

Efficient Detection and Recognition of Traffic Lights for Autonomous Vehicles Using CNN

Tayyaba Sahar^{1*}, Hayl Khadhami², Muhammad Rauf¹

Abstract:

Smart city infrastructure and Intelligent Transportation Systems (ITS) need modern traffic monitoring and driver assistance systems such as autonomous traffic signal detection. ITS is a dominant research area among several fields in the domain of artificial intelligence. Traffic signal detection is a key module of autonomous vehicles where accuracy and inference time are amongst the most significant parameters. In this regard, this study aims to detect traffic signals focusing to enhance accuracy and real-time performance. The results and discussion enclose a comparative performance of a CNN-based algorithm YOLO V3 and a handcrafted technique that gives insight for enhanced detection and inference in day and night light. It is important to consider that real-world objects are associated with complex backgrounds, occlusion, climate conditions, and light exposure that deteriorate the performance of sensitive intelligent applications. This study provides a direction to propose a hybrid technique for Traffic Light Detection (TLD) in the daytime and at night. The experimental results successfully improve the night-time detection accuracy of traffic lights from 71.84% to 79.51%.

Keywords: Object Detection; Convolutional Neural Networks; You Only Look Once; Intelligent Transportation Systems; Hough Transform; Traffic Signal Lights

1. Introduction

Humans have eyes and brains that signify their natural ability to detect and classify objects around them. Their detection and recognition capabilities are uncomparable to artificially intelligent systems. Recent progressions prompting upgrades in proficiency and execution in this field would ease human existence by facilitating them through intelligent systems. Similarly, the provision of convenient applications for drivers and road safety is crucial for Intelligent Transportation Systems (ITS). It aims to modernize the operation of vehicles and promotes driver-assisted systems and driverless cars. Traffic light detection in real-time is challenging since it is associated with real-world problems. This includes complex

backgrounds, occlusion, climate conditions, and light exposure that depreciates the performance of sensitive intelligent applications. The efficiency of ITS is crucial, whereas less efficiency is reported due to the variation of light exposures specifically day and night light.

The motivation for this study is inspired by some real-life problems associated with TLD. Problems such as biasness between traffic lights and signs, and variations in day and night light recognition efficiency. Therefore, different experiments and techniques are needed to be tested to enhance the detection accuracy of traffic lights in day and night light.

¹Department of Electronic Engineering, Dawood University of Engineering & Technology, Karachi, Pakistan

²Department of Industrial Engineering and Management, Dawood University of Engineering & Technology, Karachi, Pakistan

Corresponding Author: sahartayyaba@gmail.com

The development of ITS has evolved in two stages: data acquisition and processing followed by the development of technologies for vehicle safety such as collision detection and avoidance [1]. These systems are important for urban planning and future smart cities, because of transportation and transit efficiency [2,3]. TLS is a salient module in Driver Assistance Systems (DAS) and autonomous vehicles [4]. The placement of cameras, the distance of objects, light exposure, and the processing ability of vehicular chips affect the traditional computer vision-based systems used for TLD [4-6]. Deep learning-based CNN models have been deployed in numerous applications after the achievement of incredible results of Alex Net [7]. You Only Look Once (YOLO) version 3 is a 106-layer [8] fully convolutional underlying architecture that is a variant of the original Darknet which comprises 53 layers of a network trained on ImageNet.

Experimental studies reveal that compared to other deep learning models YOLOv3 is a faster, stronger, and more reliable real-time object detector [9]. For automated vehicles, timely detection of traffic lights and their changing states is very important whereas YOLO does generalize object representation without precision losses than other models [10]. A single neural network predicts the boundary boxes and class possibilities from an image of the video frame in a single evaluation. Usually, traffic signs and lights are placed together, these images appear relatively smaller in road view images. Thus, true recognition and detection become challenging as it covers only 1%-2% of the total image area [11]. The combination of CNN and hand-crafted techniques is applied to determine the best input features [12]. Well-acknowledged pattern analysis tools such as Hough Transform (HT) result in realistic outcomes against distortion and diffraction. On the other hand, high material needs and computation costs are two disadvantages that come with it [13]. The latest research and advancements in the field of ML enable to combine the different algorithms from similar or different domains to complement each other [14].

Throughout the years, an increment in the number of cars on roads has increased the frequency of casualties. This endangers human life and safety therefore computer vision techniques are needed to be utilized for observing the immediate data in real-time [15]. Existing studies reveal various intrusive and non-intrusive such as in situ techniques and in-vehicle technologies that are used for traffic monitoring. Computer vision-based techniques have shown better performance than traditional ones [16]. Vitas et al. proposed a hybrid model for the traffic light recognition system using adaptive thresholding and deep learning for region proposal and localization of traffic lights [17]. The researchers used an open-source LISA dataset and custom augmentation to increase the number of data samples. On the other hand, the classification part of the algorithm gave off an 89.60% true detection rate, while the regression correctly localized 92.67% of the traffic lights [17].

TLD is challenging due to the small size and colors that may be similar to the backgrounds. Faster Region-based Convolutional Neural Networks (R-CNN) and Grassmann Manifold Learning have a high degree of accuracy and are robust [18]. Comparative studies of several state-of-the-art methods show that variations and cascaded detection techniques help to deal with real-life issues that affect the detection process [18, 19]. Moreover, other kinds of traffic lights including traffic signs and pedestrian lights have been identified as the main cause of false positives. Deep Convolutional TLD shows an overall detection performance of 0.92 average precision of traffic lights [19].

The brief analysis of the state-of-the-art reveals that TLD using color space and shape detection is dominant for finding the exact parameters of traffic lights [20]. In this paper, we propose a framework for efficient TLD using deep learning and hand-crafted technique to monitor the day and night time efficiency. Moreover, the core contributions of the paper include proposing a framework for the hybrid technique for TLD. The foremost target is to improve the night light efficiency of the algorithm. Whereas, it also emphasizes the

availability of the dataset having an equal proportion of day and night light images for accurate training. In addition, a subset of this type of dataset has been prepared for conducting this experiment.

This technique is practically implementable using real-sense cameras having day and night vision capabilities mounted behind the windshield of autonomous vehicles. Enhanced real-time detection and recognition of traffic lights will help to cope with the increasing rate of road accidents and casualties.

The research is organized in sections as follows: The introduction outlines the background, motivation, brief review of the state-of-the-art, contribution, and application of this technique. The subsequent sections of

this paper include methods and then the experimental results for the applied methodology. The last section concludes the research outcomes.

2. Method

This paper outlines an efficient TLD technique to strike a balance between recognition accuracy and overall model complexity. This experimental study aims to propose a hybrid technique that has been summarized in Fig. 1 & 2. With the proposed method TLD accuracy in different lightening conditions can be detected through three major stages: image preprocessing, feature extraction, and image identification.

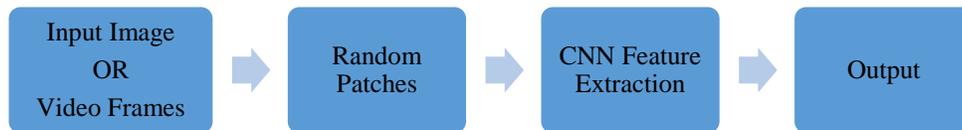


Fig. 1. Traffic Light Detection Using YOLO v3



Fig. 2. Traffic Light Detection Using Hybrid Method

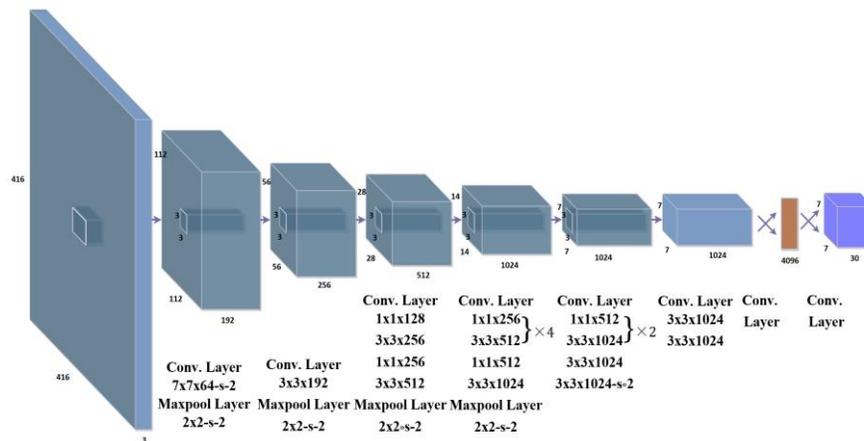


Fig. 3. YOLO Network Architecture [23]

2.1. YOLO v3

In our approach, we have used YOLO, often called a clever CNN, because it is super-fast and supposed to be run in real-time [21-22]. It incorporates features learned by a Deep Convolutional Neural Network for detecting the objects it was first defined by Joseph Redmon and Ali Farhadi in the seminal 2015 paper. This model depends on the “Darknet” architecture shown in Fig.3 [24]. It processes all image features; 2 fully connected layers are used for bounding box prediction for objects.

Mathematically, in terms of regression, an input image is divided into an $S \times S$ grid. Furthermore, the boundary boxes comprise five elements including the object’s x and y coordinates that need to be detected in the input image. The other two elements are w and h. that is the width and height of the same image. The last and fifth element is the most important that reveals the confidence score. The confidence score predicts the presence of

the object in the box, along with the accuracy of the boundary box that is indicated in Fig. 4

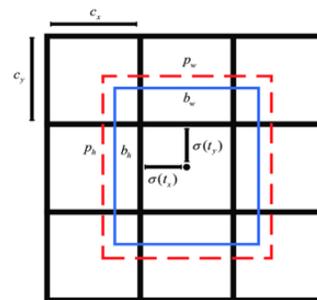


Fig. 4. Elements of boundary boxes [24]

This algorithm also has an appealing feature that enables the end-user to change the model size to make the trade-off between detection accuracy and speed [27]. Consequently, we have made changes within the config file and a number of filters are being changed as per the classes selected for traffic light detection.

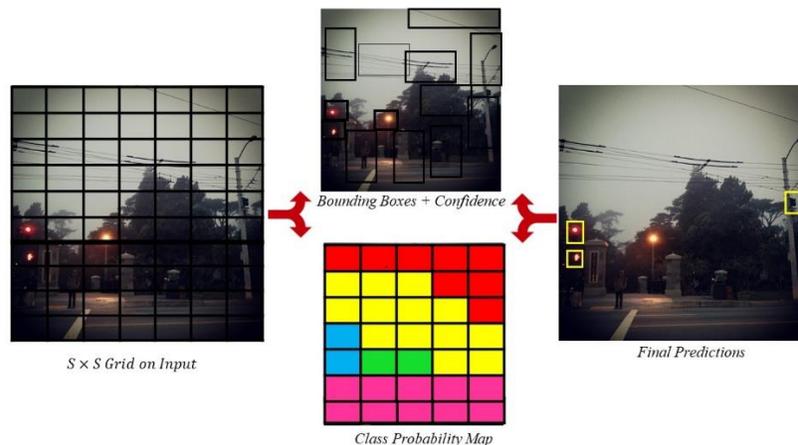


Fig. 5. Steps of YOLO algorithm for TLD

2.2. Hough Transform

To improve the results, a pattern recognition technique called Hough Transform (HT) is deployed. Usually, it is used to extract lines, circles, and similar ellipses or conic sections. All around the world traffic lights are usually circular, in this case, YOLOv3 gives

FP (False Positives) due to biasness between the traffic light and traffic signs. In addition to this, YOLOv3 also struggles while detecting small objects in larger images. Moreover, ITS need more enhanced detection and recognition capabilities in night light as well. In this regard, CHT (Circular Hough Transform) is

used for feature extraction [24] that first converts an image from cartesian to polar coordinates.

Mathematically, a circle is represented as follows:

$$(x - x_{center})^2 + (y - y_{center})^2 = r^2 \quad (1)$$

The center of the circle is (x_{center}, y_{center}) and the radius is r , similarly for using Circle Hough Transform (CHT) for TLD we have used a 2D accumulator with known r as shown in Fig. 6.

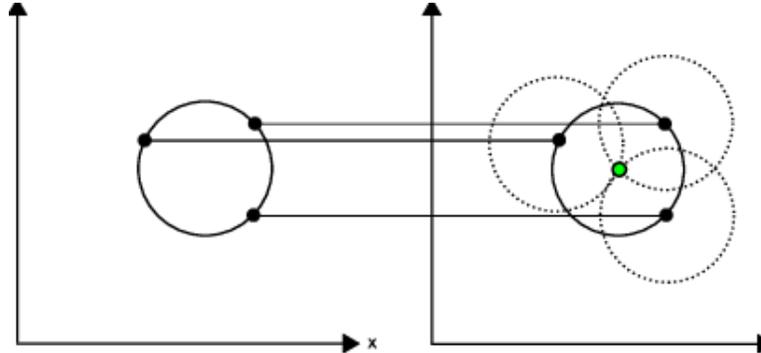


Fig. 6. Circle detection using HT [28]

In other words, the number of unknown parameters is equal to the dimension of the accumulator of a given Hough Transform. HT requires pre-processing to find edges in the original traffic light image, at first traffic light images are converted into binary operation with threshold operation that is followed by a morphological operation to remove background noise, these are depicted in Fig. 8.

As compared to CNN-based models image processing approach is quite uncomplicated, however, it undergoes critical phases such as thresholding, and filtering [26]. In this case, OpenCV is preferred because it enables to Hough Gradient method for gradient information of traffic light edges.

3. Experimental Results

The proposed model has been trained and tested on a total of 500 images at a resolution of 608 x 608 pixels of the Bosch Small Traffic Lights Dataset, 85 % of images are used for training; while 15 % are used for the test. The

annotations include bounding boxes of traffic lights as well as the color of each traffic light. It is known as an accurate dataset for vision-based traffic light detection [21]. The Bosch dataset is selected because it allows easy testing of objection detection approaches, especially for small objects in larger images in different lighting environments.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

At first, traffic light detection has performed using YOLO v3. The Bosch Lite dataset is trained on images with 3 classes (Red, Yellow, and Green) while available unseen test images are used for testing purposes. The following figures show detected traffic lights on the trained model, all three lights including red, yellow, and green are indicated within bounding boxes along with confidence scores.



Fig. 7. Output Images YOLO V3 in different lightening conditions. (a) Green light detection (b) Yellow light detection (c) Red Light detection (d) Misc. lights detection

TABLE I. SUMMARY OF RESULTS

Model	Frames	TP	FP	Day Time Efficiency	Night Time Efficiency
YOLOV3	Day: 1509	1504	5	99.68 %	71.84 %
	Night: 1450	1042	408		
	Total: 2959				

TABLE II. SUMMARY OF CHT RESULTS

Model	Frames	TP	FP	Day Time Eff.	Night Time Eff.
CHT	Day: 1509	1308	202	86.62%	79.51%
	Night: 1450	1153	297		
	Total: 2959				

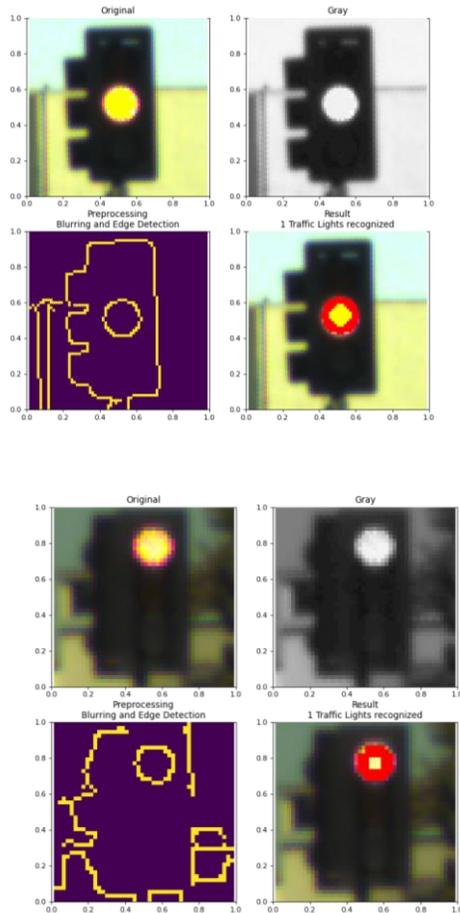


Fig. 8. Output Images of CHT

4. Conclusion

Intelligent Transportation Systems can provide convenient, sustainable, safe, and secure transportability. However, autonomous vehicles undergo several challenges including infallible detection and recognition of traffic lights. These problems can be eliminated by using enhanced computer vision techniques. Over the decades, the expeditious evolution of deep learning models introduced an object detection model YOLO that can efficiently detect traffic lights in real time. Contrary to this, real-life problems limit the accuracy and inference time. In essence, the development of a hybrid approach is needed to encounter the biasness between traffic lights and signs along

with improved efficiency at night time. This study supports the development of a hybrid technique combining deep learning and handcrafted technique to improve traffic light detection and recognition in the best possible way. The experimental results successfully improve the night-time detection accuracy of traffic lights from 71.84% to 79.51%.

REFERENCES

- [1] Rauf, Muhammad, et al. "Response Surface Methodology In-Cooperating Embedded System for Bus's Route Optimization." *Research Journal of Applied Sciences, Engineering and Technology* 5.22 (2013): 5170-5181
- [2] Sharma, S. and S.K. Awasthi, Introduction to intelligent transportation system: overview, classification based on physical architecture, and challenges. *International Journal of Sensor Networks*, 2022. 38(4): p. 215-240.
- [3] Yuan, T., et al., Machine learning for next-generation intelligent transportation systems: A survey. *Transactions on Emerging Telecommunications Technologies*, 2022. 33(4): p. e4427.
- [4] Ouyang, Z., et al., Deep CNN-based real-time traffic light detector for self-driving vehicles. *IEEE transactions on Mobile Computing*, 2019. 19(2): p. 300-313.
- [5] Li, X., et al., Traffic light recognition for complex scene with fusion detections. *IEEE Transactions on Intelligent Transportation Systems*, 2017. 19(1): p. 199-208.
- [6] Jensen, M.B., et al., Vision for looking at traffic lights: Issues, survey, and perspectives. *IEEE Transactions on Intelligent Transportation Systems*, 2016. 17(7): p. 1800-1815.
- [7] Krizhevsky, A., I. Sutskever, and G.E. Hinton, Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 2017. 60(6): p. 84-90.
- [8] Redmon, J. and A. Farhadi, Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.
- [9] Mujahid, A., et al., Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences*, 2021. 11(9): p. 4164.
- [10] Chandana, R. and A. Ramachandra, Real Time Object Detection System with YOLO and CNN Models: A Review. 2022.
- [11] Rehman, Y., et al., Small Traffic Sign Detection in Big Images: Searching Needle in a Hay. *IEEE Access*, 2022. 10: p. 18667-18680.
- [12] Sani, U.S., O.A. Malik, and D.T.C. Lai, Improving Path Loss Prediction Using Environmental Feature Extraction from Satellite Images: Hand-Crafted vs. Convolutional Neural Network. *Applied Sciences*, 2022. 12(15): p. 7685.

- [13] Kumar, S., et al. Lane and Vehicle Detection Using Hough Transform and YOLOv3. in 2022 2nd International Conference on Intelligent Technologies (CONIT). 2022. IEEE.
- [14] Sahar, T., Rauf, M., Murtaza, A., Khan, L. A., Ayub, H., Jameel, S. M., & Ahad, I. U. (2022). Anomaly detection in laser powder bed fusion using machine learning: A review. *Results in Engineering*, 100803.
- [15] Bao, C., et al. Safe driving at traffic lights: An image recognition based approach. in 2019 20th IEEE International Conference on Mobile Data Management (MDM). 2019. IEEE.
- [16] Jain, N.K., R. Saini, and P. Mittal, A review on traffic monitoring system techniques. *Soft Computing: Theories and Applications*, 2019: p. 569-577.
- [17] Vitas, D., M. Tomic, and M. Burul, Traffic light detection in autonomous driving systems. *IEEE Consumer Electronics Magazine*, 2020. 9(4): p. 90-96.
- [18] Gupta, A. and A. Choudhary. A Framework for Traffic Light Detection and Recognition using Deep Learning and Grassmann Manifolds. in 2019 IEEE Intelligent Vehicles Symposium (IV). 2019. IEEE.
- [19] Bach, M., D. Stumper, and K. Dietmayer. Deep convolutional traffic light recognition for automated driving. in 2018 21st International Conference on Intelligent Transportation Systems (ITSC). 2018. IEEE.
- [20] Iftikhar, M., et al., Traffic Light Detection: A cost effective approach. 2022.
- [21] Behrendt, K., L. Novak, and R. Botros. A deep learning approach to traffic lights: Detection, tracking, and classification. in 2017 IEEE International Conference on Robotics and Automation (ICRA). 2017. IEEE.
- [22] Yadav, P.V., et al., AquaVision: Real-Time Identification of Microbes in Freshwater Using YOLOv3, in *Soft Computing for Security Applications*. 2022, Springer. p. 437-448.
- [23] Redmon, J., et al. You only look once: Unified, real-time object detection. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [24] Li, X., et al., A modified YOLOv3 detection method for vision-based water surface garbage capture robot. *International Journal of Advanced Robotic Systems*, 2020. 17(3): p. 1729881420932715.
- [25] Chun, L. Z., Dian, L., Zhi, J. Y., Jing, W., & Zhang, C. Yolov3: Face detection in complex environments. *International Journal of Computational Intelligence Systems*, (2020). 13(1), 1153-1160.
- [26] Gothankar, N., C. Kambhamettu, and P. Moser. Circular hough transform assisted cnn based vehicle axle detection and classification. in 2019 4th International Conference on Intelligent Transportation Engineering (ICITE). 2019. IEEE.
- [27] Kulkarni, R., S. Dhavalikar, and S. Bangar. Traffic light detection and recognition for self driving cars using deep learning. in 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). 2018. IEEE.
- [28] Ribeiro, B.M., et al. Arbitrary ball detection using the circular Hough transform. in *Proc. of the 15th Portuguese Conf. on Pattern Recognition, RECPAD*. 2009.