



Automatic Traffic Red-Light Violation Detection Using AI

Le Quang Thao^{1,2*}, Duong Duc Cuong², Nguyen Tuan Anh³, Pham Mai Anh⁴, Ha Minh Duc⁵, Nguyen Minh⁶

¹ Faculty of Physics, VNU University of Science, Hanoi 100000, Vietnam

² University of Science, Vietnam National University, Hanoi 100000, Vietnam

³ Nguyen Sieu High School, Hanoi 100000, Vietnam

⁴ Le Quy Don High School, Hadong 100000, Vietnam

⁵ VNU HUS High School for the Gifted, Hanoi 100000, Vietnam

⁶ Chu Van An High School, Hanoi 100000, Vietnam

Corresponding Author Email: thaolq@hus.edu.vn

<https://doi.org/10.18280/isi.270109>

ABSTRACT

Received: 10 December 2021

Accepted: 12 January 2022

Keywords:

traffic, red light violation, machine learning, convolutional neural network

Our research is the design of a traffic signal violation detection system using machine learning that learns to prevent the increasing number of road accidents. The system is optimized in terms of accuracy by using the region of interest and location of the vehicle with a red-signal state. By modifying some parameters in the YOLOV5s and re-training the COCO dataset, we can create a model which can be predicted with a high accuracy of 82% for vehicle identification, 90% for traffic signal status change and up to 86% for violation detection. This can be used for red light violation detection which will help the traffic police on traffic management.

1. INTRODUCTION

The number of vehicles on the road is increasing rapidly in recent years. This leads to severe destructions, accidents that endanger people's lives and properties because many people don't obey the rules when using transport [1]. Traffic laws have always been one rule to break, due to manual check-up and clarification of whether a vehicle has broken the law or not [2]. Especially during rush hour, abundant vehicles are driving all over the place, making the officers' jobs much more devastating, like finding a needle in a haystack.

To solve this problem and prevent unfortunate consequences, multiple traffic violation detection has been invented and created to detect whether a car was speeding up, signal change, crossing pavements, and generally not obeying laws. These detections were useful, innovative, and have helped a lot of people. However, they were costly, and can only function in one direction, and these can be very ineffective in intersections, where most vehicles can be seen.

Li [3] On the large scale there is a problem faced by domestic traffic signal controllers. It is observed that to handle this problem there is need to develop a traffic monitoring system so this comes to the conclusion that the traffic controllers should be based on web technology. Traffic Monitoring system is an important topic to be researched in these recent years and the coming years [4]. This paper provides real-time traffic monitoring and vehicle tracking for the private and public transportation sector.

The detection process becomes more and more difficult as more vehicles appear because the scanner will not scan every vehicle, hence making the detection system almost useless. This proves that the recent detection systems and machines are not considered valuable, reliable, and efficient. Wu [5] This paper presents the recognition of the road interactions for traffic monitoring systems using traffic light recognition

methods. The YUV space method is used to check the intersection of the vehicle at the roadside. If any rules are broken it is recorded in video and sent to the E-police. The test video is recorded by the recorder of the E-Police system. Thus, the result of this experiment is better than previous research.

Rafael et al. [6] propose a red-light running detection system that analyzes the video captured by the camera inside the vehicle for the purpose of educating the driver's consciousness. The disadvantage is that it is only local, the situation of the observed vehicle must be at the same time as the offending vehicle at the intersection. Satadal et al. [7] have designed and developed a complete system to automatically generate the list of all vehicle images violating the stop line and red lights from video snapshots of roadside surveillance cameras using image processing, which takes longer with low accuracy.

A new way of detection must be made, a way that can function all of the time, detects rapidly with high accuracy, and is efficient. This can be done using multiple AI processing techniques, such as deep learning, CNN, etc. [8]. This is considered, by the general public, a good idea, since it will reduce the amount of workforces required on the streets, and the personnel can focus on other productive work.

In our study, we have applied those methods into our project of detecting vehicles that have passed red lights. We use a camera scanner on the side of the road, looking at the intersection and the traffic lights. Then, we'll be using CNN to identify the vehicles, meanwhile we'll use image processing to figure out whether it's a red light or not. If it is, then we'll record the clip of the vehicle and report, send it to the personnel so they can initialize punishments. Apart from this, we'll also be using CNN and image processing to find out the vehicle's plate number, for the personnel to find it easier. With many deep neural networks for the outdoor advertisement panel detection problem by handling multiple and combined

variabilities in the scenes [9]. We used YOLO which produced better panel localization results, detecting a higher number of true positive panels with a higher accuracy for our purpose combined with the optical character recognition technique to recognize the extracted characters and get information on violation license plates.

2. MATERIALS & METHODS

2.1 Schematic of the system

Our invention mainly consists of three parts, vehicle violation detection, red signal change monitoring and vehicle recognition with evidence recorded shown as in Figure 1. Traffic at intersections with signal lights will be captured by the camera and pre-processed to detect the stop-line and the location of the traffic lights. By using image processing techniques such as filtering, calculating the area of uniform areas, contour drawing to create awareness of the road marking as well as determining the location and changing state of the traffic lights. These parameters are used as reference to decide whether the vehicle has committed a traffic violation.

The identification of violations is performed on the server. The server will observe the stop-line on the road and the traffic signal color to generate different zones. If the vehicle is inside the violating zone while the traffic light is yellow or red, it will be detected as violating.

We used deep learning techniques with pre-trained models to recognize the vehicle that the camera captured, and in our project, we used the modified YOLOv5s pre-trained model for the detection of violating vehicles. Once the vehicle violates, the following frames will be used to identify the vehicle's license plate.

The image is given to the camera as an input, where it is further processed to compare with traffic signal change and check for violation. While the vehicle runs the red-light, a video will be recorded as evidence shown as in Figure 2, meanwhile the deep learning results will be implemented and used to extract the violated vehicle's information with high precision and also gives the vehicle type and its license plate.

To detect a vehicle license plate, we applied machine learning to border the plate combined with the optical character recognition technique, which is used to recognize the extracted characters to get the violated vehicle's information from the plate that is bordered.

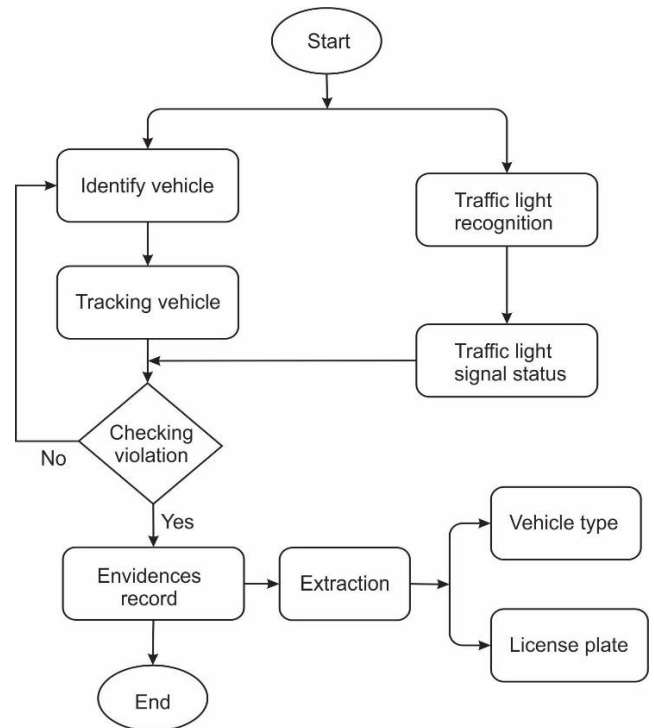


Figure 2. Algorithm of violation detection

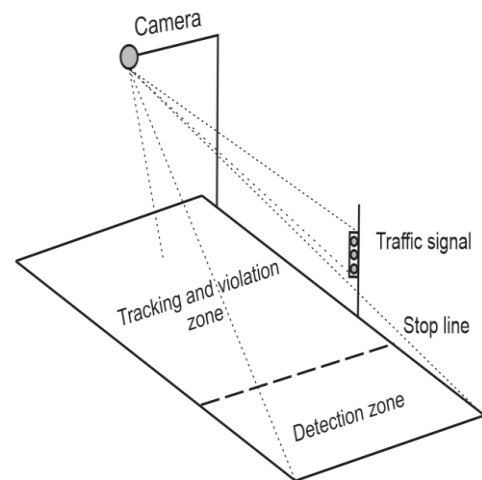


Figure 3. System configuration

Vehicles are detected using various YOLOv5s pre-trained models on the COCO dataset [10] with three different output classes: Cars, buses and motorcycles. For vehicle tracking, we will perform detection every 5 frames, using the intersection-over-union measurement of the bounding boxes after detection. If a detected box matches a tracking box with a relatively good percentage, then it's the same box that is in the tracking list. After that, all violation cases are checked. A traffic line is drawn over the road in the preview of the given video footage by the user.

The line specifies that the traffic light is red. Violation happens if any vehicle crosses the traffic line in the red state shown as in Figure 3. Stop lines are usually made by the authorities and are usually fixed on each intersection, using image processing techniques before performing identification, this is done only once and can draw virtual lines in the frame captured by the camera.

For fixed camera position at the intersection, the recognition of the stop line and the traffic light status is done through



Figure 1. Schematic of system

conventional image processing techniques with Red-Green-Blue color thresholds in the identification boxes. To identify vehicles violating traffic, we determine the area containing license plates using YOLOV5s [11], then use a segmentary algorithm to separate each letter on the license plate and build a CNN model to classify them. Characters depending on the license plate format of each vehicle, extract information for the detection process.

The stop lines on the road as shown in Figure 4 are most affected by weather conditions and brightness, however, we only need to locate once for each position at the intersection and draw the boundary lines in the frame captured from the camera.



Figure 4. Stop lines on the road

On the other hand, the status of the green - red - yellow signal lights shown in Figure 5 is usually located at a position away from the roadbed, so it is more contrasting and we can easily identify it with image processing using color filters techniques.



Figure 5. Traffic lights location

2.2 Preparation and training of data

We had prepared the training data to extract two characteristics, one is the means of transport, including cars, buses and motorcycles, the other is identifying license plates. To obtain data for the training of vehicles, we had used available images with annotation from the COCO dataset, resulting in 16,115 photos consisting of three types of cars, buses and motorcycles. Then they are re-trained according to various modified YOLOv5 models. To identify license plates, we had collected vehicle data from public parking in Hanoi, Vietnam. By using various filters, including flipping, blurring and changing light contrasts, were applied additional training materials are created for the machine so that it learns to detect the same object under different lighting conditions as shown in Figure 6.

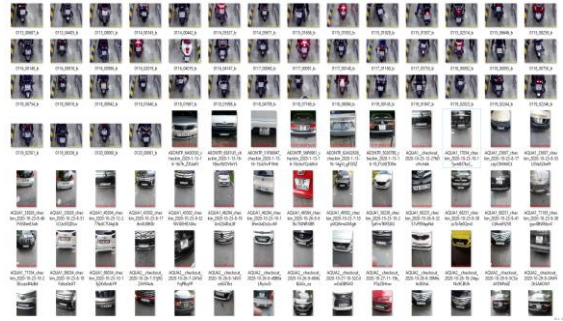


Figure 6. License plate dataset

Then, we labeled the identified objects with the labeling program as shown in Figure 7.

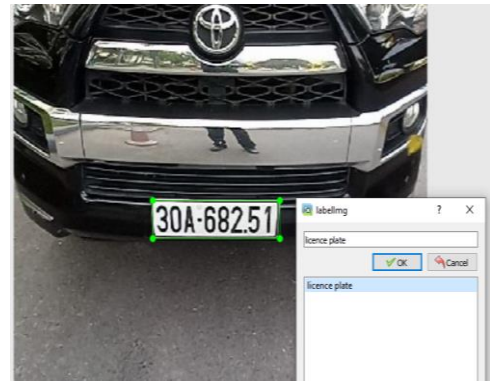


Figure 7. Labelling image process

The tool we had used is LabelImg [12], LabelImg supports labeling on both PASCAL VOC and YOLO formats with extended annotation files are .xml and .txt respectively. In case of using the YOLO architecture, we had saved the annotation file as .txt. Each line in an annotation file consists of: `<object-class> <x> <y> <width> <height>`. The data were then ready to be trained on the neural network. We use 2 independent training processes for vehicle recognition in the dataset as mentioned in data preparation and license plate recognition.

2.3 Improved CNN model

We proposed the YOLOV5s model and improved this model. With YOLOV5s origin model, there is a series of various kinds of layers with 7 million parameters in each layer that are one or a set of convolutions. We can either multiply the weights of matrix, modify the size of the matrix, or filter, then interfere with each layer's parameters to get the models v1, v2, v3 and v4 shown as in Figure 8, Figure 9, Figure 10 and Figure 11.

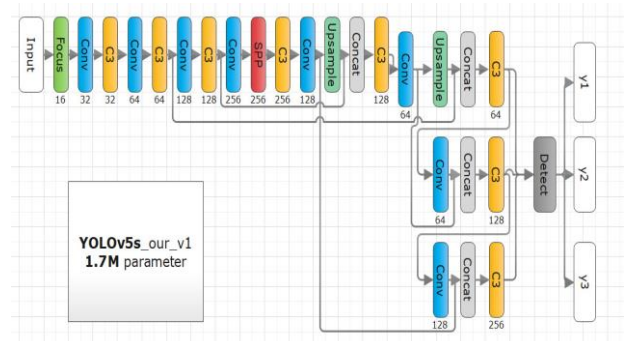


Figure 8. YOLOV5S_v1 reduce 1/2 filter in Conv

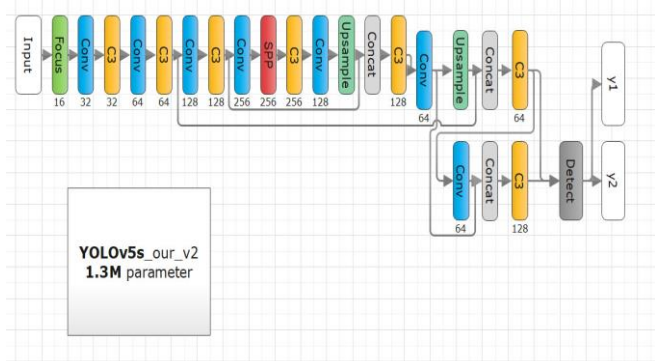


Figure 9. YOLOV5S_v2 remove y3 block

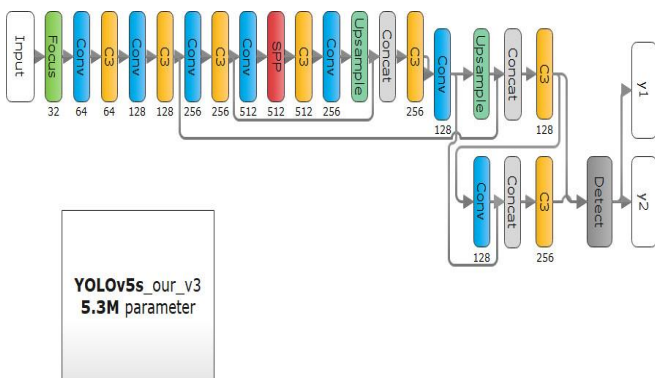


Figure 10. YOLOV5S_v3 reduce 1/2 filter in Conv, remove y3 block

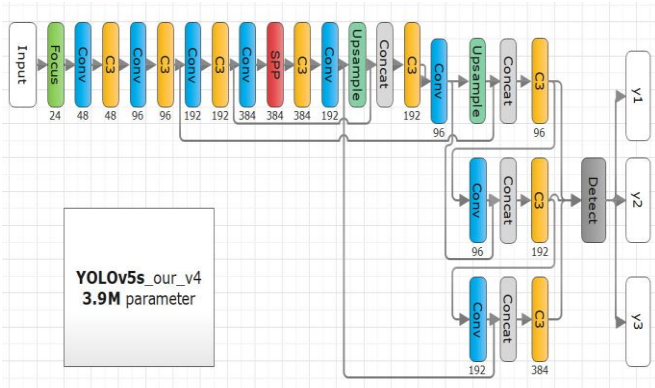


Figure 11. YOLOV5S_v4 reduce 3/4 filter in Conv

3. RESULT

3.1 Model performance

With the objects reduced quite a lot compared to the Coco dataset, if we re-train according to the origin YOLOv5s model, it will consume resources and the processing will be slower, so we change the parameters at the convolution layers, compare the accuracy with the original model we proposed and compare 4 modified yolo models by changing the filter at each convolution layer or removing the results $y1$, $y2$, $y3$ respectively at the top output of the detection layer to evaluate the performance of each model in our problem, parameter and prediction time shown as in Table 1. This training was done for 200 epochs by using GPU Tesla T6 with Pytorch environment [13], which we borrowed from Google Colab.

Table 1. Results in various modified YOLOV5S models

Model name	Modified	Parameter	Inference	mAP@.5
YOLOV5S origin	None	7.0M	0.21 sec	0.677
YOLOV5S_v4	Reduce 3/4 filter in Conv	3.9M	0.13 sec	0.657
YOLOV5S_v3	Remove y3 block	5.3M	0.19 sec	0.657
YOLOV5S_v2	Reduce 1/2 filter in Conv, remove y3 block	1.3M	0.06 sec	0.597
YOLOV5S_v1	Reduce 1/2 filter in Conv	1.7M	0.07 sec	0.628

The results in Table 1 show that the original model has the highest overall mAP value. However, this model is used to identify a lot of labels that require a very large inference time. Our case only has three labels at most, so it's not necessary to use this large model. Although the $v2$ and $v1$ model have a fairly fast prediction speed shown as in Figure 12. The mAP values are much lower than those of $v3$ and $v4$, so we use $v4$ for both vehicle recognition and license plate recognition, since $v4$ has both the high mAP value and the best possible prediction time.

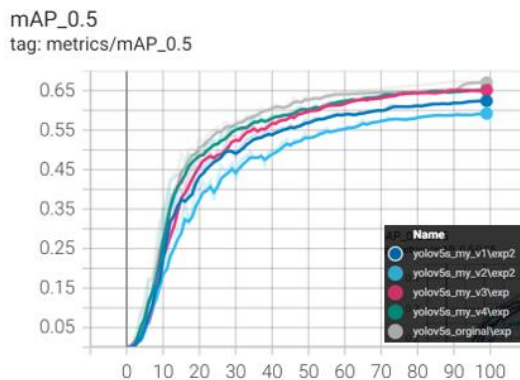


Figure 12. Accuracy graph

The object-loss graph illustrates the error when distinguishing the object from the background trust. As we can see, model $v4$ has the least error witting the best able to identify the object. Model $v4$ has fewer parameters but the approximate error of the original model shown as in Figure 13.

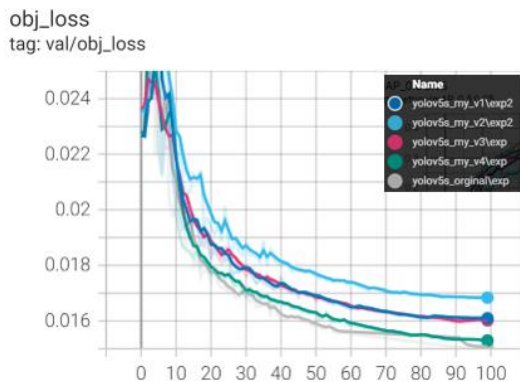


Figure 13. Object loss graph

3.2 Detection and classification

In our report, by improving the pre-trained models, we achieved better results than the feasible models in terms of *mAP* and training time. According to the comparison illustrations of those parameters after training on the same data set, we realized and selected the V4 model, then we applied extracted results after training to a video clip of a scene with traffic violations which are in real-life footage under random conditions.

The expected results were obtained as shown in Figure 14. In terms of detecting red-light signals up to 90%, classifying vehicles as cars or motorbikes approximately 86%, combined with the optical character recognition technique to recognize the extracted characters and get information on license plates up to 82%.

The trail on vehicle motion estimation shows that the method we used in this paper meets our requirements.



a. Origin frame in video



b. Detect and classification



c. Plate extraction information

Figure 14. a. Origin frame in video; b. Detect and classification; c. Plate extraction information

4. CONCLUSIONS

This paper proposes the detection of red-light traffic violation in real time using modification YOLOv5s. This neural network used pre-trained weights on the COCO dataset and then was trained on a separate dataset with various parameter changes from the original model. Results show that the detection of multiple traffic violations from a single input source is achievable. By modifying some parameters, we can create a model which can be predicted with a high accuracy of 82% for vehicle identification, 90% for traffic signal status change and up to 86% for violation detection. This can be used for red light violation detection which will help the traffic police on traffic management.

5. FUTURE WORK

With the increase in traffic density all over the world, it is a great challenge for traffic management. It should be emphasized that the large area covered and the large volume of traffic monitoring and detection from a single input source use parallel computation at the traffic control station at the intersection.

Further improvements are needed to reduce computation time on high volumes of traffic as well as to enhance the input data when training with images under various conditions, especially with vehicle photos in the night or in bad light conditions.

In the future, this idea can be enhanced to work in real-time CCTV cameras to handle all kinds of license plates.

REFERENCES

- [1] Wang, X.L., Meng, L.M., Zhang, B.B., Lu, J.J., Du, K.L. (2013). A video-based traffic violation detection system. Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), pp. 1191-1194. <https://doi.org/10.1109/MEC.2013.6885246>
- [2] The Police Department Guide to Safe Driving, www.defensivedriving.org/, accessed on June 2021.
- [3] Li, Y.B., Li, Z.H. (2011). Design and development of the monitoring system for traffic signal controller based on WEB technology. 2011 International Conference on Multimedia Technology, pp. 3920-3924. <https://doi.org/10.1109/ICMT.2011.6002741>
- [4] Azer, M.A., Elshafee, A. (2018). A real-time social network- based traffic monitoring & vehicle tracking system. 2018 13th International Conference on Computer Engineering and Systems (ICCES), pp. 163-168. <https://doi.org/10.1109/ICCES.2018.8639287>
- [5] Wu, N., Fang, H.X. (2017). A novel traffic light recognition method for traffic monitoring systems. 2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), pp. 141-145. <https://doi.org/10.1109/ACIRS.2017.7986081>
- [6] Brasil, R.H., Machado, A.M.C. (2017). Automatic detection of red light running using vehicular cameras. in IEEE Latin America Transactions, 15(1): 81-86. <https://doi.org/10.1109/TLA.2017.7827891>
- [7] Prasantha, H.S. (2020). Traffic red light violation detection using image processing. International Journal

- of Advances in Engineering and Management (IAEM), 2(5): 440-445.
- [8] AI in Transportation. www.emerj.com, accessed on 6 June 2021.
- [9] Morera, A., Sánchez, A., Moreno, A.B., Sappa, A.D., Vélez, J.F. (2020). SSD vs. YOLO for detection of outdoor urban advertising panels under multiple variabilities. *Sensors*, 20(16): 4587. <https://doi.org/10.3390/s20164587>
- [10] Common objects in context. www.cocodataset.org, accessed on 6 June 2021.
- [11] YOLO: Real-Time Object Detection, www.pjreddie.com/darknet/yolo, accessed on June 2021.
- [12] LabelImg, an annotation tool. <https://github.com/tzutalin/labelImg>, accessed on 6 June 2021.
- [13] PyTorch documentation. <https://pytorch.org>, accessed on 6 June 2021.