International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Automotive AI Revolutionizing Driving with Machine Learning

***Jivesh Kumar[1], Dipak Lungba[2], Abba Am John Justin[3], Vishal Kumar[4], Dr. K. Geetha[5]***

Student and Professor of Computer Science and Engineering (2020-24) , Excel Engineering College, NH-544, Salem Main Road, Sankari West, Pallakkapalayam, Pin:637 303. Komarapalayam Namakkal Dt., Chennai ,Tamilnadu

ABSTRACT :

To address the challenges posed by the YOLOv4 algorithm's limited sensitivity to small objects and its suboptimal precision in detecting traffic lights, this paper explores the Enhanced YOLOv4 algorithm. This enhanced approach incorporates two key mechanisms: shallow feature augmentation and bounding box uncertainty prediction. The shallow feature augmentation mechanism enhances the network's capability to pinpoint small objects and improve color resolution. This is achieved by merging shallow features from different stages with high-level semantic features obtained after upsampling iterations. By integrating these features, the network gains better discernment for intricate details, particularly useful in traffic light detection. Moreover, the bounding box uncertainty prediction mechanism introduces a layer of certainty to the prediction process. By modeling the output coordinates of the predicted bounding box and employing Gaussian modeling, the algorithm calculates uncertainty in coordinate information. This enhances the reliability of bounding box predictions, critical for precise object localization.

Experimental evaluation is conducted using the LISA traffic light dataset, separately assessing detection and recognition capabilities. The Enhanced YOLOv4 algorithm demonstrates significant improvements in both aspects. In detection experiments, it achieves an impressive 97.58% area under the PR curve, a notable increase compared to the 90.49% achieved in the Vision for Intelligent Vehicles and Applications Challenge Competition. Meanwhile, in recognition experiments, the mean average precision reaches 82.15%, surpassing the performance of the original YOLOv4 algorithm by 2.86%. These results underscore the Enhanced YOLOv4 algorithm's effectiveness as a robust and practical solution for real-time traffic light detection and recognition tasks, promising enhanced safety and efficiency in intelligent transportation systems.

Keywords: YOLOv4, Improved YOLOv4 algorithm, Shallow feature enhancement mechanism, Bounding box uncertainty prediction mechanism, Traffic light detection, Traffic light recognition, LISA traffic light dataset, PR curve, Mean average precision, Real-time detection, Robust method, Traffic signal lights

## Introduction

Detection technology for traffic lights aids drivers in swiftly discerning the status of signals, enabling prompt decision-making. This, in turn, mitigates driver distractions and curtails instances of non-compliant or illegal driving behaviors. Consequently, the pursuit of a highly accurate, real-time traffic light detection and recognition model holds immense practical significance, offering expansive avenues for enhancing road traffic safety. Typically, traffic light detection systems leverage industrial cameras to capture road conditions. Yet, the intricacies and variability of real-world traffic scenarios pose challenges, as traffic lights in images often occupy minimal pixels and exhibit sparse feature structures. Such complexities elevate the algorithmic hurdles in recognition processes. Thus, the exploration of more efficacious small-target detection algorithms for traffic signal light detection emerges as a critical imperative.

Traditional algorithms for traffic light detection have been extensively explored. In [4], researchers employed YCbCr color space conversion to locate traffic lights, augmenting detection accuracy by leveraging Gabor wavelets' sensitivity to image edges. Meanwhile, [5] converted RGB images to LUV color space for sliding window detection, employing aggregate channel features (ACFs) to identify traffic lights. [6] proposed a multi-scale detection method using ACF and enhanced tree classifiers, effectively detecting red and green lights. In [7], a fast segmentation and compression algorithm improved computational efficiency, while a time-space model enhanced detection accuracy. However, traditional methods struggle with generalization and speed.

Deep learning-based algorithms, such as YOLOv4 [8], offer promising alternatives. Enhancements include optimizing loss functions and increasing network grid units, significantly improving detection. [9] utilized Faster R-CNN, achieving faster detection by leveraging regional suggestion networks. [10] employed YOLOv2, integrating passthrough layers for fine-grained feature detection. Despite GPU acceleration, accuracy remains a challenge.

To address this, improvements to YOLOv4 are proposed. Fusion of shallow and deep features enhances resolution and positioning for small targets. Additionally, a Gaussian model enhances bounding box reliability, further improving detection performance.

## Core Concepts of YOLOv4 Algorithm

This segment provides an overview of the fundamental principles and network architecture underlying the YOLOv4 algorithm.

***Fundamental Principles of YOLOv4 Algorithm***

The YOLOv4 algorithm partitions the input network into S × S grid units, with each grid unit predicting B bounding boxes, bounding box confidence scores, and C category probabilities. The confidence score of a predicted bounding box indicates whether it contains objects and the accuracy of its position, quantified by the intersection over union (IOU) metric (Equation (1)).

confidence = Pr(object) × IOU_truth_pred (1)

Here, 'confidence' denotes the confidence of the bounding box, and 'Pr(object)' represents the probability of an object being detected within the grid. By establishing a threshold for category confidence, bounding boxes with confidence scores surpassing the threshold are retained, and the non-maximum suppression algorithm is applied to obtain the final bounding boxes. Predicted bounding boxes are defined by four parameters: tx, ty, tw, and th. To mitigate the impact of outlier samples on the network, YOLOv4 normalizes these parameters. Illustrated in Figure 1, the input image to the network is sized 608 × 608 and is divided into a 19 × 19 grid. The image's width and height, denoted as 'width_img' and 'height_img' respectively, are subdivided into s × s grid units. The dashed line represents the predicted bounding box, with its center coordinates (x0, y0) and corresponding grid position (row, col). The bounding box's width and height are represented by 'width_box' and 'height_box' respectively. The normalization process is detailed as follows:

(1) Normalize bounding box width and height using Equations (2) and (3) respectively:

tw = width_box / width_img (2)

th = height_box / height_img (3)

(2) Normalization of Center Point Coordinates:
To normalize the center point coordinates, Equations (4) and (5) are utilized:

tx = x0 * s / width_img - col (4)

ty = y0 * s / height_img - row (5)

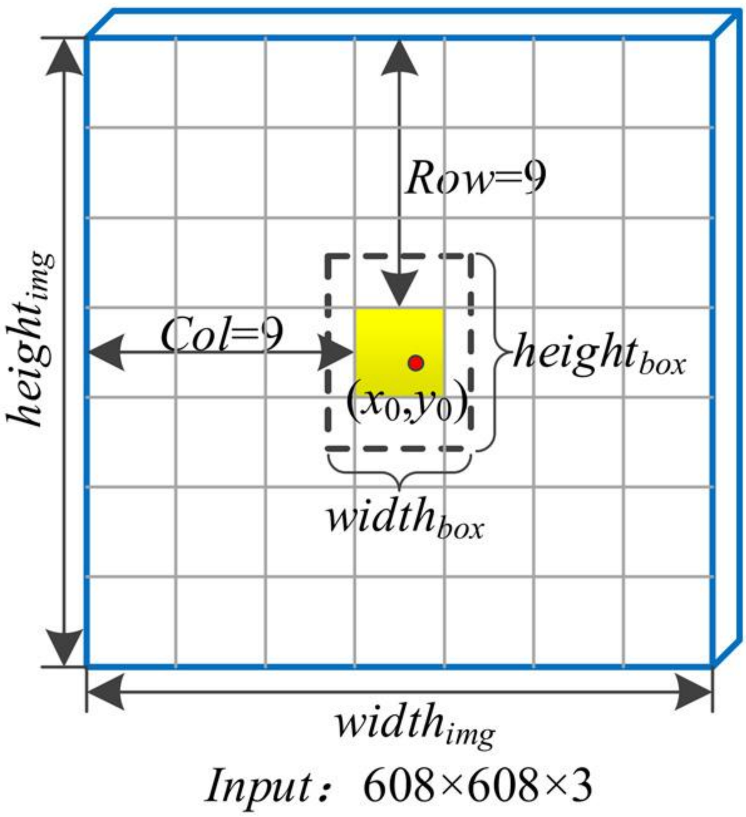
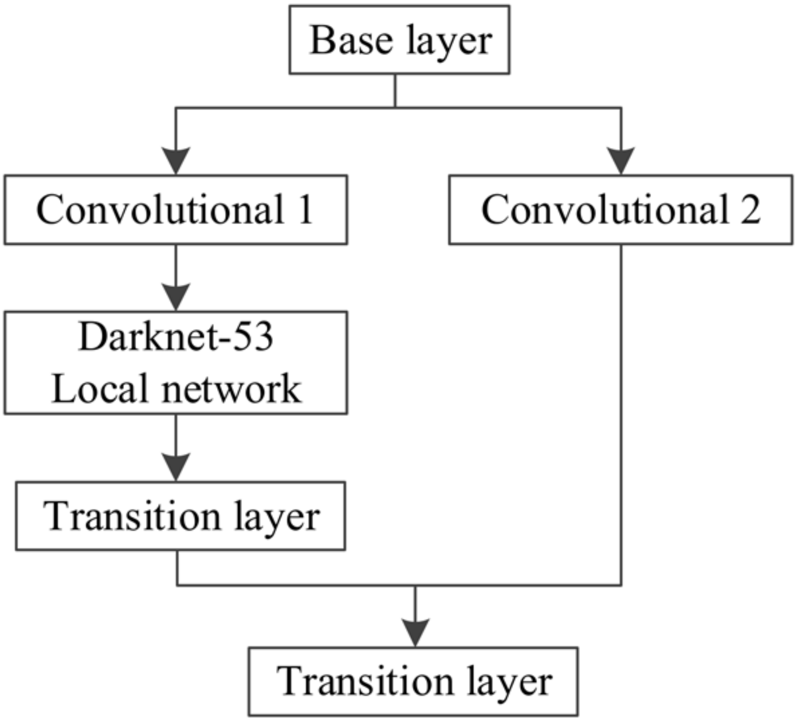


**Figure 1: Standardization of Predicted Bounding Box.**

### CSPDarknet-53 Feature Extraction Framework:

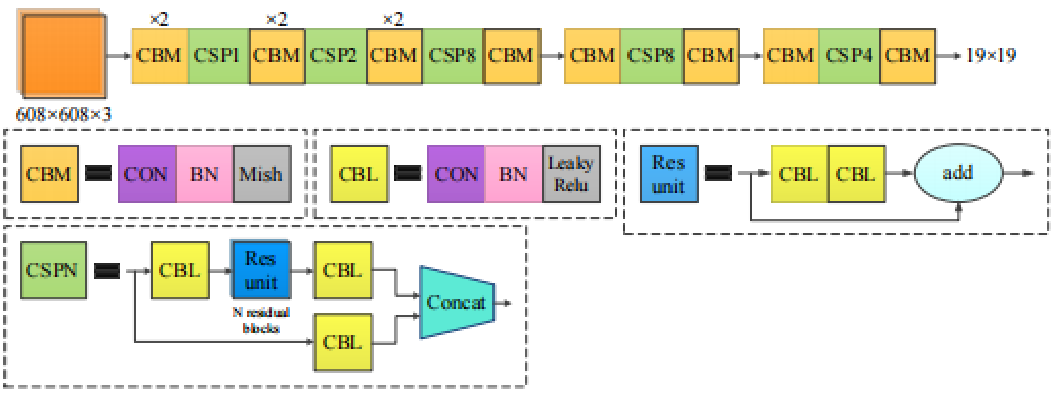
The CSPDarknet-53 feature extraction framework is derived from Darknet-53 and enhanced through the integration of a cross-stage feature fusion approach [18]. This augmentation aims to mitigate redundant gradient acquisition across different layers by introducing a scheme of splitting and fusion across stages. By truncating the flow of gradients during this process, the recurrence of gradient information is curtailed, thereby minimizing the generation of superfluous data. Implementing the cross-stage feature fusion strategy within the local network of Darknet-53 serves to diminish the computational intricacies of the feature extraction framework while concurrently bolstering the network's computational efficiency and precision. The schematic depiction of the cross-stage feature fusion strategy is delineated in Figure 2.




**CSPDarknet-53**

preserves the foundational 52 convolutional layers of Darknet-53 while incorporating a cross-stage feature fusion strategy within Darknet-53's local network. This augmentation aims to mitigate redundancy in the information integration process effectively. The network's architectural diagram is illustrated in Figure 3.

In contrast to Darknet-53, CSPDarknet-53 offers substantial benefits, notably reducing computational overhead, enhancing network inference speed, lowering network memory utilization, and elevating network accuracy.



## YOLOv4 Algorithm Network:

To enhance the YOLOv4 algorithm's capability in detecting small targets, two key improvements are proposed: shallow feature enhancement and bounding box uncertainty prediction. The original CSPDarknet-53 feature extraction network is utilized, but with modifications to enhance precise positioning and recognition of small targets. Shallow features are fused with high-level semantic features, maintaining the network structure integrity. Feature fusion strategies include merging features from layers 11 and 127, and layers 23 and 117. To avoid computational overhead, shallow feature enhancement methods involve upsampling high-level semantic features, splicing shallow features, and conducting convolution operations. This yields three-scale feature information: 152 × 152 × 255, 76 × 76 × 255, and 19 × 19 × 255. The Improved YOLOv4 algorithm's network structure necessitates input data scaling to 608 × 608 × 3, ensuring optimal detection accuracy without excessive computational burden. The 53-CSPDarknet network conducts feature extraction using alternating 3 × 3 and 1 × 1 convolution operations. Shallow feature fusion mitigates the decline in feature map dimensions with increasing convolution depth, preserving sensitivity to small targets. Ultimately, this approach enhances the detection and recognition accuracy of traffic lights by the YOLOv4 algorithm, ensuring real-time detection efficiency without compromising precision.

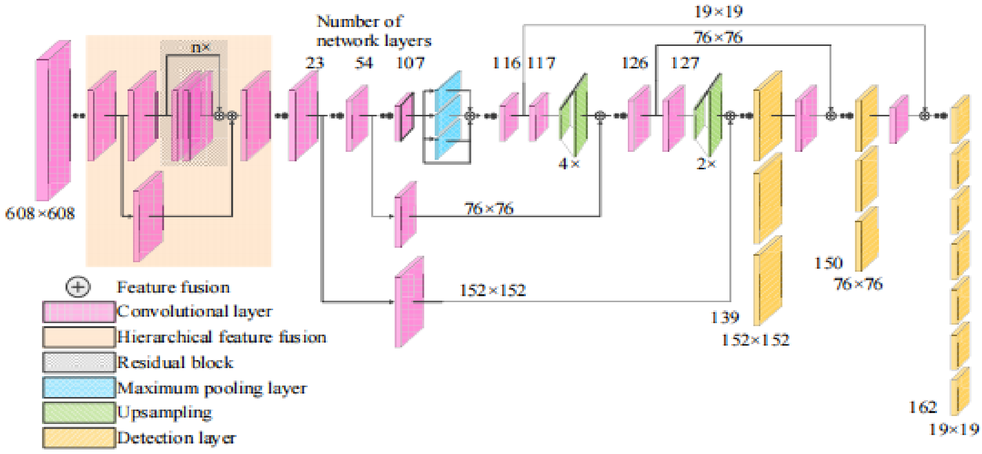



Figure 4. The network structure of the Improved YOLOv4 algorithm.

## Experiment and Analysis :

The experimental setup included Ubuntu 16.04 on an Intel(R) CPU E5-2620-v4 processor, augmented with an Nvidia GeForce GTX 1080Ti GPU for accelerated computations using CUDA 8.0 and cuDNN 6.0 within the Darknet framework. Evaluation of the Improved YOLOv4 algorithm for traffic signal detection utilized the LISA traffic light dataset from the University of California, San Diego's Intelligent and Safe Automobile Laboratory. Captured by a Bumblebee XB3 camera mounted atop a vehicle, the dataset encompasses diverse scenarios, including varying illumination, target coverage, and nighttime conditions. The dataset's complete labels facilitated robust algorithm testing, with LISA-dayTrain serving as the training set and LISA-daySeq1 as the test set

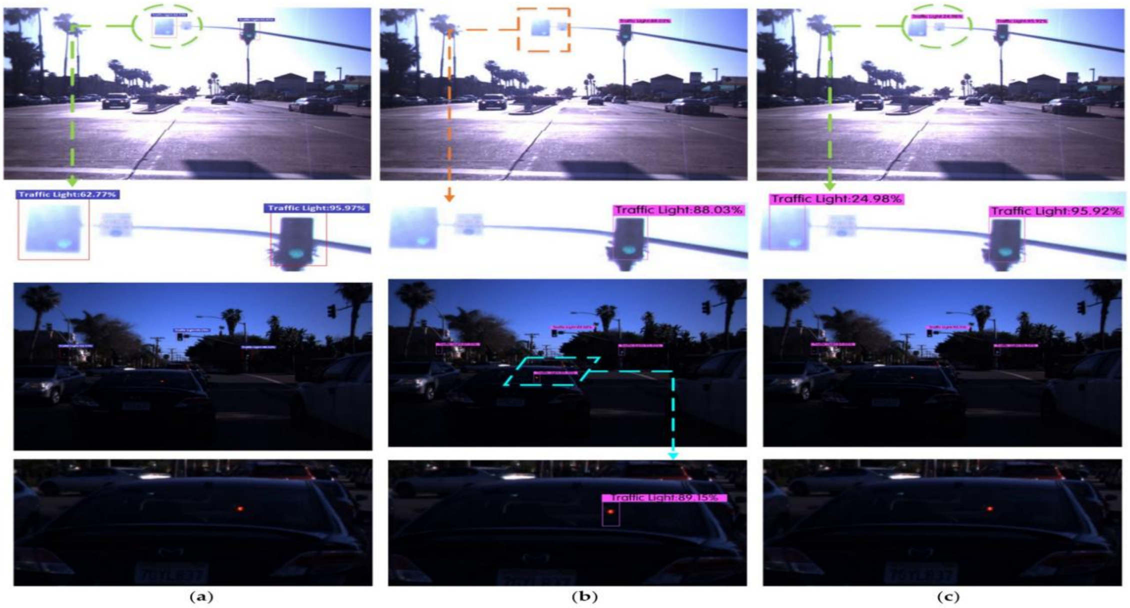
.

**Table 3.** Overview of the LISA traffic light data set.

| Sequence Name | Number of Images | Number of Tags | Image Size |
|---|---|---|---|
| LISA-dayTrain | 14,025 | 40,764 | 1280 × 960 |
| LISA-daySeq2 | 6894 | 11,144 | 1280 × 960 |
| LISA-daySeq1 | 4060 | 10,308 | 1280 × 960 |

### *Analysis of Traffic Lights Detection Performance:*

In traffic light detection, challenges arise due to the small size of the traffic light in the image and the complexity of the background. Factors such as strong lighting, evening conditions, and varying weather can impact detection accuracy. To validate the proposed algorithm's effectiveness in detecting small target traffic lights amidst complex backgrounds, we utilized the AUC value specified in the Vision for Intelligent Vehicles and Applications (VIVA) Challenge Competition as the evaluation metric. The intersection over union (IoU) of true positive samples needed to exceed 0.5 for calculation. We conducted six experiments on the LISA traffic light dataset, employing Faster R-CNN, YOLOv3, YOLOv4, YOLOV4-V1, YOLOV4-V2, and Improved YOLOv4 algorithms. Detection results were compared across different scenes, with yellow rectangles indicating missed detections, blue parallelograms representing incorrect detections, and green circles denoting correct detections.




displays detection results and enlarged images of traffic lights under various algorithms in strong illumination and evening scenarios. YOLOv3 exhibited missed and false detections, while Faster R-CNN, YOLOv4, and Improved YOLOv4 algorithms showed improved performance with minimal errors. Testing on the LISA-daySeq1 dataset revealed the Improved YOLOv4 algorithm significantly reduced missed and false detections, leading to enhanced detection accuracy. Performance metrics in Table 6 indicate Improved YOLOv4 outperformed other algorithms in terms of AUC value, precision, and recall. These findings highlight the effectiveness of proposed enhancements for real-time traffic light detection.


## Conclusions:

We proposed the utilization of the Improved YOLOv5 algorithm for enhancing traffic light detection and recognition. This approach involved incorporating a shallow feature enhancement mechanism and a bounding box uncertainty prediction mechanism. Through the adoption of the Improved YOLOv5 algorithm, we effectively addressed the YOLOv5 algorithm's insensitivity to small targets, significantly improving the accuracy of traffic light detection and recognition. Experimental analyses were conducted using the LISA traffic light dataset and the LaRa traffic light dataset, yielding the following conclusions.Firstly, the implementation of the shallow feature enhancement mechanism optimized the YOLOv5 algorithm, leading to notable improvements in traffic light detection and recognition accuracy. For both the LISA and LaRa datasets, the Area Under Curve (AUC) increased to 97.03% and 95.31%, respectively, in traffic signal detection experiments, while mean Average Precision (mAP) increased to 81.34% and 78.88% in recognition experiments, respectively. Although this enhancement resulted in a slight increase in network calculations and detection time, the method maintained real-time capabilities while effectively enhancing detection and recognition accuracy. Secondly, the integration of a bounding box uncertainty prediction mechanism further improved the YOLOv5 algorithm's accuracy in traffic light detection and recognition. AUC values increased to 96.84% and 94.73% for the LISA and LaRa datasets, respectively, in traffic signal detection experiments, while mAP improved to 79.93% and 78.23% in recognition experiments. Notably, detection time was reduced to 27.59 ms and 33.45 ms, respectively. This method, compared to the Improved YOLOv5 algorithm, exhibited minimal differences in detection and recognition speed while effectively enhancing accuracy.Lastly, combining both optimization methods in the Improved YOLOv5 algorithm resulted in further improvements. For the LISA and LaRa datasets, AUC increased by 1% and 1.19% in traffic light detection experiments compared to the original YOLOv5 algorithm, respectively. In recognition experiments, mAP improved by 2.86% and 2.56%, respectively. The Improved YOLOv5 algorithm demonstrated robustness in reducing missed and false detections, particularly under complex traffic light backgrounds. Additionally, testing on datasets collected by different cameras validated the algorithm's scalability.Looking forward, future investigations should focus on target tracking to predict traffic light movement trajectories and statuses relative to vehicles, further enhancing the reliability of the Improved YOLOv5 algorithm for traffic light detection and recognition.

The Improved YOLOv5 algorithm exhibits promising advancements in traffic light detection and recognition. Through the integration of shallow feature enhancement and bounding box uncertainty prediction mechanisms, it effectively tackles challenges such as insensitivity to small targets. Experimental results on the LISA and LaRa datasets demonstrate significant enhancements in detection accuracy, with minimal impact on real-time capabilities. By achieving higher AUC and mAP values compared to the original YOLOv5 algorithm, it showcases robustness against complex traffic scenarios. Future investigations may explore further enhancements in target tracking, ensuring continued reliability and applicability in real-world traffic environments.
nments.

References :

1. Wang, Z.-Y.; Miao, D.-Q.; Zhao, C.-R.; Luo, S.; Wei, Z.-H. Pedestrian tracking and detection combined algorithm based on multi-granularity features. *Comput. Res. Dev.* **2020**, *57*, 996–1002.

2. Xiang, N.; Cao, Z.; Wang, Y.; Jia, Q. A Real-Time Vehicle Traffic Light Detection Algorithm Based on Modified YOLOv3. In Proceedings of the 2021 IEEE 4th International Conference on Electronics Technology, Chengdu, China, 7–10 May 2021; pp. 844–850.

3. Sun, Y.-C.; Pan, S.-G.; Zhao, T.; Gao, W.; Wei, J.-S. Traffic Light Detection Based on YOLOv3 Optimization Algorithm. *Acta Opt. Sin.* **2020**, *40*, 143–151.

4. Cai, Z.-X.; Gu, M.-Q. Accurate Recognition and State Estimation of Traffic Lights in Urban Environment. *Control. Decis.* **2014**, *29*, 163–167.

5. Philipsen, M.P.; Jensen, M.B.; Møgelmose, A.; Moeslund, T.B.; Trivedi, M.M. Traffic Light Detection: A Learning Algorithm and Evaluations on Challenging Dataset. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Gran Canaria, Spain, 15–18 September 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 2341–2345.

6. Du, X.-P.; Li, Y.; Guo, Y.; Xiong, H. Vision-Based Traffic Light Detection for Intelligent Vehicles. In Proceedings of the 2017 4th International Conference on Information Science and Control Engineering, Changsha, China, 21–23 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1323–1326.

7. Li, Z.-X.; Qin, B.; Wang, M.-Q. Real-time Detection and Recognition of Traffic Signal Lights Based on Time-space Relationship Model. *Comput. Sci.* **2018**, *45*, 314–319.

8. Behrendt, K.; Novak, L.; Botros, R. A Deep Learning Approach to Traffic Lights: Detection, Tracking, and Classification. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation, Singapore, 29 May–3 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1370–1377.

9. Pan, W.-G.; Chen, Y.-H.; Liu, B.; Shi, H.-L. Detection and Recognition of Traffic Lights Based on Faster R-CNN. *Transducer Microsyst. Technol.* **2019**, *38*, 147–149.

10. Morten, B.J.; Kamal, N.; Thomas, B. Evaluating State-of-the-art Object Detector on Challenging Traffic Light Data. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 882–888.