Heliyon 8 (2022) e11570

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Multi-object detection at night for traffic investigations based on improved SSD framework

Qiang Zhang [a,b,c], Xiaojian Hu [a,b,c,*], Yutao Yue [d], Yanbiao Gu [d], Yizhou Sun [e]

[a] Jiangsu Key Laboratory of Urban ITS, Southeast University, China
[b] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, China
[c] School of Transportation, Southeast University, Southeast University Road #2, Nanjing, 211189, China
[d] Jiangsu JITRI Deep Perception Technology Research Institute Co., Ltd, Wuxi, 214028, China
[e] Dulwich College, London, UK

ARTICLE INFO

Keywords:
Object detection
Night condition
SSD
Medium object
Small object

ABSTRACT

Despite significant progress in vision-based detection methods, the task of detecting traffic objects at night remains challenging. Visual information of medium and small stationary objects is deteriorated due to poor lighting conditions. And the visual information is important for traffic investigations. For meeting the needs of night traffic investigations, this study focuses on presenting a nighttime multi-object detection framework based on Single Shot MultiBox Detector (SSD). Considering the need of traffic investigations, the applicable detection framework is presented for detecting traffic objects, especially medium and small stationary objects. In the framework, the Dense Convolutional Network (DenseNet) and deconvolutional layers are introduced to enhance the feature reuse, and the effectiveness of the optimization is finally verified. In this paper, qualitative and quantitative experiments are presented. The results show that our presented framework has better detection performance for medium and small stationary objects. Moreover, the results show that presented framework has better performance for nighttime traffic investigations at intersections.

1. Introduction

Some hot issues such as traffic object location in complex scenes [1], intelligent traffic video analysis [2], intelligent traffic control [3], traffic safety analysis [4], etc., urgently need to get accurate object information in night scenes. Therefore, it is of great practical significance to explore an effective object detection framework for night scenes, which is not only conducive to the analysis of subsequent object behaviors and trajectories in the field of traffic video surveillance [1], but also conducive to tasks such as data transmission, storage, and annotation [5, 6]. At the same time, it also has important theoretical research value to improve the application level of image analysis, image understanding and image processing technology [7]. It is foreseeable that the object detection framework for night scenes will be applied to more and more tasks [6, 8], and the detection at night will also play a huge role in improving the accuracy of full-time detection in actual security monitoring [9].

For visual traffic surveillance at night, one of the challenges is to monitor medium and small stationary objects [7, 10]. At night, the sharp decline in the ability of traffic surveillance equipment limits its wider application [11]. And the videos recorded by traffic surveillance equipment at night have shortcomings such as less information, more noise, and blurred images [11, 12], which can increase the difficulty of detecting medium and small stationary objects. But these medium and small stationary objects are important for traffic investigation [7]. Therefore, for achieving traffic investigations, applicable methods are presented to detect objects, especially medium and small stationary objects at night [13, 14, 15, 16].

For night detectors in recent years, Kim et al [8] proposed a method of pedestrian detection at nighttime using a visible-light camera and faster region-based convolutional neural network (R–CNN). Although this method improves the accuracy of pedestrian detection at night to a certain extent, it may be difficult to reliably detect small and weak pedestrians in a single image. For overcoming this problem, Wei et al [7] proposed a small and weak target detection method. The method is used to detect small and weak targets in a specific image. But the performance in object classification may not be excellent. For overcoming this

* Corresponding author.
E-mail address: huxiaojian@seu.edu.cn (X. Hu).

https://doi.org/10.1016/j.heliyon.2022.e11570
Received 28 April 2022; Received in revised form 14 July 2022; Accepted 7 November 2022

problem, Kuang et al [17] combined feature selection based on tensor decomposition, and a new object proposal approach to detect night-time multiclass vehicles. Although this method improves detection performance at night to a certain extent, it may be insufficient for detecting stationary objects in a complex environment. Ala et al [11] proposed an embedded system for multi-object detection in traffic surveillance, which includes a new architecture of a deep detector adopted from the Faster R–CNN and an original specialization framework for a traffic object detector. Although this method improves the performance of multi-object detection at night to a certain extent, it may be insufficient for detecting medium objects in a complex environment. Jisoo et al [18] developed a CNN-based human detection approach that can perform pixel-wise segmentation and make fine-grained predictions in terms of the object neighborhood. Although this method improves the accuracy of object detection, the method can only be used for limited objects and visual images. For overcoming above problem, Sudha et al [1] proposed an advanced deep learning method called enhanced you only look once v3 and improved visual background extractor algorithms are used to detect the multi-type and multiple vehicles in an input video. Furthermore, there are a lot of excellent detectors [19, 20, 21, 22] that have been improved for better performance. Carion et al [21] presented DETR that can achieve comparable results to an optimized Faster R–CNN baseline. Wang et al [22] developed the YOLOv7 series of object detection systems. And Liu et al [23] proposed a method to detect vehicles with a small size and partial occlusion. Although the above methods improve detection performance at night to a certain extent, they may be insufficient for detecting medium and small stationary nighttime objects in a complex environment.

Aiming at the problem that the existing methods cannot accurately detect the medium and small stationary traffic object at night, we focus on presenting the effective solution based on Single Shot MultiBox Detector (SSD). Our main contributions are as follows.

(1) Considering the need of traffic investigations, the applicable detection framework is presented for detecting traffic objects, especially medium and small stationary objects.
(2) We present the nighttime traffic data to advance the development of object detection at night.
(3) The Dense Convolutional Network (DenseNet) and deconvolutional layers are introduced to enhance the feature reuse, and the effectiveness of the optimization is finally verified.
(4) It is verified that the presented framework has better performance for nighttime traffic investigations at intersections.

2. Methods

In this section, we introduce the details of our presented framework. The main idea of SSD is to use Convolutional Neural Networks (CNN) as the feature extraction network [24]. And relevant modules of SSD are modified to improve the performance of detecting medium and small stationary objects at night.

2.1. Structure summary for improved SSD

As shown in Figure 1, the dense connection is added to the original network, and a deconvolution layer is added after the last convolutional layer. Through the deconvolution process, the size of feature map is expanded to the same size as the Conv4_3 feature map of the original convolutional layer. Then the size of the feature map is enlarged by performing the deconvolution process on the down-sampled feature map, which can get richer semantic information. In the prediction stage, the feature map generated by the convolution process is stitched with the feature map generated by the deconvolution process, and the stitched result is sent to the detector and the classifier for processing.

2.2. Replacement of feature extraction module

For the feature extraction module of SSD, we refer to the replacement rule of Deep Supervised Object Detector (DSOD) [25], and replace the VGG-16 with the DenseNet in the convolution process. For DenseNet, all forward layers are concatenated as input [26], as shown in formula (1).

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{1}$$

Where $H_l(\cdot)$ is the non-linear transformation; x_0, x_1, …, x_{l-1} are all forward layers. When the module in DenseNet is used to replace the VGG-16 structure, the sizes of the down-sampled feature maps are the same as the standard, which can ensure that the sizes of the final input feature maps are consistent with the sizes of the original feature maps.

Compared with the VGG-16 network, the advantages of DenseNet are as follows.

(1) DenseNet has fewer parameters;
(2) DenseNet can enhance feature reuse;
(3) DenseNet has the advantages of implicit supervision.

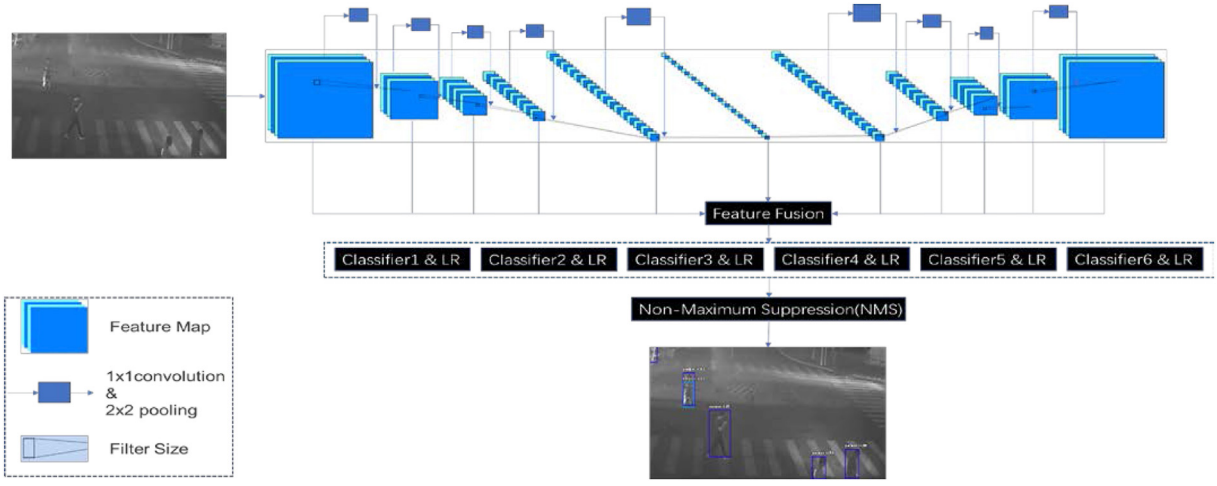

Figure 1. Procedure for improved SSD framework.

2.3. Addition of deconvolution module

In the convolution process, the feature maps with large size tend to have small receptive fields. And the feature maps with large receptive fields are small in size and insensitive to small objects, which can lead to inaccurate detection of small and medium objects by SSD network.

The deconvolution process is introduced, which can not only enlarge the size of the feature map, but also increase the reuse of features by using dense connections. As shown in Figure 2, the deconvolution process is performed on the feature maps of small size. The 3×3 convolution and batch normalization [27] are performed on the feature maps of enlarged size. By means of the dot product, the feature map after deconvolution processing is fused with the feature map obtained by the convolution process.

The input of the deconvolution module is the output of the last convolution layer. And formula (2) is used to determine the parameter of the convolution kernel.

$$F_{\text{out}} = s \times (F_{\text{in}} - 1) + a - 2p_1 \tag{2}$$

Where F_{out} is the size of the output feature map; F_{in} is the size of the input feature map; s is the stride; a is the size of the convolution kernel; p_1 is the meaning of padding.

2.4. Enhancement of feature reuse

Based on the mentioned content, the introduction of DenseNet and deconvolution layers can enhance the feature reuse. In this process, the details of feature reuse can be described as follows.

The SSD network propagates the feature maps in turn according to the layer-by-layer convolution method, and the feature map generated by each convolution layer is only used once. The feature maps are densely connected and reused after fusion. The concatenating of feature maps needs to ensure the same size, so the feature map generated by the previous convolution layer is halved in size.

As shown in Figure 3, h and w represent height and width of the image respectively. P_1 and P_2 are the channel number of the image. As shown in Figure 3, the feature map of the previous convolutional layer is halved using a 1×1 convolutional layer and a 2×2 maximum pooling layer, and then stitched with the feature map obtained by the current convolutional layer. For avoiding the problem of excessive increase in the channel number caused by the dense connection of feature maps, 1×1 convolution is used for dimensionality reduction after each stitching process.

2.5. LOSS calculation

The objective function of this kind of algorithm is divided into two parts, including the confidence loss and the location regression. The calculation is shown in formula (3).

$$L(x, c, l, g) = \frac{1}{N}\left(L_{conf}(x, c) + \alpha L_{loc}(x, l, g)\right) \tag{3}$$

Where N is the number of default boxes that are matched to Ground Truth; L_{conf} is the confidence loss; the α parameter is used to adjust the

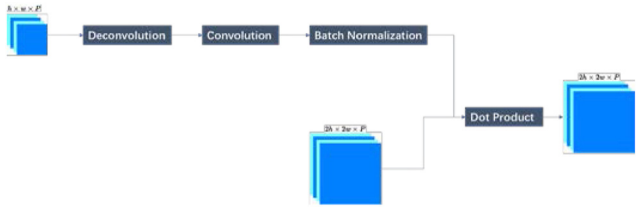

Figure 2. Deconvolutional module.

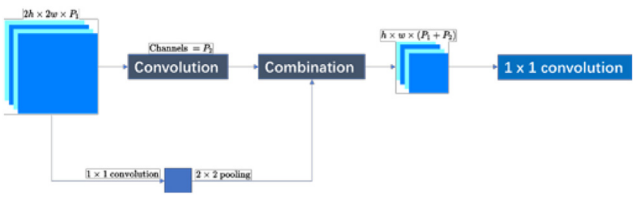

Figure 3. Reusing feature map.

ratio between confidence loss and location loss, and the default α = 1. Location regression uses smooth_{L1} loss, and the calculations are shown in formula (4)-(8):

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^{N} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^{k} smooth_{L1}\left(l_i^m - \hat{g}_j^m\right) \tag{4}$$

$$\hat{g}_j^{cx} = \left(g_j^{cx} - d_i^{cx}\right) \Big/ d_i^{w} \tag{5}$$

$$\hat{g}_j^{cy} = \left(g_j^{cy} - d_i^{cy}\right) \Big/ d_i^{h} \tag{6}$$

$$\hat{g}_j^{w} = \log\left(\frac{g_j^{w}}{d_i^{w}}\right) \tag{7}$$

$$\hat{g}_j^{h} = \log\left(\frac{g_j^{h}}{d_i^{h}}\right) \tag{8}$$

Confidence loss is the softmax loss in multi-category confidence. And the calculations of confidence loss are shown in formula (9)(10).

$$L_{conf}(x, c) = \sum_{i \in Pos}^{N} x_{ij}^{p} \log(\hat{c}_i^{p}) - \sum_{i \in Neg} \log(\hat{c}_i^{0}) \tag{9}$$

$$\hat{c}_i^{p} = \frac{\exp(c_i^{p})}{\sum_{p} \exp(c_i^{p})} \tag{10}$$

For ensuring that the positive and negative samples are as balanced as possible, we apply the hard negative mining used by SSD to control the ratio of positive and negative samples. The difficult negative samples are easy to be classified into the wrong category. Therefore, the core function of the hard negative mining is to select the hard samples among the negative samples for training the network, to ensure the balance of

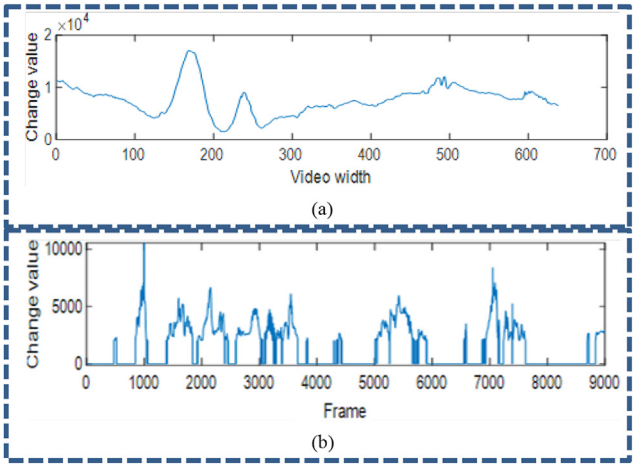

Figure 4. Summary of recorded video data.

Table 1. Object sizes on the nighttime traffic data.

Small (Pel^2)	Medium (Pel^2)	Large (Pel^2)
Size<17250	17250 ≤ Size <33150	Size≥33150

positive and negative samples, and to improve the effectiveness of training.

3. Experiments and results

In this section, we carry out experiments and analyses on the nighttime traffic data and PASCAL VOC data [28] to validate the performance of the presented framework.

3.1. Implementation details

Our training and testing experiments are conducted based on the hardware of CPU (Core i7-9700) and GPU (RTX 2060) processors. The collected images are annotated. Based on the training data set, the improved SSD network is trained. The basic learning rate is set to 0.1 and the momentum is set to 0.9. The learning rate is divided by 10 for every 10,000 iterations. The continuous iteration learning makes the model converge, and the model parameters are stored. Finally, the presented data is used for testing.

3.2. Datasets and evaluation criteria

The datasets consist of the nighttime traffic data and PASCAL VOC data [28]. And object categories in the datasets include person, car, bus, bike, motorbike, etc. On the basis of the Pascal VOC benchmark, we supplement the data of medium and small stationary objects at night. Roadside traffic surveillance equipment is used to collect the nighttime traffic data, especially medium and small stationary objects at night.

The recorded video data is shown in Figure 4. And the object sizes on the nighttime traffic data can be shown in Table 1. As shown in Figure 4 (a), it can show the main activity region of objects. As shown in Figure 4 (b), it can show existences and pixel changes of objects in each frame of the videos. And some specific regions in videos have significant clutter, inconsistent contrast, various levels of illumination, and pose variation of objects.

We use the 11-Point calculation method of the Pascal VOC benchmark [28], which is the approved evaluation method. According to the Pascal VOC benchmark, the performance evaluations are analyzed on the basis of the benchmark metrics such as Average Precision (AP) and mean Average Precision (mAP) [29].

3.3. Ablation study

It is verified whether the optimization of the SSD network structure is effective. The testing results on the nighttime traffic data are shown in Table 2. The replacement of the feature extraction network can improve baseline model by 4% mAP. For the addition of deconvolution module, the size of the relevant feature map is enlarged, and the detection of medium and small stationary objects is improved, which improves baseline model by 5.2% mAP. For the enhancement of feature reuse, the information of medium and small objects is reused. The replacement of the feature extraction network can be used to assist in the feature reuse. And the deconvolution module can be used to increase the feature reuse. Finally, the 47.2% mAP can be achieved.

Table 2. Result of ablation studies on the nighttime traffic data.

Model	Replacement of feature extraction network	Addition of deconvolution module	Enhancement of feature reuse	mAP (%)
Baseline				36.9
Improved SSD	✓			40.9
		✓		42.1
	✓	✓	✓	47.2

Table 3. AP for each category of road traffic objects on PASCAL VOC data.

Model	Data	Person (%)	Car (%)	Bus (%)	Bike (%)	Motorbike (%)
DSOD300 [19]	07 trainval +07 test +12 trainval	84.6	80.6	83.6	85.3	86.8
DSSD 513 [20]	07 trainval +07 test +12 trainval	86.4	85.0	84.3	86.6	87.8
Ours	07 trainval +07 test +12 trainval	**89.6**	**92.5**	**89.4**	**90.8**	**91.2**

3.4. Framework evaluation

As shown in Table 3, Table 4, Figures 5 and 6, the evaluation results are obtained from the presented framework and the excellent methods. Table 3 and Table 4 show the results of quantitative evaluations. And Figures 5 and 6 show the results of qualitative evaluations.

As shown in Table 3, the evaluation results on PASCAL VOC data (07 trainval +07 test +12 trainval) are presented. For person detection, our presented framework is 5% and 3.2% higher than those achieved by DSOD300 and DSSD 513, respectively. For car detection, our presented framework is 11.9% and 7.5% higher than those achieved by DSOD300 and DSSD 513, respectively. For bus detection, our presented framework is 5.8% and 5.1% higher than those achieved by DSOD300 and DSSD 513, respectively. For bike detection, our presented framework is 5.5% and 4.2% higher than those achieved by DSOD300 and DSSD 513, respectively. For motorbike detection, our presented framework is 4.4% and 3.4% higher than those achieved by DSOD300 and DSSD 513, respectively.

As shown in Table 4, the 47.2% mAP of the overall classes can be achieved on the nighttime traffic data. And our presented framework can achieve an efficiency of 29 frames/s. For mAP values, our presented framework is 8.4% and 1.7% higher than those achieved by Faster RCNN-R101-FPN+ and YOLOv7, respectively. For AP_{medium} values, our presented framework is 8.4% and 1.3% higher than those achieved by Faster RCNN-R101-FPN+ and YOLOv7, respectively. For AP_{small} values, our presented framework is 10.2% and 5.6% higher than those achieved by Faster RCNN-R101-FPN+ and YOLOv7, respectively.

As shown in Figures 5 and 6, the results include representative examples of scenes with medium and small stationary objects. As shown in Figure 5, the presented framework has better detection performance for medium and small stationary traffic object. As shown in Figure 6, in these challenging situations, the presented framework can accurately detect relevant traffic objects.

From the quantitative and qualitative results, we conclude that the presented framework can achieve better performance in traffic

Table 4. Comparison of detection performance on the nighttime traffic data.

Method	AP_{large} (%)	AP_{medium} (%)	AP_{small} (%)	mAP (%)	FPS on our processors (frames/s)
Faster RCNN-R101-FPN+ [21]	51.0	43.1	22.2	38.8	8
YOLOv7 [22]	**59.5**	50.2	26.8	45.5	**64**
Ours	57.6	**51.5**	**32.4**	**47.2**	29

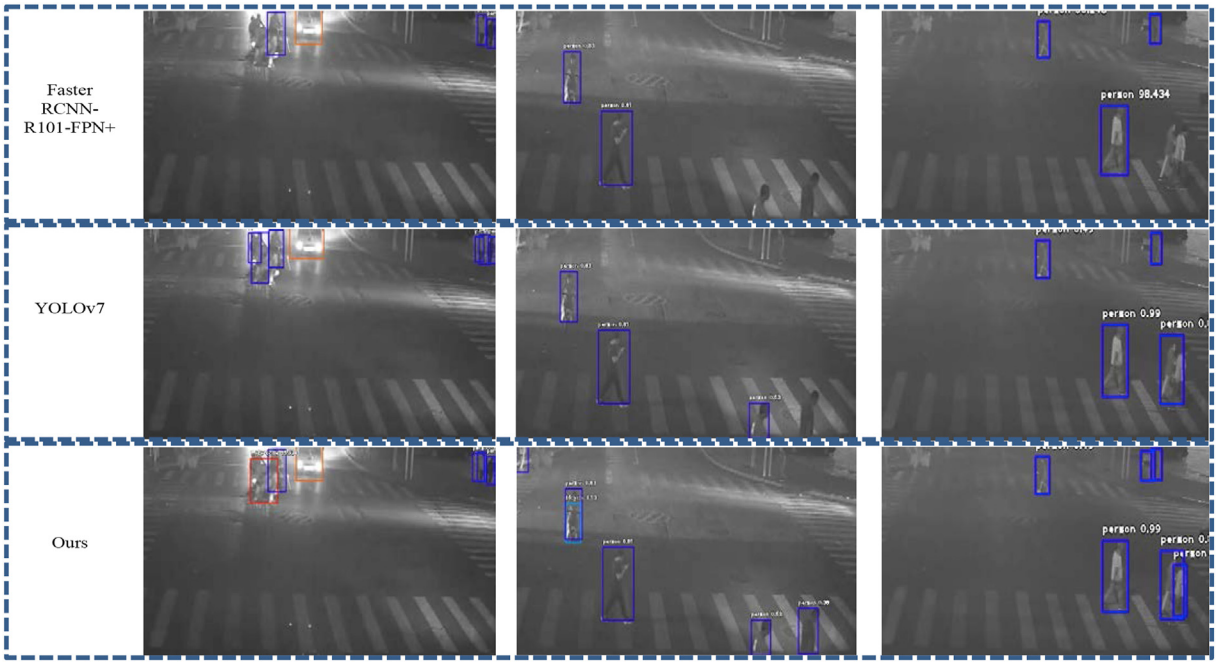

Figure 5. Qualitative evaluation of detection performance.

Figure 6. Detection results by the presented framework in the presence of significant interference.

investigations, especially for detecting medium and small stationary objects at night.

4. Conclusions

In this paper, aiming at the problem that the existing methods cannot accurately detect the medium and small stationary traffic object at night, we focus on presenting the effective solution based on Single Shot MultiBox Detector (SSD). Considering the need of traffic investigations, the applicable detection framework is presented for detecting traffic objects, especially medium and small stationary objects. And we present the nighttime traffic data to advance the development of object detection at night. In the framework, the DenseNet and deconvolutional layers are introduced to enhance the feature reuse, and the effectiveness of the optimization is finally verified. Finally, it is verified that the presented framework has better performance for nighttime traffic investigations at

intersections. Furthermore, we found that some occluded objects in the adverse condition may not be detected accurately. For future work, traffic objects in adverse weather will be detected for traffic investigations. On the other hand, we plan to further study on how to apply the more effective CNN for dynamic videos to get a practical application.

Declarations

Author contribution statement

Qiang Zhang and Xiaojian Hu: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Yutao Yue and Yanbiao Gu: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yizhou Sun: Analyzed and interpreted the data; Wrote the paper.

Funding statement

This work was supported in part by National Natural Science Foundation of China (grant number 52272344), Key Research and Development Program of Jiangsu Province (grant number BE2019713), Key Research and Development Program of Jiangsu Province (grant number BE2018754) and the Fundamental Research Funds for the Central Universities.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] D. Sudha, J. Priyadarshini, An intelligent multiple vehicle detection and tracking using modified vibe algorithm and deep learning algorithm, Soft Comput. 24 (22) (Nov, 2020) 17417–17429.

[2] G. Lee, R. Mallipeddi, G.-J. Jang, M. Lee, A genetic algorithm-based moving object detection for real-time traffic surveillance, IEEE Signal Process. Lett. 22 (10) (Oct, 2015) 1619–1622.

[3] G. Zhang, S. Lu, W. Zhang, CAD-net: a context-aware detection network for objects in remote sensing imagery, IEEE Trans. Geosci. Rem. Sens. 57 (12) (Dec, 2019) 10015–10024.

[4] Y. Xing, C. Lv, H. Wang, D. Cao, E. Velenis, F.-Y. Wang, Driver activity recognition for intelligent vehicles: a deep learning approach, IEEE Trans. Veh. Technol. 68 (6) (Jun, 2019) 5379–5390.

[5] L. Zhu, F.R. Yu, Y. Wang, B. Ning, T. Tang, Big data analytics in intelligent transportation systems: a survey, IEEE Trans. Intell. Transport. Syst. 20 (1) (Jan, 2019) 383–398.

[6] T. Akilan, Q.M.J. Wu, W. Zhang, Video foreground extraction using multi-view receptive field and EncoderDecoder DCNN for traffic and surveillance applications, IEEE Trans. Veh. Technol. 68 (10) (2019) 9478–9493, Oct.

[7] M.-S. Wei, F. Xing, Z. You, A real-time detection and positioning method for small and weak targets using a 1D morphology-based approach in 2D images, Light Sci. Appl. 7 (May 4, 2018).

[8] J.H. Kim, G. Batchuluun, K.R. Park, Pedestrian detection based on faster R-CNN in nighttime by fusing deep convolutional features of successive images, Expert Syst. Appl. 114 (Dec 30, 2018) 15–33.

[9] V.A. Sindagi, V.M. Patel, A survey of recent advances in CNN-based single image crowd counting and density estimation, Pattern Recogn. Lett. 107 (May 1, 2018) 3–16.

[10] Y. Liu, T. Qiu, J. Wang, W. Qi, A nighttime vehicle detection method with attentive GAN for accurate classification and regression, Entropy 23 (11) (Nov, 2021).

[11] A. Mhalla, T. Chateau, S. Gazzah, N.E. Ben Amara, An embedded computer-vision system for multi-object detection in traffic surveillance, IEEE Trans. Intell. Transport. Syst. 20 (11) (Nov, 2019) 4006–4018.

[12] Y. Zhu, Z. Jia, J. Yang, N.K. Kasabov, Change detection in multitemporal monitoring images under low illumination, IEEE Access 8 (2020) 126700–126712.

[13] J. Chu, Z. Guo, L. Leng, Object detection based on mult-layer convolution feature fusion and online hard example mining, IEEE Access 6 (2018) 19959–19967.

[14] H. Tayara, K.G. Soo, K.T. Chong, Vehicle detection and counting in high-resolution aerial images using convolutional regression neural network, IEEE Access 6 (2018) 2220–2230.

[15] Y. Tian, J. Gelernter, X. Wang, W. Chen, J. Gao, Y. Zhang, X. Li, Lane marking detection via deep convolutional neural network, Neurocomputing 280 (Mar 6, 2018) 46–55.

[16] T. Yang, X. Long, A.K. Sangaiah, Z. Zheng, C. Tong, Deep detection network for real-life traffic sign in vehicular networks, Comput. Network. 136 (May 8, 2018) 95–104.

[17] H. Kuang, L. Chen, L.L.H. Chan, R.C.C. Cheung, H. Yan, Feature selection based on tensor decomposition and object proposal for night-time multiclass vehicle detection, Ieee Transactions on Systems Man Cybernetics-Systems 49 (1) (Jan, 2019) 71–80.

[18] J. Park, J. Chen, Y.K. Cho, D.Y. Kang, B.J. Son, CNN-based person detection using infrared images for night-time intrusion warning systems, Sensors 20 (1) (Jan, 2020).

[19] Z. Shen, L. Zhuang, J. Li, Y.G. Jiang, X. Xue, DSOD: learning deeply supervised object detectors from scratch, 2017.

[20] C.Y. Fu, W. Liu, A. Ranga, A. Tyagi, A.C. Berg, DSSD : Deconvolutional Single Shot Detector, 2017.

[21] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko, End to End object detection with transformers, Computer Vision – ECCV (2020) 213–229.

[22] C.Y. Wang, A. Bochkovskiy, H. Liao, YOLOv7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors, 2022.

[23] Y. Liu, T.T. Qiu, J.W. Wang, W.T. Qi, A nighttime vehicle detection method with attentive GAN for accurate classification and regression, Entropy 23 (11) (Nov, 2021).

[24] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, SSD: single Shot MultiBox detector, 2016.

[25] Z. Shen, Z. Liu, J. Li, Y.-G. Jiang, Y. Chen, X. Xue, Ieee, DSOD: learning deeply supervised object detectors from scratch, in: IEEE International Conference on Computer Vision, 2017, pp. 1937–1945.

[26] G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, IEEE, "densely connected convolutional networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2261–2269.

[27] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, Proceedings of Machine Learning Research (2015) 448–456.

[28] A. Moffat, J. Zobel, Rank-biased precision for measurement of retrieval effectiveness, ACM Trans. Inf. Syst. 27 (1) (2009).

[29] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, SSD: single Shot MultiBox detector, in: J. Matas, N. Sebe, M. Welling (Eds.), *Computer Vision - Eccv 2016, Pt I*, Lecture Notes in Computer Science B. Leibe, 2016, pp. 21–37.