ISSN: 0974-5823 Vol. 7 No. 2 February, 2022

International Journal of Mechanical Engineering

Car Detection and Features Identification Based on YOLOV5

Asst. Prof. Dr. Ali Abdulazeez Mohammedbaqer Qazzaz/Education Faculty/ University Of Kufa/ Iraq

MSc. student Ahmed Younus Abdulkadhim /University of AL-kafeel, Najaf, Iraq

Abstract

Due to the advancement in technology at last years and increasing in the number of cars in the roads, it is necessary to develop techniques and algorithms for detection cars and performing some related tasks like identification the type of each car in the image . This research perform three important tasks, the first one related to discovering the existence of car in the specific image, the second task related to the classification process of cars into pre decided classes, and the third one related to discovering the direction of the car into (front, rear and side) for the purpose of treatment the presence of several non-accepted actions in the street. The proposed system use the network YOLO-v5 algorithm for performing the proposed tasks, where the data acquired by using suitable camera for cars in the streets of Iraq country, which were pre-classified into five classes (personal cars, taxi cars, minibus cars, trucks, and tuk-tuk). The number of acquired images is (7500) images. The training was conducted in the environment of https://colab.research.google.com/, where the results after training on 100 EPOCH were Odds Ratio mAP @ 0.924.

Keywords: You only look once, Object detection, Convolution NN, direction, minibus, images.

Introduction

The increasing in the number of cars in the streets from different types as well as the difficulties in the monitoring process for all these cars for detecting abnormal actions for the purpose of fining them, the great development in the cars number from different types leads to many problems or risks related to people, led to the need of developing special systems that work to reduce these risks by using technologies or algorithms to discover the presence of cars, determine the direction of cars and classify cars by exploiting special techniques like CNN, SSD and Yolo.

Object recognition is a critical component of automated driving & traffic monitoring [1]. While conventional vehicle detection methods such as the Gaussian mixed model (GMM) [2] have demonstrated promise, they are not optimal due to changes in illumination, background clutter, and occlusion. Vehicle detection remains a significant computer vision task.

Recent years have seen substantial advancements in object detection because to the resurgence of DNN [3]. Object detection is a well-known technique for locating items in an image or video. Object detection is mostly accomplished through the use of machine learning and deep learning approaches [4]. Humans recognize objects by examining a photo or video; researchers want the developing the machine to have the same capability through the use of intelligence algorithms. Object detection is a technique that is utilized in video surveillance systems as well as a variety of other applications. Several deep learning-based object recognition techniques include the Region convolutional neural network (R-CNN), the SSD, and YOLO in various flavors. [5]. Object detection process is a procedure of two-steps, first step concern to train the model on large number of object positions and types and statuses and second step, the learned model used to perform detection operation of desired object [3] also, system can be used previous or already trained model like YOLO and SSD. The choice of deep learning or machine learning as a training approach is determined by the model's objective. The algorithm used in this research is YOLO As illustrated in the following figure (1)

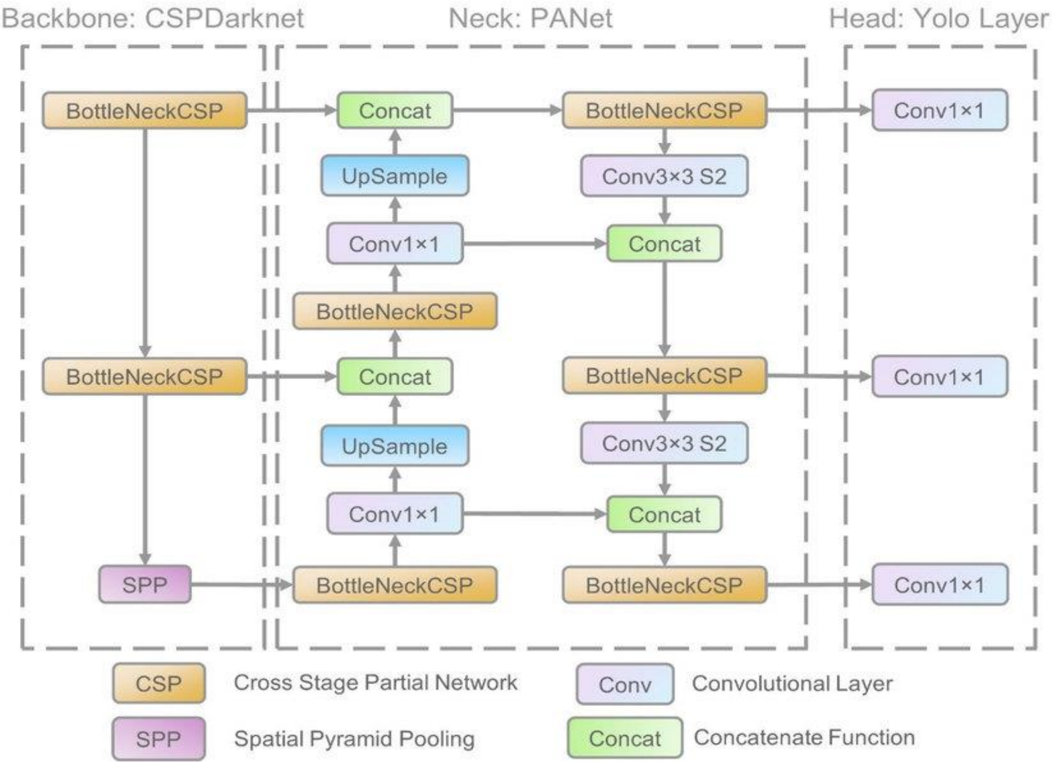


Figure (1) Yolov5

The architecture of Yolov5 network consists of three parts:
(1) Backbone: CSPDarknet,
(2) Neck: PANet, and
(3) Head: Yolo Layer.
The data are initially input to the first parts(CSPDarknet) for the purpose of extracting features, and then to (PANet) for the purpose of performing fusion for the extracted features, Finally, results fed to the Yolo Layer for the purpose of introduce the output results of detection process like class, score, location and size.[6]

Literature review

1- V Sowmya and R Radha in [7] proposed real time car detection technique by using yolo-4, Transfer learning methods are used to avoid over-fitting and to optimize training pace. Following that, a tuning YOLOv4 method will be used used to recognize the heavy truck. The Vision by Computer science is used to validate the algorithm in real-world scenarios with varying traffic density. The suggested system provides an accuracy of 96.54 percent (mAP) for detection, as demonstrated by experimental findings. 3500 photos of buses and trucks comprise the custom vehicle dataset. Vehicle photos are downloaded from crawlers. Several of the photographs were captured from live traffic video while others were downloaded from the net.

2- Quanfu Fab et.al. in [8] undertake a complete investigation of the suggested model's underlying structure by conducting a variety of experiments. They demonstrate that by appropriately tweaking parameters and modifying the algorithm, the system can considerably increase the accuracy of vehicle detection in Faster R-CNN and reach accepted results on the KITTI dataset for vehicle. These papers provide useful information for other researchers interested in applying Faster R-CNN to the specific issues and their datasets. The system divided pictures in the KITTI dataset's which is (7481) images into two parts, 2/3 for training and 1/3 for testing, resulting in (11042) accepted training samples and 3105 accepted testing samples.

3- Bilel Benjdira et.al. in [9] The purpose of this paper is to compare the performance of two cutting-edge CNN methods that are Faster R-CNN and YOLOv3 for the purpose of recognition car depending on aerial photos. These two models were trained and validated using a huge car dataset collected via unmanned aerial vehicles. The researchers demonstrated in this work that YOLO-v3 better than Faster R-CNN in the sensitivity and spend time for processing, despite the fact that their precision metrics are comparable. To conduct the experimental portion of this study, the system generated a collection of photos divided into a training and test set. There are (218) photos and (3365) occurrences of labeled automobiles in the training set. (52) photos and (737) occurrences of automobiles are included in the test set. We used the Tensor flow Object Detection API for training the Faster R-CNN.

4- Sushmitha.S et.al. in [10] designed system with the goal of reducing highway congestion, improving street design, identifying people at crime scenes, reducing road accidents, detecting car theft, and creating a living zone, among other things. In the current environmental conditions, deciding the outline of an unfixed car in a video frames is quite challenging. To address this issue, the research presented a number of strategies, including video segmentation , scene preprocessing and car tracking. The system provides an overview of the detection, recognition, and tracking of several cars. Following frame conversion, the operation of preprocessing is performed by applying a median filter, which takes the frames in video and converts them to gray scale images. As the car moves, it is tracked by looking it in a succession of images in which the mark car appears in the top of the box, indicating that a wanted object has been identified and recognized, until the car has entirely moved.

5- Zhi Xu et.al. in [11] proposed The creation of a deep neural network model is centered on the detection of tiny targets using a UAV platform. Due to the innovative structure and broad industrial application possibilities of one-stage detection models such as YOLO, this article explains a new model for detection objects based on the YOLO-v2 algorithm. Faced with the challenge of missed identification of small targets, a set of enhanced programs is proposed that are perfect for detecting far vehicles from an aerial images and capable of detecting cars in real-time environment, including an optimal pooling technique and dense topology. Because a YOLO model has such a wide potential application and the capacity to detect in real time, it has a lot of room for development in terms of detection accuracy. Predicting Networks The network will provide a block of matching features in accordance with its initial grid parameters. The network forecasts the anchor box's offsets, which are the bounding box's four offsets. x_t , y_t , w_t , and h_t.

6- DANIEL PADIlIA CARRaSCO et.al. in [12] offer a redesigned deep model for detecting desired objects depending on the YOLO-v5 network. The introduced technique capable of detecting objects of various sizes, from large to small and in some cases to extremely small. To be more precise, they suggest the use of a multi-scale models for learning deep wanted feature by considering representation of that feature at different sizes and choosing the scales produced best suitable cases for object detection. By contrast to the architecture of original YOLO-v5 model, the suggested model minimizes the number of parameters used in training operation. Additionally, the testing results show a significant increase in precision. As demonstrated in the studies, proposed model resulted in a modest reduction of the profiles of YOLO-v5-S from (7.28) million parameters to (7.26) million parameters. Comparing to the profiles of YOLO-v5-L/X, the system reduced the speed of detecting objects by (30 fps). When compared to the profile of YOLO-v5-X, the efficiency of detecting micro vehicles was greatly enhanced by 33%.

7- Yan Miao et. al. in [13] designed an effective nighttime vehicle detection approach Firstly, the original nighttime images enhanced by an optimal MSR algorithm. Then, using the improved images, a pre-trained YOLO-v3 network was selected and fine-tuned. Finally, the detection network was utilized to recognize vehicles in nighttime photos, outperforming two commonly used approaches. namely the Faster R-CNN and SSD, on the precision and detection efficiency. The average precision of the proposed method reaches 93.66%, which is 6.14% and 3.21% higher than that of the Faster R-CNN and SSD, respectively. The network was used to detect vehicles from the nighttime images in the testing dataset by outputting locations and classifications of the detected objects. Prediction Every box calculates the probable categories in the bounding box through multi-label classification.

The proposed system

The proposed system can be illusterated as in figure (2)

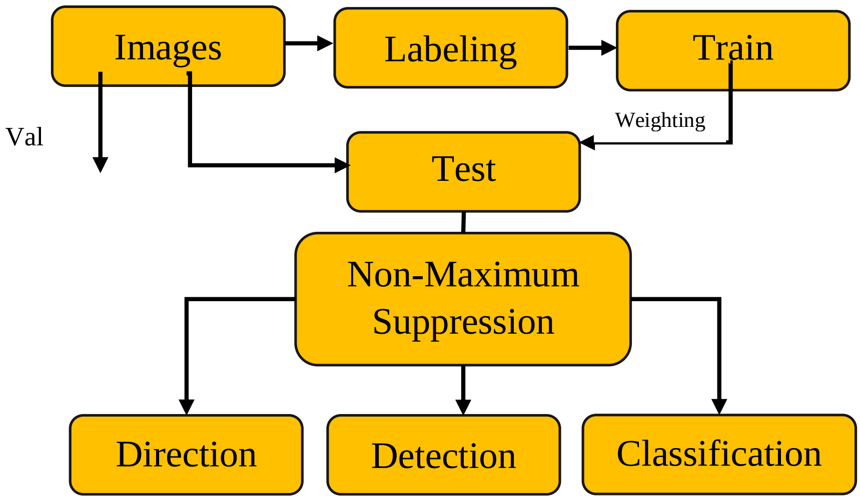


Figure (2) the block diagram of the proposed

A. Experimental Data

The proposed system used data set captured from the Iraqi street, it was filmed in more than one city by using the AF_P Dx NIKKOR 18_55mm f/3.5-5.6G VR- Nikon MEA then labelImg [8] will be used in labeling process for the acquired pictures. The dataset images will be divided into five essential classes of cars (personal cars, taxi cars, Minibus cars, truck, and tuk-tuk). The resolution of the images in the datasets is (1280×720) pixels. Some of these classes will be divided into subclasses that are (front car, rare car, side car) as illustrated in figure (3).

B. Experimental Environment

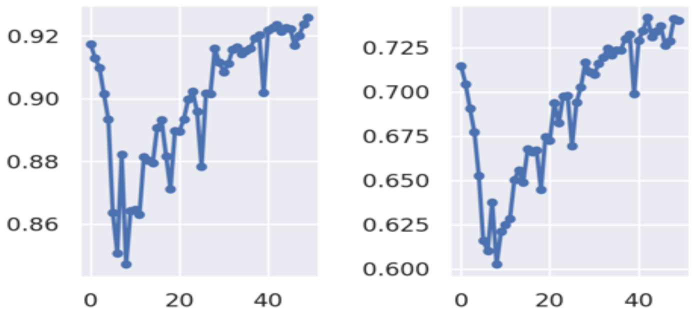
The processor is Intel(R) Core(TM) i7-7200U CPU @ 2.50GHz (4CPU), memory 8G. The operating system is Windows 10 Pro, 64-bit. The depth learning framework is https://colab.research.google.com/.

C. Experimental Results and Analysis

Set the network epochs to (100), batch-size to (32), and img-size to [416, 416], and compare different parameters to obtain the best training results.

Figure (3) classes and sub classes in the proposed system

The training process after 100 iterations with Yolo v5x, the values of mAP@ 0.5 recorded during the training process are shown in figure (4).



a-mAP@0.5 b- mAP@0.5:0.95

Figure (4) training process

The data sets contain images from three directions (front, rare, and from the side) as well as taking video clips through a Canon-style camera... The process of preparing datasets took (10) days after that, the images will be processed, then the process of tagging the database of (7500) images was carried out through the website https://roboflow.com/, specializing in the process of tagging as will as specializing in the training process, where the data divided into (70%) images for training, (20%) images for testing and (10%) image validation, data was trained on the yolov5x algorithm on epoch (10) on (7500) images with a training time of (18min 41 s) , then on epoch (25) during a training time of (58min 40 s), then on epoch (50) with a training time of (1hour 52min 36 s), and finally on epoch 100 training time was (3 hour 45min 12 s) as illustrated in table (1).

Table (1) training time vs number of epoch

epoch	Training-time	mAP@	data
10	18min 41s	0.622	7500
25	58min 40s	0.867	7500
50	h1 52min 36s	0.91	7500
100	h3 45min 12s	0.924	7500

There are some challenges in the proposed database that reduce the percentage of probability, including suggesting more than one class, some of proposed classes are general, so system need very large number of images to get better accuracy, especially classes of personal cars, as well as other challenges, including the direction of the car's shape, whenever the direction is difficult or unclear, reducing accuracy. Data was firstly preprocessed for discovering the car Secondly, the direction of the car was determined as one of three directions (front, rare and side), and then the cars were classified into five subclasses (personal cars, taxi cars, Minibus cars, truck, and tuk-tuk) as illustrated in figure (5).

Figure (5) results of the proposed system

Discussion

After studying results of the proposed system as illustrated in Table (2) figure (6), we can calculate some metrics for calculating the efficiency of the proposed system as follows

a- True positive

b- False Negative

c- False Negative

d- $precision = \frac{TP}{TP+FN}$

e- $recall = \frac{TP}{TP+FN}$

f- $F_1 = 2 * \frac{recall*precision}{recall+precision}$

Table (2) efficiency metrics of the proposed system

Classes	TP	FP	FN	F1	recall	precision
CAR-F	0.92	0.13	0.08	90	92.000	87.619
CAR-R	0.9	0.9	0.09	91	90.909	90.909
CAR-S	0.9	0.1	0.1	90	90.000	90.000
TAXI-F	0.98	0.08	0	96	100.000	92.453
TAXI-R	0.97	0.07	0.01	96	98.980	93.269
TAXI-S	0.93	0.06	0.02	96	97.895	93.939
MINIBUS-F	0.93	0.11	0.06	92	93.939	89.423
MINIBUS-R	0.94	0.06	0.05	94	94.949	94.000
MINIBUS-S	0.9	0.06	0.07	93	92.784	93.750
TRUCK	0.93	0.11	0.03	93	96.875	89.423
TUK-TUK	0.93	0.13	0.06	91	93.939	87.736

The accuracy of detecting each suggested classes and subclasses will be explained in the figure (6)

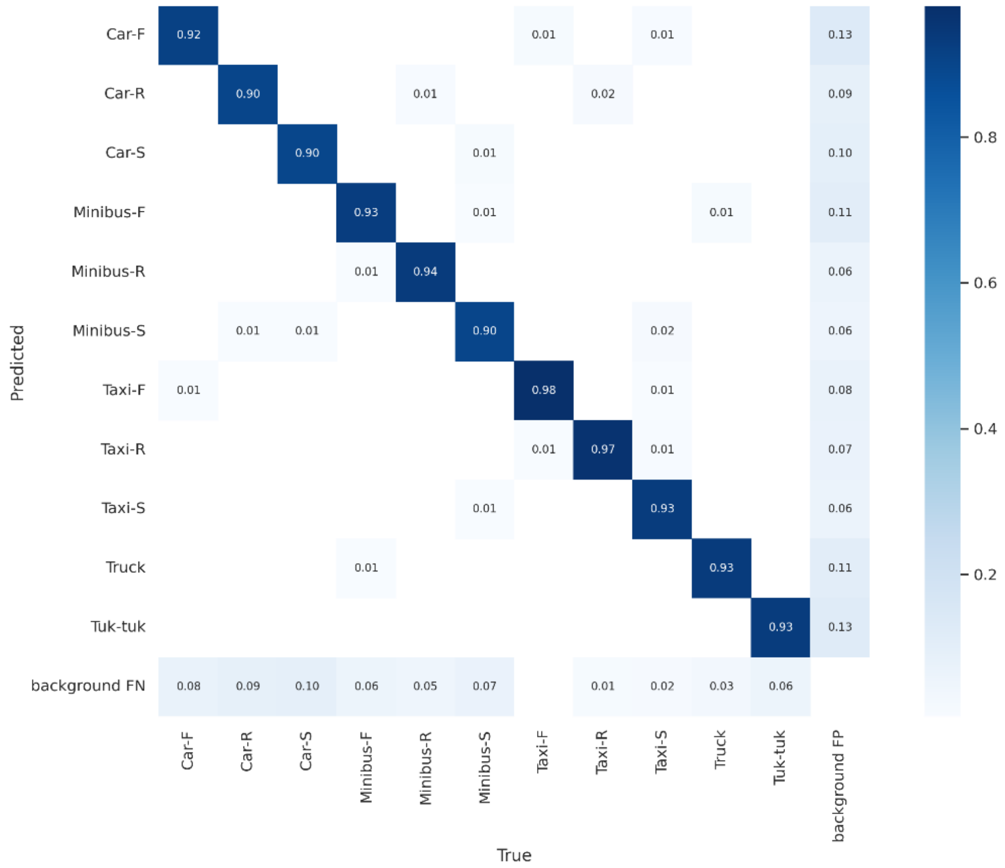


Figure (6) accuracy of predicted all classes and subclasses

Figure (7) illustrated the efficiency metrics of the proposed system

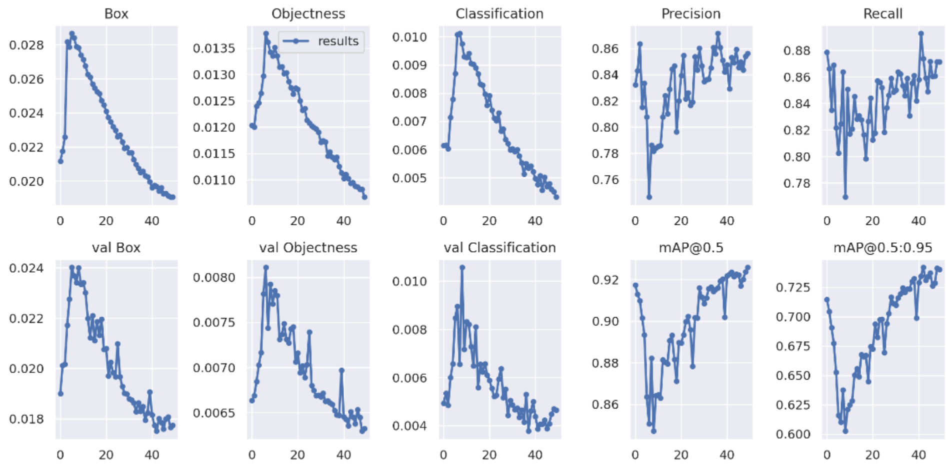
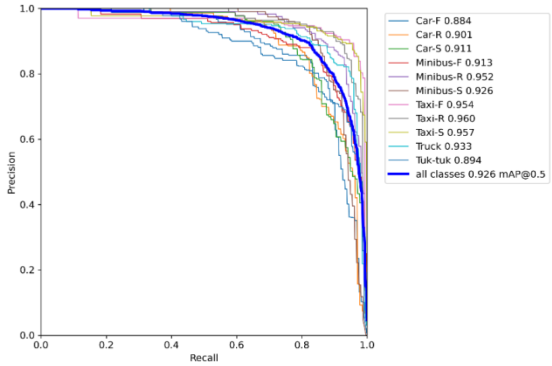
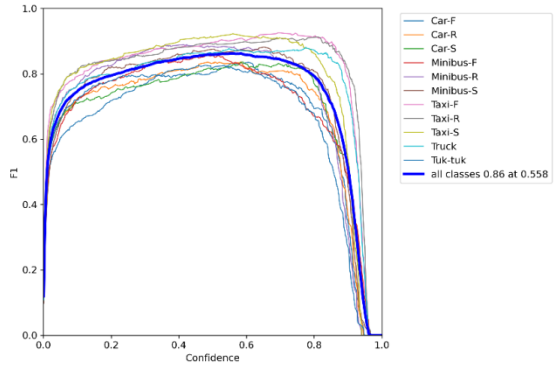


Figure (7) efficiency charts of the proposed system

Figure (8) illustrated the relation between produced efficiency metrics





a-(recall/precision) chart b-(confidence/f1) chart

Figure (7) relations between produced metrics

Figure (8) illustrated the produced efficiency

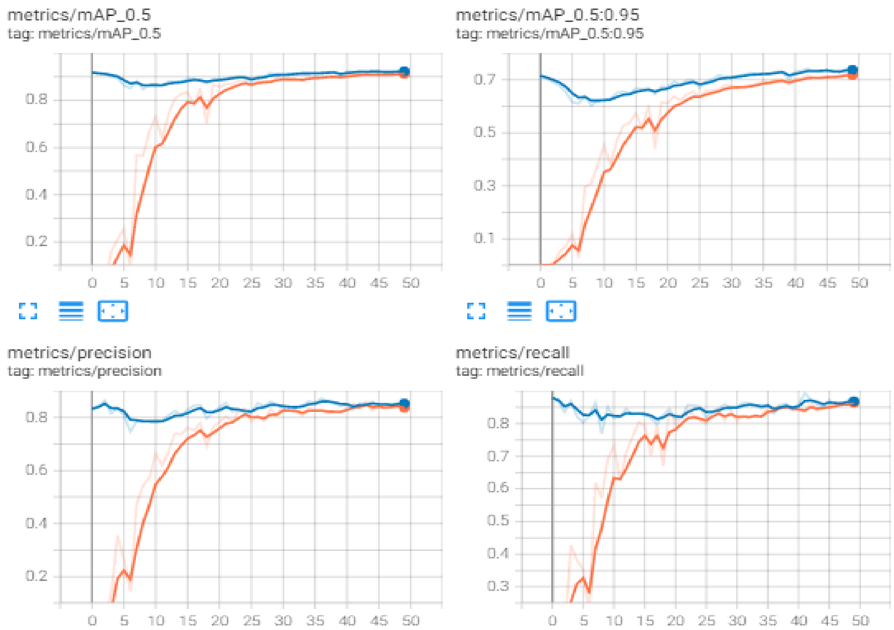
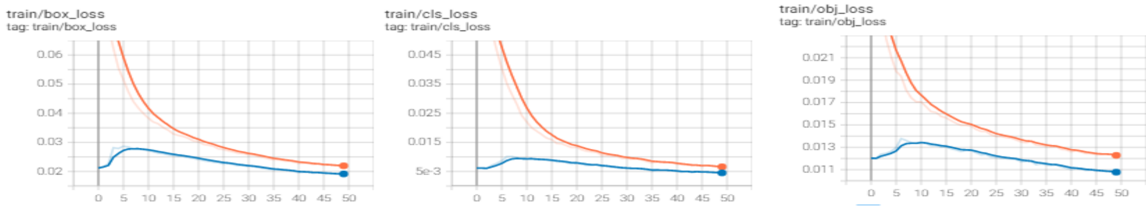


Figure (8) efficiency charts

Figure (9) illustrated the produced loss



a-box loss b-classification loss c-objectness loss

figure (9) Loss charts

Conclusion

The purpose of determining the direction of the car has several benefits, some of these benefits related to people with special needs, including determining the direction of travel and knowing the cars in the wrong side through the direction of the car as well as the determining the direction of the car can be useful for blind-human to knowing the direction of traffic. The result of the proposed system introduce that the using of the Yolo network can be a useful way for detecting objects with perfect accuracy and probability that well be reflect reliable results that can be depended by real time systems that need high degree of accuracy and spend very small period of time accepted as a good response without taking time produce problems in reacting.

References

1. Hu, X., et al., *SINet: A scale-insensitive convolutional neural network for fast vehicle detection.* IEEE transactions on intelligent transportation systems, 2018. **20**(3): p. 1010-1019.
2. Stauffer, C. and W.E.L. Grimson. *Adaptive background mixture models for real-time tracking*. in *Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149)*. 1999. IEEE.
3. Hinton, G.E., S. Osindero, and Y.-W. Teh, *A fast learning algorithm for deep belief nets.* Neural computation, 2006. **18**(7): p. 1527-1554.
4. Fang, L., X. Zhao, and S. Zhang, *Small-objectness sensitive detection based on shifted single shot detector.* Multimedia Tools and Applications, 2019. **78**(10): p. 13227-13245.
5. Fang, W., L. Wang, and P. Ren, *Tinier-YOLO: A real-time object detection method for constrained environments.* IEEE Access, 2019. **8**: p. 1935-1944.
6. Xu, R., et al., *A Forest Fire Detection System Based on Ensemble Learning.* Forests, 2021. **12**(2): p. 217.
7. Sowmya, V. and R. Radha. *Heavy-Vehicle Detection Based on YOLOv4 featuring Data Augmentation and Transfer-Learning Techniques*. in *Journal of Physics: Conference Series*. 2021. IOP Publishing.
8. Fan, Q., L. Brown, and J. Smith. *A closer look at Faster R-CNN for vehicle detection*. in *2016 IEEE intelligent vehicles symposium (IV)*. 2016. IEEE.
9. Benjdira, B., et al. *Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3*. in *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*. 2019. IEEE.
10. Sushmitha, S., N. Satheesh, and V. Kanchana. *Multiple car detection, recognition and tracking in traffic*. in *2020 International Conference for Emerging Technology (INCET)*. 2020. IEEE.
11. Xu, Z., et al. *Vehicle detection under uav based on optimal dense yolo method*. in *2018 5th International Conference on Systems and Informatics (ICSAI)*. 2018. IEEE.
12. Carrasco, D.P., et al., *T-YOLO: Tiny vehicle detection based on YOLO and multi-scale convolutional neural networks.* IEEE Access, 2021.
13. Miao, Y ,.et al. *A Nighttime Vehicle Detection Method Based on YOLO v3*. in *2020 Chinese Automation Congress (CAC)*. 2020. IEEE.