vehicle detection algorithms, laying a foundation for future research in related fields. However, the study is limited in that experiments were conducted solely under rainy conditions, which somewhat restricts the comprehensiveness of the research. Future studies should focus on collecting datasets under more complex and adverse weather conditions to comprehensively evaluate the performance of YOLOv5.

References

[1] Chen, Z., et al. (2021). A Review on Recent Advances in Vehicle Detection Algorithms. IEEE Access, 9, 23212-23229.

[2] Zhang, H., et al. (2020). A Survey on Recent Advances in Vehicle Detection from Remote Sensing Images. Remote Sensing, 12(10), 1644.

[3] Duan, K., et al. (2020). CenterTrack: Tracking Objects as Points. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 9416-9426).

[4] Barnich, O., & Van Droogenbroeck, M. (2011). ViBe: A Universal Background Subtraction Algorithm for Video Sequences. IEEE Transactions on Image Processing, 20(6), 1709-1724.

[5] Hofmann, M., Tiefenbacher, P., & Rigoll, G. (2012). Background Segmentation with Feedback: The Pixel-Based Adaptive Segmenter. In 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (pp. 97-104). IEEE.

[6] Zach, C., Pock, T., & Bischof, H. (2007). A Duality Based Approach for Realtime TV-L1 Optical Flow. In Proceedings of the Joint Pattern Recognition Symposium (pp. 214-223). Springer, Berlin, Heidelberg.

[7] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems (NIPS) (pp. 1097-1105).

[8] Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1440-1448).

[9] John. (2021). YOLOV5-DeepSORT-Vehicle-Tracking-Master. Retrieved from [GitHub Repository URL]

[10] Redmon, Joseph and Farhadi.YOLOv3:An Incremental Improvement[D].arXiv.2018.