IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 7, 2022**

Improving the Efficiency of Image Processing with Deep Learning for Vehicle Detection and Tracking

Priyanka Ankireddy[1], V. Siva Krishna Reddy[2,*], Dr.V. Lokeswara Reddy[1]

[1]Department of CSE,K.S.R.M. College of Engineering (Autonomous), Krishnapuramu, Kadapa, Y.S.R. (District), Andhra Pradesh, India.

2 Dr.V.Lokeswara Reddy, Professor & HoD, Department of CSE, K.S.R.M. College of Engineering, Kadapa, AndhraPradesh India.

*Corresponding Author e-mail id: vskreddy.ysrafu@gmail.com

Abstract: Vehicle identification and tracking is an extremely important function of traffic surveillance systems that is necessary for efficient traffic management and the protection of drivers and passengers. Finding and following the path of vehicles is the primary goal of this research. The goal of this research is to develop methods for the automated identification of cars in digital photographs and moving pictures. One of the numerous uses for Deep Learning, which may include fuzzy logic, neural networks, and evolutionary algorithms, is in the detection and tracking of automobiles. The purpose of this project is to apply deep learning to the problem of vehicle recognition and tracking; the primary-stage target detection techniques will be YOLOv5 and Single Shot MultiBox Detector (SSD). This is the main topic of the article. The Single Shot MultiBox Detector (SSD) model architecture is then employed as the major foundation for vehicle detection. Focus loss, in addition to the standard SSD, is an optimization component that improves feature extraction speed. Therefore, the procedure begins with a series of training procedures on the photos included inside the publicly accessible road vehicle dataset. The vehicle recognition model is then trained using YOLOv5 and SSD algorithms; these two algorithms work together to show how effective they are at detecting vehicles. Comparing the models' detection rates on different cars is the key to locating it. The fundamental objective of this study is to develop an automated technique for detecting and tracking autos in both static and dynamic scenes. In the end, the trained network model is applied to the analysis of the vehicle camera video, and the detection performance is tested experimentally. The study's results show that the approach may enhance vehicle identification success to 97.65%. From video and picture inputs, it can reliably identify vehicles.

Keywords: Vehicle Detection, Image Processing, Vehicle Tracking, Deep Learning, Object Tracking.

1. INTRODUCTION

Both the vehicle information system and the intelligent traffic system make use of automatic vehicle data recognition. Academics have paid a lot of attention to it since the turn of the decade, thanks to developments in digital photography and processing power. The recognition of vehicles automatically is a cornerstone of many cutting-edge traffic management programs [1-3]. Among them are automated vehicle accident detection, automatic traffic density estimates, lane departure warning systems, traffic signal controllers, and traffic response systems. The strain on those in charge of managing the population and its associated infrastructure grows with each passing year.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 7, 2022**

The global population is expanding at a breathtaking pace. There was a subsequent increase in the manufacturing of automobiles and other mechanical equipment. But it's crucial to handle new problems like traffic, accidents, and other challenges with caution. In order for humanity to continue making progress toward their objectives, new discoveries and inventions have had to be developed and implemented. Congestion on main thoroughfares and in big cities is a prime example. Some of the solutions used to this issue include a traffic signal and a sign. These answers seem to be inadequate by themselves.

2. RELATED WORKS

Vehicle detection technology is analogous to target detection technology. Both vehicle detection and target detection aim to accomplish similar core goals, which may be broken down into the locations and types of targets. Historical data-based algorithms, deep learning-based algorithms, YOLO-based algorithms, and path-following algorithms are the four primary categories of vehicle identification algorithms.

i. Traditional Vehicle Detection Algorithm

Conventional techniques of vehicle detection also make use of a vehicle identification system predicated on a very simple application of machine learning. This system can identify vehicles based on their individual characteristics thanks to the integration of a vehicle-centric algorithm and a machine learning algorithm [4-6]. While the use of SVM and HOG features into shallow machine learning has the potential to improve detection accuracy, there are several downsides that must be considered. Even while this technique boosts detection accuracy by simple cascade, it also uses a larger model for detection, which requires more calculations overall. Even when using a shallow machine learning technique, feature selection is necessary to get the highest possible accuracy in vehicle recognition. The substantial modeling work required to simulate complex and ever-changing traffic circumstances is a primary cause for the delay in its development. Traditional techniques for identifying automobiles include frame difference approaches, streamer methods, and background modelling methods [7, 8]. These methods are among the most popular in use today. There are a number of other ways to detect vehicles. There are a variety of factors, such as lighting and weather, that might affect the accuracy of the identification results produced by such an algorithm [9, 10]. This makes it difficult to adapt to the ever-changing reality of traffic on the roadways and hampers the demands of real-time vehicle identification. It also makes it harder to avoid collisions with other cars. Therefore, it will be an extremely difficult task to satisfy all of these prerequisites simultaneously.

ii. Vehicle Detection Algorithm Based on Deep Learning

R-CNN with slow, medium, and quick speeds R-CNN is a common technique for two-stage target identification [11, 12]. After producing a predetermined number of targets based on regional suggestion, the most notable feature of this technique is the use of a convolution neural network to deal with prospective targets. In order to accomplish sparse sampling, (ER-CNN first takes the original picture as input and uses candidate areas. Once potential areas have been located, a convolutional neural network (CNN) collects features, and a support vector machine (SVM) assigns labels. With R-CNN, detection accuracy has dramatically increased while the algorithm's bounds have become much more reasonable [13, 14], making it the de facto standard in target detection. Faster R-CNN [15, 16] is an improved version of Fast R-CNN that speeds up region formation by using the properties of an RPN network. To begin, it checks to verify whether the candidate box contains the proper characteristics for the detection aim using a multi-task loss function. If this is the case, the procedure moves onto the next step. This will allow us to assign a name to the detected item.

iii. Vehicle Detection Using YOLO (You Only Look Once)

Initially, YOLO [17] approached object identification as a regression issue inside a single neural network. The method's rapid adoption as the de facto standard in object detection is a direct result of its stellar performance. Consistent development since YOLO's inception has resulted in five generations of the architecture: YOLO [18], YOLOv2 [19], YOLOv4 [20], and YOLOv5 [21]. The original YOLOv1 combined the three processes of feature extraction, object localisation, and classification into a single operation. Even though it had a high mAP, this network was SOTA when measuring mean average precision. The foundation of the first incarnation of the YOLO architecture were layers of convolutional and then maxpool activation functions. The network is now adaptive to picture resolution thanks in large part to the elimination of the fully-connected layer that existed at the very end of YOLOv1. The third iteration, dubbed YOLOv3, builds upon the foundation laid by its predecessors. Two prior generations, ResNet [22] and the feature-pyramid network (FPN) [23], served as inspiration for this new generation's architecture. Fast models such as YOLOv3, Faster-RCNN [24], single shot multibox object detection (SSD) [25], and Center Net [26] may achieve comparable mAPs on the COCO-2017 dataset.

iv. Vehicle Tracking

In order to find a series of images that best matches an object, conventional tracking systems often begin by recognizing objects in the first frames and then scanning the surrounding environment for characteristics that correlate to those objects. Conventional detectors including contour-based target identification [27-29], the Harris corner detector [30], symmetric integral and fluctuating transform (SIFT), and feature point-based approaches [31,32] all suffered from the same problem of false detection. Better performance was achieved, however, by using DL models to identify the objects first, and then going to match features through the traditional tracking approaches. We use DeepSORT, a tracking methodology, in combination with low-confidence track filtering, to implement the strategies presented in [33] for tracking through detection. This meant that the default DeepSORT algorithm produced less false positives. Using 3-D constrained multiple kernels, [34] recently described a method for following objects recognized by a YOLOv3 network [35].

3. PROPOSED METHODOLOGY

This study's authors suggest investigating the vehicle-recognition and tracking technique using deep learning. In this study, we use first-stage target recognition methods such the Single Shot MultiBox Detector (SSD) and YOLOv5 algorithms. The Single Shot MultiBox Detector (SSD) model architecture is then employed as the major foundation for vehicle detection. The fundamental objective of this study is to develop an automated technique for detecting and tracking autos in both static and dynamic scenes. The suggested procedure consisted of three separate actions. To begin with, YOLOv5 takes N frames at set intervals to search for and locate vehicles. To gather and evaluate characteristics of objects, we next utilize K-means clustering and the KLT tracker to follow the corner points as they travel over N-frames. The article concludes by detailing a dependable method for assigning vehicle trajectories to each of the highlighted bounding boxes. This method ensures that the labels applied to the trajectories of individual vehicles are unique from one another. You can view a diagram of the suggested solution architecture in Figure 1. Next, we'll provide a high-level overview of the most important aspects of the proposed method for detecting, tracking, and counting the number of cars.

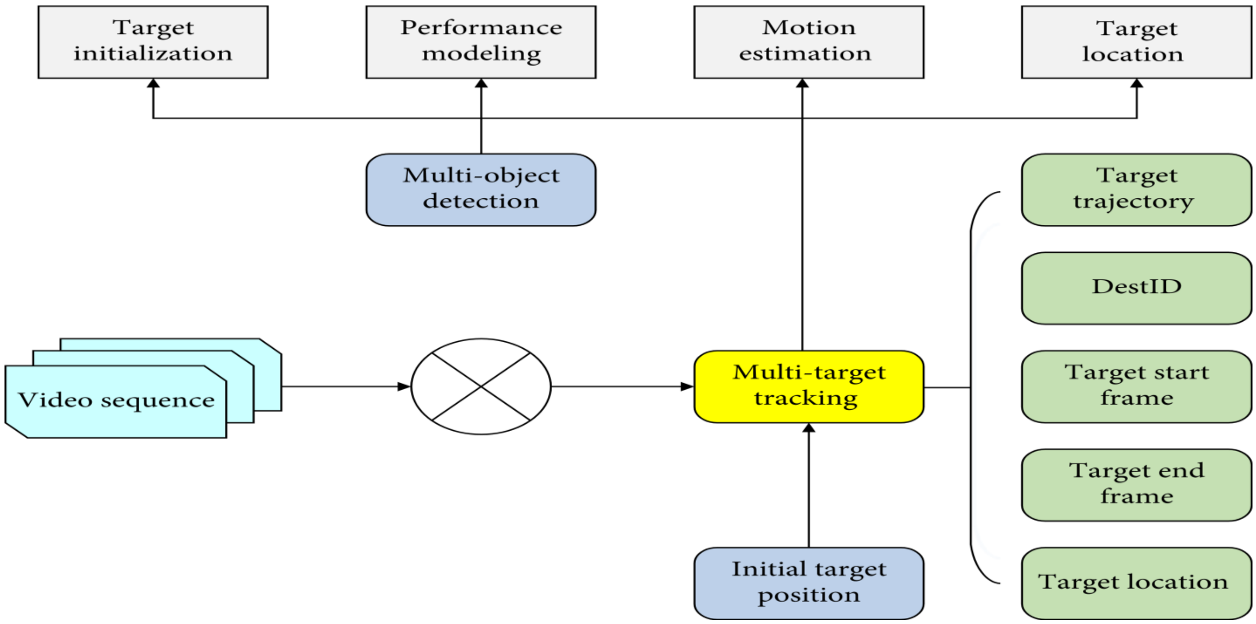

Figure 1. Architecture of vehicle detection and tracking model

i. Vehicle Detection

In this study, we use the YOLOv5 architecture to rapidly apply transfer learning to the detection process. In addition, we can make a precise assessment and substantially improve detection performance by combining the optical flow data into the counting approach suggested in [14]. In the last few layers, we use transfer learning by exchanging the softmax 1000 classes for the softmax 2 classes. Transfer training makes use of pre-trained convolutional neural network models to expedite subsequent training. These models needed a lot of training data to become this good. After developing our architecture using pre-trained models up to the last, fully-connected layer, we train it from scratch on the vehicle dataset. Our hard work has finally paid off with the completion of this layer. In [22], you can find a more thorough description of the transfer learning approach. [22] Using transfer learning, we settled on the Resnet-50 [16] as the primary neural network model for the YOLOv5 framework. Figure 2 is a block schematic of the YOLOv5 and SSD car-identification models. This shows the model's inner workings in great detail. The convolutional neural network ResNet50 was trained using over a million photos from the ImageNet dataset. There are almost a thousand different types of tags used to organize the over 1.2 million pictures in this collection. We proposed a strategy for training YOLOv5, then tested it out over three independent sessions with varying sets of photos.

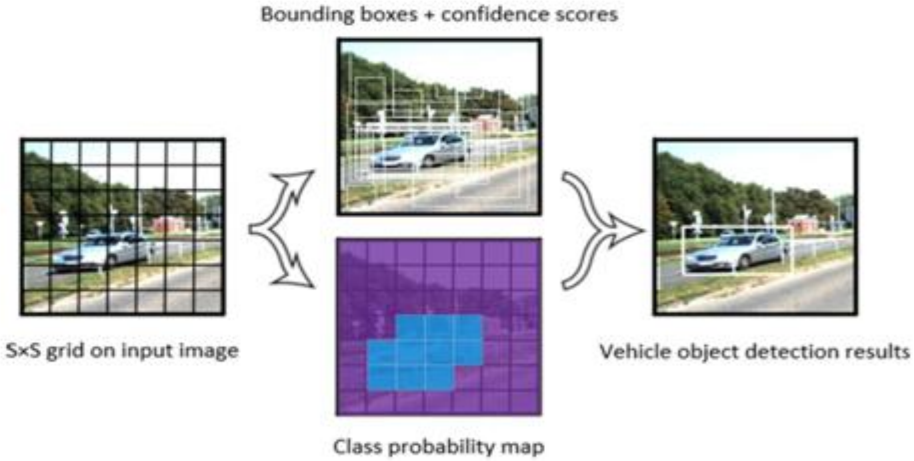

Figure 2. Architecture of YOLOv5 and SSD vehicle Detection models

ii. Vehicle features refinement and clustering

At this point, we don't only group automobiles together, we also separate them from their backgrounds by wiping them clean. Processing speed and precision in matching features are both increased by optical flow tracking. As a result, we monitor the feature points between frames f and f + 1 using the optical flow Kanade-Lucas method [2]. Combining two successive images creates a new set of optical flow vectors, V, with elements Vi = (Mi, i), where S and are defined as follows:

$$M_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

$$\theta_i = Arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \quad (2)$$

We use the notation X1 and Y1 for the X and Y coordinates of the previous frame, and X2 and Y2 for the X and Y coordinates of the next frame. In V, each number represents an individual corner point Pi captured between frames f and f+1, whereas M and denote the magnitude and angle of the corresponding displacement. Because of the erratic detections, many trackers end up failing. As noisy detections can only be followed for a limited length of time (9 consecutive frames), this study only considers a foreground detection to be a vehicle object if it has been tracked for a sufficient number of consecutive frames. After that, we use k-means clustering to group the remaining automobiles in the foreground of the picture. For further information, you may check out [36].

iii. Vehicle Counting

After identifying the most trustworthy characteristics, we put them into their own categories for each car. We assign unique ID numbers to each of these vehicle components and track them until they are no longer visible. The allocation method starts with a region determined by the intersection of rectangular bounding boxes based on historical tracking and current detection. Two automobiles are deemed to be a perfect fit if they share a certain

ISSN PRINT 2319 1775 Online 2320 7876
Research paper , **Volume 11, Iss 7, 2022**

minimum percentage of space. The vehicle will get a new label if the junction area is either less than or does not exist at all.

4. IMPLEMENTATION OF VEHICLE DETECTION AND TRACKING USING YOLOv5

Small-sized target objects, certain size scaling in the process of continuous detection, the complexity of the vehicle environment, too many targets in a single image in the dataset, and overlapping of targets are just some of the difficulties encountered when attempting to detect the vehicle target using the vehicle dataset. The correct detection technique is very important since the success of a vehicle detection system is contingent on achieving a number of requirements. Despite the challenges, there are a few things about cars that are immediately clear.

i. Analysis of Algorithm Selection

The algorithmic development process should prioritize speed and accuracy in target detection systems, since they are crucial in practical applications. Better item recognition is possible with larger detection accuracies, and faster detection rates make the technique applicable to more devices. Think about it In Table 1, we compare the detection rates and accuracy ratings of a number of widely used target identification techniques. Table 1 shows that Faster R-CNN has a higher detection accuracy than other one-stage detection networks, but a considerably different detection rate.

Training Set	Algorithms			
	FRCNN	YOLO	SSD	YOLOv5
VOC2007 + 2012	0.765	0.772	0.783	0.792
VOC2007 + 2012	0.958	0.963	0.972	0.987
VOC2007 + 2012	0.865	0.871	0.883	0.896
VOC2007 + 2012	0.784	0.789	0.793	0.798
Training Set	2010	2010	2010	2010
MAP	78.4	68.8	81.3	91.7
FPS	21	49	43	54

Table 1. Comparing the detection-rate and accuracy-rate of a number of popular target-detection algorithms.

ii. Designing the Model

a. Initialization Operation of Candidate Box

The most important part of the detection process, network training, requires initialization of the network model's parameters.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 7, 2022

b. Detection Module of Network

YOLOv5 is able to do feature extraction because to the support of DarkNet-53. The network is able to extract more feature information because of its deep-level structure. But as seen in Figure 3, there are also problems for the network at extremely deep levels. As a result, training a deep network will result in a decrease in the network's effectiveness while attempting to recognize small objects. Users of both the deep neural network and other one-stage detection algorithms have recently noticed these issues. DarkNet-53 is able to considerably enhance the learning capacity to image features since it is built on the concept of residual networks and makes use of residual connections. In addition, it compensates for the fact that it can't pick up on the tiniest of details.

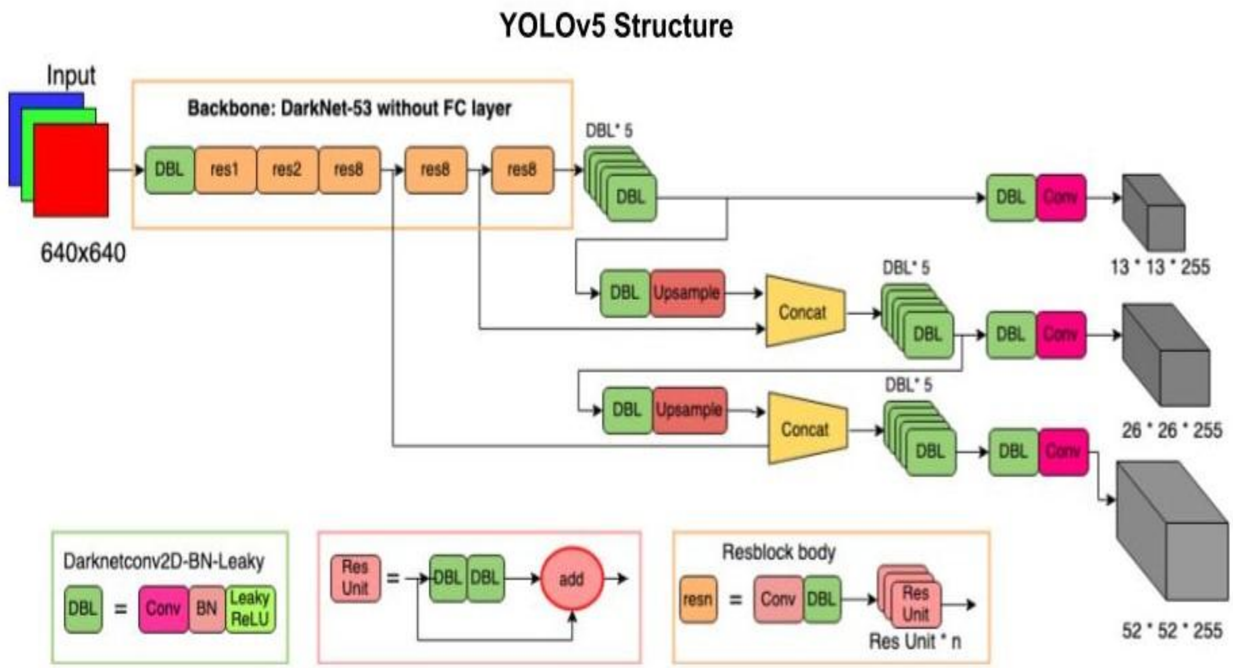

Figure 3. Network architecture of YOLOv5 with a backbone ofDarkNet-53.

5. EXPERIMENTAL RESULTS AND ANALYSIS

This section primarily compares the speed and performance of the proposed vehicle recognition and tracking method against that of current high-performance CF trackers using publicly accessible datasets, and analyzes the results. In this subsection, we will compare the performance of the proposed algorithm for vehicle recognition and tracking to that of the high-performance CF trackers.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 7, 2022

i. Dataset selection

When developing a system for target identification using a deep learning technique, the selection of an appropriate dataset is a crucial and challenging step. In this study, we are developing a model and have chosen to train it using the BDD100K picture dataset. Figure 4 displays some representative data from the set. All of the images included in the offered data were taken from a moving vehicle on a public route. These pictures show a broad variety of vehicles, as well as human and non-human victims. We only shot at six of the 10 target categories available to us. Vehicles, buses, passengers, trucks, motors, and bicycles all count as means of transportation.


(a)


(b)


(c)


(d)

Figure 4. Data images of BDD100K

ii. Processing of Data

Both a "picture" and a "label" portion make up the downloaded data. All the fresh information was actually saved as JSON files, and each JSON file does indeed map to an image with the same name. The naming conventions reveal this to be true. It includes details such as the image's filename, the category to which it belongs, the category's name, and the image's coordinates for the indicated box.

iii. Detection of Model

To complete the process of building a vehicle detection model, one must first train the model before loading the trained weight onto the model to carry out the actual detection. The training technique may make use of the parameter sets shown in Table 2. Target identification throughout the three backbone network layers is where YOLOv5 often shines in both training and detection. These are the lowest, middle, and highest tiers of the network's central infrastructure.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 7, 2022

Models	**mAP/%**	**Vehicle Detection and Tracking/FPS**	**Memory size**
FRCNN	92.4	20	200.3
YOLOv3	85.7	30	243.6
SSD	92.5	35	15.7
YOLOv5	97.6	39	15.9

Table 2. Parameter settings of the training process on various models

iv. Result Analysis of vehicle detection and tracking

Figure 5 displays the results for both the Track Maintenance dataset and the BDD100K dataset.Our method's recognition accuracy is higher than that of Faster R-CNN, YOLOv3, SSD, and YOLOv5, as shown in Figure 5. Our improved YOLOv5 model has a 5.1% higher mAP value (mAP at 0.5: 0.05: 0.97) than the previous version. Our enhanced model is 1.12 times faster than YOLOv5, 1.45 times faster than YOLOv4, and 2.9 times faster than Faster R-CNN in terms of detection speed, indicating that it can match the requirements of real-time detection. Compared to the YOLOv5 algorithm, our enhanced technique uses somewhat more memory, but still only about a seventeenth as much as YOLOv3 and about a thirteenth as much as Faster R-CNN.

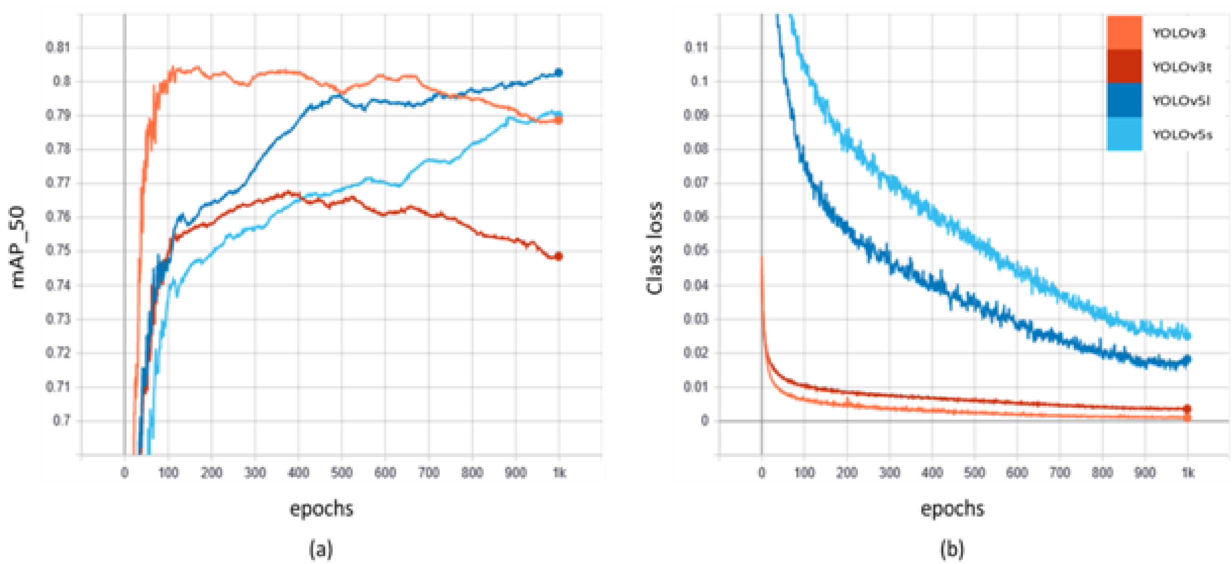

Figure 5. Training results of FRCNN, SSD, YOLOv3t, YOLOv5l, and YOLOv5s on the BDD100K dataset. (a) mAP at IoU = 50 (b) loss of classification.

The image's orange curve shows the loss value for the test set, while the image's blue curve shows the loss value for the training set. After a few repetitions of training, the initial high loss numbers begin to decrease. Values start to converge and the rate of decrease slows. Comparing results after training and testing reveals that the model is steadily nearing convergence, even though the loss value is still rather large. The values of the training set loss tend to decline with time, stabilize, and then fluctuate within a small range. Adjusting the learning rate in accordance with the breakpoint continuation strategy is necessary to keep training going, even if the final results won't be that

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 7, 2022**

different from the second portion of Figure 6. The detection box scores in Figure 6 remain constantly above 0.7, indicating that the model can correctly categorize the great majority of cars. The detection impact of trucks is around average, that of bus types is a little lower, and that of detection boxes is typically minimal. In Figure 7, we can see how well the network can recognize both people and bicycles.

Figure 7 demonstrates that although the detection model does not substantially affect either the person or bicycle categories, the human category scores much higher than the bicycle category scores. The extension of available box possibilities and the subsequent increase in "person" detection box accuracy are other noteworthy developments. Figure 8 demonstrates the detection method's outcomes, providing evidence that the bus category detection effect is general.

Figure 6. Detection of vehicles on the road from different camera views by day scenes and night scenes

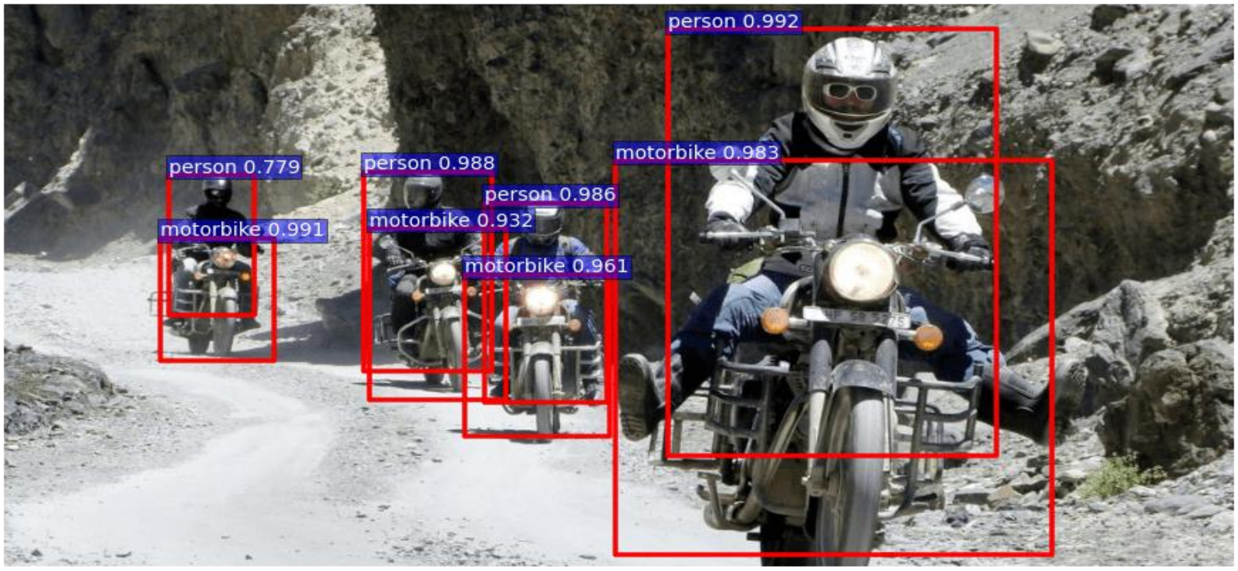

Figure 7. Illustrating the detection on bike category and person category

Figure 8. Vehicle detection on bus category

6. CONCLUSION

In this study, we use deep learning to automate the process of music production. For tracking down people and machines on the track, we also suggest updating the YOLOv5 and SSD algorithms. We make substantial use of the YOLOv5 and SSD algorithms to dramatically enhance the mean absolute performance (mAP) at IoU and the loss of classification for the different sets of training outcomes. Newer YOLOv5 and SSD algorithms allow for faster convergence and better recognition of obstructed vehicle objects and tiny vehicle objects. These benefits are cumulative. The testing findings demonstrate the excellent resilience of the newly developed YOLOv5 and SSD technique. By applying these algorithms, we are able to perform thorough inspections of construction workers and equipment, addressing the problem of low detection accuracy for complex scene issues like occluded vehicle objects and small vehicle objects, and meeting the practical requirements for vehicle detection in the context of track construction safety. The findings of this work provide credence to the practical use of intelligent detection tools and lend momentum to the thorough investigation and advancement of track safety vehicle detection technology. According to the measured KPIs, the combination of the YOLOv5 and SSD algorithms is the most successful in terms of both vehicle detection and tracking precision.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 7, 2022**

REFERENCES

[1] Liu, Z., Wang, S., Yao, L. et al. Online Multi-Object Tracking Under Moving Unmanned Aerial Vehicle Platform Based on Object Detection and Feature Extraction Network. J. Shanghai Jiaotong Univ. (Sci.) (2022). https://doi.org/10.1007/s12204-022-2540-4

[2] H. Song, J. Zhu, and Y. Jiang, "Two-stage merging network for describing traffic scenes in intelligent vehicle driving system," IEEE Transactions on Intelligent Transportation Systems, no. 99, pp. 1–12, 2021.

[3] H. (omas and H. Michael, "Camera-based method for distance determination in a stationary vehicle, corresponding computer program and corresponding parking assistance system with camera," 2020, https://patents.google.com/ patent/WO2012076400A1/en.

[4] Y. Chen, W. Li, C. Sakaridis, and D. Dai, "Domain adaptive faster R-CNN for object detection in the wild," 2018 IEEE, https://arxiv.org/abs/1803.03243.

[5] I. Papakis, A. Sarkar, A. Svetovidov, and J. S. Hickman, "Convolutional neural network-based In-vehicle occupant detection and classification method using second strategic highway research program cabin images," Transportation Research Record Journal of the Transportation Research Board, 2021.

[6] G. Goedert, D. Jungen, J. Beck et al., Weight -responsive Vehicle Seat Occupant Detection and Classification Method and System, US20180244172A1[P], 2018.

[7] M. A. A. Al-qaness, A. A. Abbasi, H. Fan, R. A. S. H. Ibrahim, and A. Hawbani, "An improved YOLO-based road traffic monitoring system," Computing, vol. 103, no. 2, pp. 211–230, 2021.

[8] H. Yin, Bo Chen, C. Yi, and Y. D. Liu, "Overview of target detection and tracking based on vision," Acta Automatica Sinica, vol. 42, no. 10, pp. 1466–1489, 2016.

[9] M. Qi, Y. Pan, and Y. Zhang, "Preceding moving vehicle detection based on shadow of chassis," Journal of Electronic Measurement and Instrument, vol. 26, no. 1, pp. 54–59, 2012.

[10] T. Schamm, C. Von Carlowitz, and J. M. Zollner, "On-road vehicle detection during dusk and at night," in Proceedings of the Intelligent Vehicles Symposium, pp. 418–423, IEEE, La Jolla, CA, USA, 21 June 2010.

[11] J. Hu, Research on Fast Tracking Algorithm of Moving Target Based on Optical Flow Method, Xidian University, Xi'An, City, 2014.

[12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587, IEEE, Columbus, OH, USA, 23 June 2014.

[13] R. Girshick, "Fast R-CNN," in Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448, IEEE, Santiago, Chile, 7 December 2015.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R- CNN: towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2015.

[15] K. Han, H. Zhang, Y. Wang et al., "A vehicle detection algorithm based on faster R-CNN," Journal of Southwest University of Science and Technology, vol. 32, no. 4, pp. 65–70, 2017.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 7, 2022

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN:towards real-time object detection with region proposal networks," in Proceedings of the Advances in neural information processing systems, pp. 91–99, Montreal, Quebec, Canada, 7 December 2015.

[17]. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 Jun–1 July 2016; pp. 779–788.

[18]. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.

[19]. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. arXiv 2018, arXiv:1804.02767.

[20]. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. arXiv 2020, arXiv:2004.10934.

[21]. Jocher, G.; Nishimura, K.; Mineeva, T.; Vilari no, R. Yolov5. 2020. Available online: https://github.com/ultralytics/yolov5 (accessed on 15 June 2021).

[22]. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

[23]. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.

[24]. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 2015, 28, 91–99.

[25]. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 21–37.

[26]. Duan, K.; Bai, S.; Xie, L.; Qi, H.; Huang, Q.; Tian, Q. Centernet: Keypoint triplets for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27–28 October 2019; pp. 6569–6578.

[27]. Wu, C.W.; Zhong, M.T.; Tsao, Y.; Yang, S.W.; Chen, Y.K.; Chien, S.Y. Track-clustering error evaluation for track-based multi-camera tracking system employing human re-identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 1–9.

[28]. Wende, F.; Cordes, F.; Steinke, T. On improving the performance of multi-threaded CUDA applications with concurrent kernel execution by kernel reordering. In Proceedings of the 2012 Symposium on Application Accelerators in High Performance Computing, Argonne, IL, USA, 10–11 July 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 74–83.

[29]. Meyer, D.; Denzler, J.; Niemann, H. Model based extraction of articulated objects in image sequences for gait analysis. In Proceedings of International Conference on Image Processing, Santa Barbara, CA, USA, 26–29 October 1997; IEEE: Piscataway, NJ, USA, 1997; Volume 3, pp. 78–81.

ISSN PRINT 2319 1775 Online 2320 7876
Volume 11, Iss 7, 2022

[30]. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988.

[31]. Henriques, J.F.; Caseiro, R.; Martins, P.; Batista, J. High-speed tracking with kernelized correlation filters. IEEE Trans. Pattern Anal. Mach. Intell. 2014, 37, 583–596.

[32]. Cui, Z.; Xiao, S.; Feng, J.; Yan, S. Recurrently target-attending tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1449–1458.

[33]. Hou, X.; Wang, Y.; Chau, L.P. Vehicle tracking using deep sort with low confidence track filtering. In Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, 18–21 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.

[34]. Liu, T.; Liu, Y. Deformable model-based vehicle tracking and recognition using 3-D constrained multiple-Kernels and Kalman filter. IEEE Access 2021, 9, 90346–90357. [CrossRef]

[35]. Wende, F.; Cordes, F.; Steinke, T. On improving the performance of multi-threaded CUDA applications with concurrent kernel execution by kernel reordering. In Proceedings of the 2012 Symposium on Application Accelerators in High Performance Computing, Argonne, IL, USA, 10–11 July 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 74–83.

[36]. Gomaa A, Abdelwahab MM, Abo-Zahhad M, Minematsu T, Taniguchi R (2019) Robust vehicle detection and counting algorithm employing a convolution neural network and optical flow. Sensors 19(20):4588