Research Article

A Dynamic Multitarget Detection Algorithm in front of Vehicle Based on Embedded System and Internet of Things

Huili Dou\(^1\) and Guohua Wang\(^2\)

\(^1\)Institute of Rail Transit, Zhejiang Institute of Communications, Hangzhou, Zhejiang 310012, China
\(^2\)Transportation Mechanical and Electrical Engineering Institute of Planning & Design, Zhejiang Provincial Institute of Communications Planning, Design and Research Co., Ltd., Hangzhou, Zhejiang 310030, China

Correspondence should be addressed to Huili Dou; wanggh@zjic.com

Received 18 January 2022; Revised 13 February 2022; Accepted 18 February 2022; Published 20 March 2022

Academic Editor: Muhammad Usman

Copyright © 2022 Huili Dou and Guohua Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are few studies for the classification detection and dynamic multitarget detection of the targets in front of vehicles. In order to solve this problem, a dynamic multitarget detection algorithm is proposed. First, a dynamic multitarget detection with displacement at any time is suggested; secondly, a multitarget detection algorithm based on improved You Only Look Once version 3 (YOLOv3) is proposed for the detection of multitarget high probability risk events in front of the vehicle. The YOLOv3 algorithm model is a lightweight backbone network that uses embedded real-time detection technologies. In this paper, we use a lightweight Mobilenetv2 to replace Darknet-53 for feature extraction. Moreover, an optimizer is used for multiobjective feature extraction, group normalization, and multiobjective feature extraction. The results show that in comparison with the original YOLOv3 algorithm, the detection leakage rate of the improved YOLOv3 multitarget detection algorithm is less than 5%, and the amount of model parameters in this paper is reduced by 95% as compared to the traditional data and CPU intertime is reduced to 78%.

1. Introduction

The introduction of smart Internet of Things technology within the traffic management system is essential in light of the rapid expansion of urban modernization [1]. It is inseparable from the support of developing innovations to effectively adapt to the complex and changing traffic management needs. Sophisticated computer system, Internet of Things innovation, telecommunication, automatic control technology, and other technologies have been used to create an intelligent traffic management system. Its features include real-time, accurate, and efficient monitoring. To achieve efficient traffic control, the objective of a smart transportation system is to create an evolutionary combination of three parts of traffic involvement: driver, vehicle, and road [2].

Target detections major function and task are to accurately locate the relevant target object information in an image using location and recognition. The main purpose of the Internet of things is to provide network capabilities to the equipment and supplies used in our daily life to form information networking and interconnections and maintain the coverage of an information network service [3]. Vehicle signal recognition and vehicle information image acquisition in the Internet of things mode mainly focus on dynamic multitarget acquisition, which is the mainstream improvement direction of multitarget detection technology algorithm at present [4]. Traditional target detection technologies, such as hog feature technology combined with support vector machine (SVM), have some problems: first, the region selection strategy of sliding window is lacking pertinence, high time complexity, and window redundancy [5]; the second problem is that the combination of manual design features and object detection has no good robustness [6]. Multitarget detection can achieve a technical breakthrough and improve the intelligent level of automobile driving faster by using an algorithm with
international advanced experience, which is based on the advantages of embedded systems and the deep learning convolution neural network of Internet of things [7]. For using a unified algorithm architecture of pedestrian and vehicle joint recognition technology, it is necessary to adopt highly significant regional or redundancy strategies to select specific targets [8]. Through the multitarget online self-learning tracking and multitarget tracking method, the multitarget recognition in front of the vehicle can be realized, and the technology can automatically upgrade the recognition algorithm to make the system track the target more stably [9]. Some scholars have studied a basic network architecture of single-shot detector (SSD) network and visual geometry group (VGG)-16. The assumption is replaced by residual network Resnet-26, which improves the detection algorithm and the real-time accuracy of traffic detection [10]. Some scholars applied the full convolution network technology to the three-dimensional scanning data to combine the three-dimensional point dimension with the two-dimensional grid as the feature extraction method by forming different 2D end-to-end full convolution networks through the candidate regions and detect the vehicle target and frame and achieved good results. Scholars of vehicle technology proposed to design a feature convolution kernel library composed of multiple forms and color Gabor [11] to train and screen the optimal feature extraction convolution kernel group by replacing the low-level convolution kernel group of the original network, so as to improve the detection accuracy [12]. In the upgrading iteration from single image detection to multitarget image detection, the adaptive threshold strategy is added to reduce the missed alarm rate and false alarm rate and realize the target detection in complex traffic scenes [13]. In addition, there are some innovative researches including traffic hazard’s warning sign recognition technology, which turns target detection in complex traffic environment into a method of detection together with signs, forming an effective network combination method.

The above studies solve and put forward relevant methods and specific improvement directions from different angles and technical levels. The main research value of the dynamic multitarget detection algorithm in front of the vehicle in the embedded system combined with the Internet of things lies in the following: a lot of research focuses on how to improve the algorithm, and the detection algorithm of multiple targets is upgraded. The improvement of multitarget detection technology must insist that the multitarget detection algorithm can meet the specific requirements and operation value. Combined with the traffic development trend, single category target detection can no longer meet the needs of the traffic scene in front of the vehicle. The complexity of all target detection has become complex and diverse, and the detection accuracy still needs to be improved. In the next few years, road traffic will be more complex, and it is necessary to find approaches of how to effectively and accurately identify the process of the development using the Internet of things. Thus, the Internet will become more critical when it comes to identifying multiple targets, such as a variety of different vehicles, mixed pedestrians and objects, electric motorcycles, and two-wheeled vehicles which often appear and become dangerous targets, so it is necessary to upgrade and iterate the multi-objective algorithm of vehicle automatic driving technology detection.

By considering the statement of the above problems and according to the analysis of this paper, the target needs to be inspected in front of the vehicle in a complex traffic environment and divided into vehicle dynamic target and vehicle static target. The dynamic target is the displacement in front of the vehicle at any time. The main body applicable to the road includes four-wheel vehicles and two-wheel vehicles. Cars, trucks, and buses are in category of four-wheel vehicles while bicycles, motorcycles, and people are placed in the category of two-wheel vehicles. There are many potential safety problems for pedestrians and cyclists. The static target is about the target that the front of the vehicle will not be displaced. The road auxiliary reference is the traffic signal. The single-stage detection algorithm YOLOv3 algorithm is a basic algorithm aiming at the problem that the model is large and is not suitable for embedded devices; the interfree problem is improved during CPU detection.

The rest of this paper is organized as follows. In Section 2, the basic principle of the proposed algorithm and its optimization is presented. In Section 3, the experimental results are discussed. Finally, this paper is concluded in Section 4.

## 2. Algorithm Principle and Optimization

The multitarget detection algorithm based on deep learning can be divided into regions with convolutional neural networks (r-CNN) series in terms of detection mode. Its multitarget detection method can adopt the two-stage algorithm via the single-stage detection algorithm represented by YOLO series. The two-stage detection algorithm concept is as follows: first, employ algorithm for location information [14]; secondly, produce classification for category information. According to the proposed model, the focus is on real-time detection [15]. The single-stage detection method introduces a novel concept: the dynamic multitarget image in front of the vehicle is transformed into network output [16] by returning the category of bounding box position in the output layer [17] and transforming multitarget detection problem to regression problem by improving the detection speed [18–20].

### 2.1. YOLOv3 Algorithm

YOLO is a series of algorithms. YOLOv3 is the third version of YOLO algorithm. The proposed research work has a network structure that describes the YOLOv3 algorithm’s principle; the network structure diagram is shown in Figure 1.

YOLOv3 architecture does not use classic backbone network structures such as VGG-16 and Resnet-50. The ImageNet classification is used as the backbone network for object detection. ImageNet is recognized as authoritative datasets for evaluating the capabilities of deep convolutional neural networks. Many new networks are developed to improve the performance of existing networks. The
proposed model produces a backbone network as the combination of Darknet-53 feature extraction method. The structure of network is shown in Figure 2. The structure of basic YOLOv3 has neither pool layer nor the full connection layer. The front propagation process is size transformation and is realized by changing the step size of convolution layer. The image edge is reduced to half size and the area is reduced to 1/4 of its original size. Using five samples, the characteristic image is 1/32 of the original image. The YOLOv3 algorithm adopts the fixed pattern noise (FPN) idea [21, 22] in the construction by using multiscale targets of different sizes for detection, and as an output, three characteristics of different scales, 13 × 13, 26 × 26, and 52 × 52, are produced, which makes the detection effect of YOLOv3 more significant than that of Yolo algorithm.

By comparing the YOLOv3 algorithm with the fast r-CNN two-order detection algorithm, the former has obvious advantage in terms of detection speed, but there are some shortcomings: the trained model is large, not suitable for embedded devices, and the influence time in the detection under the CPU is high. In order to solve this problem, the following methods are adopted for the optimization of YOLOv3 algorithm. Model optimization is carried out through the aspects of backbone network optimization and model pruning optimization. This paper considers two aspects of optimization, backbone network optimization adjustment and normalized optimization adjustment, as shown in Figure 3.

2.2. Lightweight Improved Model. MobileNetv2 is utilized as the backbone network to replace Darknet-53 for feature extraction in the lightweight model’s design. MobileNet has been optimized into Mobilenetv2 [23–25]. The model is useful for resolving the compatibility issue between mobile terminals and embedded devices. MobileNet builds depth neural networks using depth separable convolution, based on the streamlined design. In Figure 3, the first layer is a convolution layer, followed by depth convolution and point-by-point convolution layers. The used lightweight model architecture is the segregated deep neural network.

Set the input feature mapping according to the model F, its size is (DF, DF, M), the standard convolution K used is (DK, DK, M, N), and the output feature map is G, size (DG, DG, N). The standard convolution is calculated by

\[
G_{k,l,m} = \sum_{i,j,m,n} F_{k+i-1,l+j-1,m,n} \cdot K_{i,j,m,n}
\]

(1)

The number of inserted channels is M and the number of output channels is N. The corresponding calculation quantity is expressed as

\[
\Omega = D_K \cdot D_K \cdot M \cdot N \cdot DF \cdot DF.
\]

(2)

The modified standard convolution is \(K=(D_K, D_K, M, N)\). Then, it is divided by depth convolution and point-by-point convolution. Specifically, depth convolution is responsible for filtering, and its size is \((D_K, D_K, 1, M)\) and its output characteristic is \((D_G, D_G, M)\). The point-by-point convolution is responsible for the conversion channel, and its size is \((1, 1, M, N)\) and its output is \((D_G, D_G, N)\). The convolution formula of depth convolution is

\[
G_{k,l,m} = \sum_{i,j,m,n} K_{i,j,m,n} \cdot F_{k+i-1,l+j-1,m,n}
\]

(3)

In (3), \(K\) is the depth convolution and convolution kernel is \((D_K, D_K, 1, M)\). Among them, \(m_{th}\) is the application of convolution kernel in \(F\) and \(m_{th}\) is the channel, \(G\) and \(m_{th}\) outside.

The calculated equation of depth convolution and point-by-point convolution is

\[
\Omega = D_K \cdot D_K \cdot M \cdot N \cdot DF + M \cdot N \cdot DF.
\]

(4)

From the perspective of comprehensive calculation, the reduction of comprehensive calculation is more obvious. Through the derivation of the above formula, it can be found that the amount of parameters can be significantly reduced by using depth separation convolution. Thus, MobileNet is improved in aspect of linear bottleneck and inverted residual block, as shown in Figure 4.

The algorithms of RELU6 and PW1x1 are introduced in MobileNets, built in the previous portion. For feature extraction in the lightweight model’s design.
extraction, the point-to-point convolution feature approach is used with deep convolution [26]. Because calculating the characteristics of DW convolution defines its ability to adjust the number of channels, flowing from the top layer to the output, the improvement primarily involves PW convolution added to DW convolution as illustrated in Figure 5. The number of channels flowing to the upper layer is small. DW can only extract features in low-dimensional space, which is not good. Therefore, PW activation function can be added before each DW, generally linear bottleneck, because the activation function can effectively increase nonlinearity in high-dimensional space and destroy features in low-dimensional space. The function of the second PW is dimension reduction.

2.3. Normalization Adjustment. In training of depth convolution neural network, the problems of network deepening, difficult training, and slow convergence arise [27]. BN and GN are shown in Figure 6.

The gradient will disappear when the parameters reach to low level of neural network. Based on BN, the optimization algorithm adds the group normalization method to normalize the same group of the same feature map, while the group is divided into channel dimension. Therefore, the normalization operation is independent from the size of the batch and avoids its influence.

3. Experimental Analysis and Visualization

The experimental platform is divided into two portions, the hardware platform and the software platform. NVIDIA GeForce GTX 1080ti (double) GPU; Intel Core i7-7700 CPU; 32 GB disc TensorFlow1.13.1 GPU deep learning framework; PyCharm Community IDE; Linux-Ubuntu 16.04; and 10 types of targets in the bdd100k data set are analyzed, Python script files are generated, and the seven types of target image data and label files studied in this paper are extracted, all with the goal of studying the dynamic multitarget in front of the vehicle.

3.1. Experiment Data Set. The experimental datasets are shown in Table 1. The statistics of data set adopts instance object, BDD100 K, and team-test. Datasets in the experiment are car, bus, truck, bike, motor, rider, and person.

| Type       | Filters Size | Output |
|------------|--------------|--------|
| Convolutional | 32 3 3 3     | 256 256|
| Convolutional | 64 3 3/2    | 128 128|
| Convolutional | 32 1 1      |        |
| Residual    | 64 3 3      |        |
| Convolutional | 128 3 3/2   | 64 64  |
| Convolutional | 64 1 1      |        |
| Residual    | 128 3 3     | 64 64  |
| Convolutional | 256 3 3/2   | 32 32  |
| Convolutional | 128 1 1     |        |
| Residual    | 256 3 3     | 32 32  |
| Convolutional | 512 3 3/2   | 16 16  |
| Convolutional | 256 1 1     |        |
| Residual    | 512 3 3     | 16 16  |
| Convolutional | 1024 3 3/2  | 8 8    |
| Convolutional | 512 1 1     |        |
| Residual    | 1024 3 3    | 8 8    |
| Avgpool     | 8 8         |        |
| Connected   | 1000        |        |
| SoftMax     | 3           |        |

Figure 2: YOLOv3 algorithmic network.

Figure 3: YOLOv3 optimal adjustment of standard convolution to depth convolution.
3.2. Visual Analysis. Visual analysis is carried out using YOLOv3 algorithm and the optimized YOLOv3 algorithm, respectively. The loss change value of models is recorded, as shown in Figure 7.

After a comparison analysis, it can be shown that using different algorithms introduced in the training process, the optimized algorithms curve change is steeper, the decrease in the curve descent process is slower, and the mitigation algorithm’s effectiveness is lowered. Missed detection rate and false detection rate of the target detection test set in the field of unmanned technology environment perception are considered incredibly essential evaluation indicators in the target detection model, which directly increases the model’s trustworthiness.

3.3. Visualization of Target Detection. The visualization of target detection effect is shown in Figures 8 and 9, the algorithm YOLOv3 MobileNetV2 obtains YOLOv3, and the model visualizes the dynamic multitarget detection in front of the test object vehicle.

In the real scene detection effect, I Xian multitarget detection can be carried out, including all detection objects such as cars, trucks, bicycles, pedestrians, and motorcycles. The effect of multitarget detection in front of vehicles has been significantly improved, which is proved in experiments. The algorithm YOLOv3 MobileNetV2 obtains YOLOv3. The model has significantly improved the visualization of dynamic multitarget detection in front of the test object vehicle.

4. Conclusion
A dynamic multitarget detection algorithm in front of vehicle based on improved YOLOv3 is proposed in this paper. The lightweight network MobileNetv2 is used as a replacement of Darknet-53 to form the backbone network for feature extraction; the normalization and optimizer are adjusted to accelerate the convergence of the network. The optimization algorithm trained the two models and from the

Table 1: Data sets used in the experiment.

| Instance object | BDD100 K | Team-test |
|-----------------|----------|-----------|
| Car             | 1134567  | 11876     |
| Bus             | 17890    | 1 198     |
| Truck           | 43689    | 1092      |
| Bike            | 11783    | 913       |
| Motor           | 4535     | 716       |
| Rider           | 6576     | 6021      |
| Person          | 138762   | 7023      |

Figure 6: Normalization operation.

Figure 7: Visual analysis diagram.

Figure 8: Visualization of real scene detection effect of test object 1.

Figure 9: Visualization of real scene detection effect of test object 2.
perspectives of loss visualization, missed detection rate and false detection rate, model size, and effect time, the fundamental algorithms are compared and analyzed. The experiments show that the optimization model constructed in this paper improved by 0.5% in the map, the parameter quantity is reduced to about 89% compared with the YOLOv3 basic model, and the influence time under the CPU is reduced to about 70%. The visualization of test set and actual scene detection qualitatively verifies that the proposed algorithm has a good detection effect for dynamic multi-target detection in front of the vehicles. From the current research in the field of intelligent networked vehicle environment perception, the image level research is using camera as a sensor, including target detection, target tracking, and semantic or instance segmentation. There is still room for improvement in accuracy and speed, the common practice in the industry is sensor fusion, which matches and fuses the cloud information obtained by radar with the image information obtained by the camera to achieve higher perception accuracy. The significance of this paper is that it can optimize the speed of camera sensor target detection, saved computing time, and space for the subsequent fusion. In this paper, the algorithm is only for those embedded systems that are based on the environmental detection and has not been transplanted to the other embedded platforms, and the robustness of the model has not been specifically analyzed. The follow-up research will focus on this aspect.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Ibrar, A. Akbar, R. Jan et al., “Artnet: Ai-based resource allocation and task offloading in a reconfigurable internet of vehicular networks,” IEEE Transactions on Network Science and Engineering, vol. 9, no. 1, pp. 67–77, 2020.

[2] Z. Wang and Y. Ma, “Detection and recognition of stationary vehicles and seat belts in intelligent Internet of Things traffic management system,” Neural Computing & Applications, vol. 34, no. 5, pp. 3313–3322, 2021.

[3] A. J. Moshayedi, A. Kolahdooz, and L. Lief, Unity in Embedded System Design and Robotics: A Step-by-step Guide, CRC Press, Routledge, 2022.

[4] Y. Ye, Q. Wang, and J. Wang, “Green city air monitoring and architectural digital art design based on IoT embedded system,” Environmental Technology & Innovation, vol. 23, Article ID 101717, 2021.

[5] J. D. J. C. Robles, I. Torres-Zúñiga, J. G. Avina-Cervantes, and M. A. Ibarra-Manzano, “FPGA-based architecture to estimate the input glucose in a dark fermenter using a super-twisting observer,” Microprocessors and Microsystems, vol. 85, Article ID 104297, 2021.

[6] P. Singh and N. Senroy, “Steady-State models of STATCOM and UPFC using flexible holomorphic embedding,” Electric Power Systems Research, vol. 199, Article ID 107390, 2021.

[7] N. Niu, F. Fu, B. Yang et al., “DLIB: A low overhead wear leveling algorithm for embedded systems with hybrid memory,” Microelectronics Reliability, vol. 123, Article ID 114154, 2021.

[8] B. Zeng, Z. Zhang, X. Zhang, and J. Xiang, “Design of refueling control system with multi-controller based on network,” Electronics, vol. 10, no. 13, p. 1559, 2021.

[9] W. Huang, C. Zhang, X. Wu, J. Shen, and Y. Li, “The detection of defects in cellular cell phone backplane with embedded system,” Measurement, vol. 181, Article ID 109598, 2021.

[10] R. Tabish, J.-Y. Wen, R. Pellizzoni et al., “An analyzable inter-core communication framework for high-performance multicore embedded systems,” Journal of Systems Architecture, vol. 118, Article ID 102178, 2021.

[11] J. Shi, J. Bian, J. Richter et al., “MODES: Model-based optimization on distributed embedded systems,” Machine Learning, vol. 110, no. 6, pp. 1527–1547, 2021.

[12] E. Cardillo, C. Li, and A. Caddemi, “Embedded heating, ventilation, and air-conditioning control systems: From traditional technologies toward radar advanced sensing,” Review of Scientific Instruments, vol. 92, no. 6, Article ID 061501, 2021.

[13] M. Calixto-Rodriguez, J. S. V. Martinez, M. A. Meneses-Arcos et al., “Design and development of software for the SILAR control process using a low-cost embedded system,” Processes, vol. 9, no. 6, p. 967, 2021.

[14] J. Kim and J. Cho, “Low-cost embedded system using convolutional neural networks-based spatiotemporal feature map for real-time human action recognition,” Applied Sciences, vol. 11, no. 11, p. 4940, 2021.

[15] Z. Zhang, X. Wang, Q. Hao et al., “High-efficiency parallel cryptographic accelerator for real-time guaranteeing dynamic data security in embedded systems,” Micromachines, vol. 12, no. 5, p. 560, 2021.

[16] J. H. Choi, K. M. Kim, and J. W. Kwak, “TA-CLOCK: Tendency-aware page replacement policy for hybrid main memory in high-performance embedded systems,” Electronics, vol. 10, no. 9, p. 1111, 2021.

[17] S. Luo, L. Zhang, and Y. Fan, “Dynamic multi-objective scheduling for flexible job shop by deep reinforcement learning,” Computers & Industrial Engineering, vol. 159, Article ID 107489, 2021.

[18] W. Yin, Z. Ming, and T. Wen, “Scheduling strategy of electric vehicle charging considering different requirements of grid and users,” Energy, vol. 232, Article ID 121118, 2021.

[19] Z. Li, C. Peng, G. Yu, X. Zhang, Y. Deng, and J. Sun, “Detnet: A backbone network for object detection,” arXiv preprint arXiv:1804.06215, 2018.

[20] A. Thamallah, A. Sakly, and F. M'Sahli, “Dynamic multi-objective matrix control for a class of switched systems,” Transactions of the Institute of Measurement and Control, vol. 43, no. 10, pp. 2229–2245, 2021.

[21] H. Liu, Y. Wang, S. Liu, Q. Liu, Y. Xie, and X. Ma, “May. Research on multi-objective optimal scheduling strategy of photovoltaic and energy storage based on dynamic programming,” IOP Conference Series: Earth and Environmental Science, vol. 781, no. 4, IOP Publishing, Article ID 042011, 2021.

[22] M. Chen and Y. Ma, “Dynamic multi-objective evolutionary algorithm with center point prediction strategy using ensemble Kalman filter,” Soft Computing, vol. 25, no. 7, pp. 5003–5019, 2021.
[23] W. Yankai, W. Shilong, L. Dong, S. Chunfeng, and Y. Bo, "An improved multi-objective whale optimization algorithm for the hybrid flow shop scheduling problem considering device dynamic reconfiguration processes," Expert Systems with Applications, vol. 174, Article ID 114793, 2021.

[24] B. H. Nguyễn, T. Vo-Duy, C. H. Antunes, and J. P. F. Trovão, "Multi-objective benchmark for energy management of dual-source electric vehicles: An optimal control approach," Energy, vol. 223, Article ID 119857, 2021.

[25] A. Ahrari, S. Elsayed, R. Sarker, D. Essam, and C. A. C. Coello, "A heredity-based adaptive variation operator for reinitialization in dynamic multi-objective problems," Applied Soft Computing, vol. 101, Article ID 107027, 2021.

[26] H. Karimi, R. Bahmani, and S. Jadid, "Stochastic multi-objective optimization to design optimal transactive pricing for dynamic demand response programs: A bi-level fuzzy approach," International Journal of Electrical Power & Energy Systems, vol. 125, Article ID 106487, 2021.

[27] L. Zhang, J. Lu, and Z. Yang, "Optimal scheduling of emergency resources for major maritime oil spills considering time-varying demand and transportation networks," European Journal of Operational Research, vol. 293, no. 2, pp. 529–546, 2021.