A Survey of Small Object Detection Based on Deep Learning

Zhang zhenghua^{a*}, Jiang ling^a, Hong qingqing^a

^aYangzhou University, 88 South University Road, Yangzhou City, Jiangsu Province, China

*Corresponding Author: 3265850874@qq.com

Abstract

As a basic visual recognition problem in computer vision, object detection has made great progress based on traditional manual features and deep learning algorithms. However, researches on small object detection have only begun to appear in recent years, which has become a hot and difficult point in the field and most of them are improved on the basis of existing object detection algorithms to enhance the detection accuracy. With the rapid development of deep learning, small object detection based on deep learning has made great progress, which has wide application requirements in the fields of automatic driving, remote sensing image detection, criminal investigation and other fields, so the research on small object detection has strong practical values. In this paper, the existing research on small target detection is reviewed in detail. Firstly, the existing algorithms are divided into one stage and two stages according to the number of detection stages, and then the characteristics of these algorithms are analyzed; Secondly, the small object detection datasets commonly used are introduced. Finally, the challenges of small object detection are summarized, and the future research directions are prospected.

Keywords: small object detection, deep learning, feature extraction, convolutional neural network.

1. Introduction

Small object detection has extensive and important applications in many fields [1]. For example, in the field of automatic driving, pedestrian objects or traffic signs in photos collected by cars are too small, but the accurate detection of these small objects is an import guarantee for the realization of safe automatic driving; In the field of medicine, the successful detection of small tumor is the premise of early and accurate diagnosis of tumors; The small defects on

the surface of the automatic industrial inspection and positioning materials also reflect the importance of small object detection; In satellite image analysis, it is necessary to detect objects such as cars and ships, which are often difficult to detect because of their small scale; In the field of criminal investigation, the efficiency of solving cases can be greatly improved by effectively identifying abnormal small furnishings. Therefore, it is an import and urgent research topic to study the effective methods of small object detection and improve the detection performance of small object.

There are two main definitions of small objects [2]: the first is absolute small objects. The COCO dataset [3][43] specifies that when the pixel number of objects is less than 32×32 , the object can be regarded as a small object; The second is relatively small object. It is considered as relatively small object when the object size is less than 0.1 times the size of the original drawing. Small object is difficult to detect. Firstly, the small object itself has fewer pixels and less feature information; Secondly, because the small object has fewer pixels, it is more difficult to extract features after multiple down sampling [4]. Therefore, small object detection faces great challenges. At present, there are few algorithms for small object detection, and most of the existing algorithms are also modified on the basis of object detection algorithms, thus improving the detection effect of small objects [5][6][7].

In the traditional object detection based on machine learning, small objects are mainly detected by constructing image pyramid at the bottom of the pyramid. With the emergence and development of deep learning, the method of using image pyramid to detect objects of different scales is gradually replaced by convolutional neutral network [8], which can effectively improve the detection performance of different scale objects by forming multi-level and rich feature expressions. In the convolutional neural network, the underlying features contain abundant details, which is conducive to small object detection; High level features

contain rich semantic information, which is conducive to large object detection. With the continuous in-deep research, the detection effect of small objects has been greatly improved, but there is still a big gap between the detection performance of medium and large objects.

Regarding the progress of research on small object detection, Mask R-CNN proposed by He et al. [29] in 2017 has made a breakthrough, which not only solves the long-standing problem of precision loss of R-CNN, but also improves the situation of small object missing detection; And the contour of the detected object can be segmented, which makes the positioning of the object more accurate. Based on the good performance of Mask R-CNN in object detection, Yang et al. [30] proposed a ship detection method and a training strategy of hard sample relearning. Hard sample training improves the ability to adapt to difficult samples, so as to improve the network's ability to locate and recognize objects in complex scenarios, effectively solving the problems of cloud occlusion and shore-based interference, and reducing the situation of small object ships missing inspection; Secondly, the learning strategy of hard sample relearning is used to weaken the interference of cloud occlusion and land background, which solves the problem of ship segmentation and recognition in complex scenes to a certain extent. Hu P et al. [31] proposed an algorithm for small face detection, which studied the problem of small face recognition from three aspects of scale invariance, image resolution, and contextual reasoning. The number of faces that can be detected in a picture with 1000 faces is about 800.

2. Research status of small object detection algorithms

At present, small object detection algorithms are based on convolutional neural networks [9][10][11]. The popular small object detection algorithms can be divided into two categories: the first category is one-stage object detection method, which uses a convolutional neural network to directly predict different object categories and locations without extracting the object candidate areas. The problem of object frame location is transformed into the problem of regression, which directly generates the object category probability and location coordinates, and has the advantage of speed; The second is two-stage object detection method, that is, firstly, candidate regions that may contain objects are generated by selective search [12], edge detection, region extraction network and other methods, and then convolutional neural network is used to extract features of

candidate regions for accurate object class estimation and boundary box position regression, which needs to go through two stages to obtain the final detection effect. The kind of method has high accuracy, but the processing efficiency and the inference speed is low, which is not suitable for real-time detection tasks. Both one-stage object detection method and two-stage object detection method are faced with the difficulty of small objects detection.

2.1 Roadmap of small object detection

In the past ten years, small object detection based on deep learning is mainly divided into two stages, one-stage object detection method and two-stage object detection method. The main development roadmap is shown in Fig.1.

Fig.1 Roadmap for small object detection

2.2 One-stage small object detection algorithm

(a) YOLO algorithm

YOLO [13] was proposed by Joseph et al. in 2015. It is the first single-stage object detector in the era of deep learning. YOLO is the abbreviation of "You Only Look Once". We can see from its name that only one feature extraction is performed on the input object. The author completely abandoned the previous "regional candidate + classification" detection mode. On the contrary, it follows a completely different science: a single neural network is applied to the entire input image, which divides the image into multiple regional grids and simultaneously predicts the bounding box and probability of each grid. Later, Joseph made a series of improvements on the basis of YOLO and proposed the YOLOv2[14], which further improved the detection accuracy and maintained a high detection speed. Although the detection speed of YOLOv2 has been greatly improved, compared with the two-stage detector, the positioning accuracy of YOLOv2 still lags behind, especially for some small objects. Afterwards, the improved YOLOv3 [15] learned from the residual network structure, formed a deeper network level and multi-scale detection, and

improved the detection effect of small objects.

(b) SSD algorithm

SSD [16] was proposed by Wei Liu et al. in 2015, which is the second single-stage detector based on deep learning. The main contribution of SSD is to extract feature maps of different scales for detection and adopt anchor frames of different scales and aspect ratios, which have significantly improved the detection accuracy of single-stage detector, especially for certain small objects. SSD has advantages in both detection speed and accuracy. The main difference between SSD and previous detectors is that the former detects on the feature map of different scales, while the latter only detects the last layer. Although the SSD is simple and fast, its detection accuracy is somewhat inferior to that of a two-stage detector.

(c) Retina-Net algorithm

Lin [17] et al. found the reason for the low accuracy through experiments, and proposed Retina-Net in 2017. They believe that the main reason is the unbalanced distribution of positive and negative samples in the training process. In order to solve this problem, the cross entropy loss has been improved and a loss function named focal loss is designed. This function can reduce the weight of easy-to-classify samples so that the network will focus more on difficult-to-classify samples in training, thereby successfully solving the problems that the positive and negative sample areas are extremely unbalanced, and the object detection loss is easily controlled by a large number of negative samples. The focal loss can make the single-stage detector achieve the accuracy of two-stage detector, while maintaining a high detection speed.

Since the one-stage small object detection method is fast, but the accuracy is low, the two-stage small object detection method is slow, but the accuracy is high. It is hoped that the accuracy of the one-stage small object detection method can reach the two-stage small object detection method, without affecting the original speed.

2.3 Two-stage small object detection algorithm

(a) R-CNN

In 2014, Girshick et al. proposed the R-CNN network structure. Firstly, about 2000 candidate regions were extracted by using the selective search [12] method, and then the convolutional neural network was used to calculate the feature map of each candidate region. After the feature maps were obtained, it was classified by using the support vector machine to distinguish the background and object. Since there will be multiple overlapping candidate regions near an

object, the non-maximum suppression [19] method is used to remove the overlapping candidate regions, and finally the frame regression method is used to correct the position of the candidate frame. Because of its creative fusion of convolutional neural networks and object detection, it is known as the pioneer of modern object detection.

(b) SPPNet

In 2015, He Kaiming et al. proposed the Spatial Pyramid Pooling Network [20]. Previous convolutional neural networks require fixed input size, for example, the input size of AlexNet [21] is 224x224. The main contribution of SPPNet is the introduction of the Spatial Pyramid Pool layer, which enables the convolutional neural network to generate feature maps of fixed size, independent of the size of the image or region of interest, without rescaling. When SPPNet is used for object detection, only one feature extraction is needed for the whole image, and then SPP layer is used to adjust the resolution of the feature map of the candidate region to a fixed size to train the detector, so as to avoid the repeated calculation of convolutional features. The speed of SPPNet is more than 20 times that of R-CNN without sacrificing any detection accuracy.

However, although SPPNet effectively improves the detection speed, there are still some shortcomings. Firstly, the training is still multi-stage; Secondly, SPPNet only fine-tunes its fully connected layer, ignoring the previous convolutional layer.

(c) Fast RCNN

In 2015, Girshick proposed Fast RCNN [22], which is a further improvement of R-CNN and SPPNet [20] [23]. Fast RCNN enables us to train detectors and the boundary box regressors simultaneously in the same network configuration. Due to the large time, memory consumption of RCNN algorithm and the slow test speed, the Fast RCNN algorithm is improved to solve these problems. Fast RCNN successfully integrates the advantages of R-CNN and SPPNet, but there is still room for improvement in the method of proposing candidate regions. Then, a question naturally arises: "Can we use the CNN model to generate object proposals?" Later, Faster R-CNN [24] answered this question.

(d) Faster RCNN

In 2015, Ren Shaoqing et al. proposed the Faster RCNN detector [24][25] shortly after the Fast RCNN was released. Faster RCNN is the first near real-time object detector based on deep learning and the first real end-to-end object detector. The greatest improvement is to propose a region candidate network, which makes the proposed candidate regions

almost free of extra computation. From R-CNN to Faster RCNN, most independent modules of object detection system (such as candidate regions, feature extraction, border regression, etc.) have been gradually integrated into a unified end-to-end learning framework. Although Faster RCNN breaks through the speed bottleneck of Fast RCNN, there is still computational redundancy in the subsequent detection phase. Various improvement measures were proposed later, including RFCN [26] and Light head RCNN [27].

(e) Feature Pyramid Networks

In 2017, Lin et al. proposed feature pyramid network [28] for small object detection based on Faster RCNN. Before using the feature pyramid structure, most detectors only detect at the top level of the network. Although the high-level feature map extracted by the convolutional neural network is conducive to category recognition, it is not good for object location. Therefore, a top-down architecture with horizontal connections is proposed in the feature pyramid network to fuse high-level semantic information and low-level location information. Therefore, the feature pyramid network shows great progress in detecting objects of various scales. FPN has now become a basic part of many of the latest detectors.

3. Dataset of small object detection

At present, the object detection algorithms based on deep learning are data-driven. However, in the general object detection dataset, the number of large and medium objects is far higher than that of small objects. For the research of small objects, a specific dataset for small objects is needed. At present, the datasets for small objects include public datasets and datasets applied in various fields, among which several commonly used datasets and their parameters are shown in Table 1.

Table 1. Common datasets and their parameters

dataset	years	unit	Class	Number of pictures
COCO	2014	Microsoft Corporation	91	328000
PASCAL VOC	2005	Pattern Analysis, Statical Modeling and Computational Learning	20	17125

Tiny Person	2019	Chinese Academy of Sciences University	1	1610
DOTA	2017	Wuhan University	15	2806
NWPU VHR-10	2014	Northwestern Polytechnical University	10	800
UCAS-AOD	2015	University of Science and Technology of China	2	2420
Tsinghua-Tencent100K	2016	Tsinghua University and Tencent Joint Laboratory	128	100000
RSOD-Dataset	2015	Wuhan University	4	976

3.1 Public dataset

COCO [3] is a common object detection dataset, including natural pictures and common object pictures in life. The background is more complicated, the number of objects is larger, and the size of object is smaller. For small object detection tasks, the current standard for measuring the quality of a model is more inclined to use the detection results on the COCO dataset.

As the benchmark dataset for object detection, PASCAL VOC [32] has 20 categories, including 11,530 images for training and evaluation. From 2005 to 2012, an image recognition competition will be held every year. The main purpose of the competition is to recognize some categories of objects in real scenes, which is a problem of supervised learning. The training set is given in the form of pictures with labels.

3.2 Application datasets in various fields

Tiny Person [33] is a dataset proposed by the University of Chinese Academy of Sciences, which only contains the category of person. The training set contains 794 images and the test set contains 816 images.

Tsinghua-Tencent100K is a large-scale traffic sign dataset [34], which provides 100,000 images and contains 30,000 traffic sign instances. The dataset is small enough to become a classic dataset for small object detection of traffic signs. The images in the dataset cover most of the lighting

and weather conditions, and all the images are manually labeled, which can be divided into three categories: prohibitions, instructions, and warning signs. Tsinghua-Tencent100K is the only public traffic sign dataset based on real scenes in China with high resolution and large number of images, which fill the gap of lack of benchmark datasets in China.

Aerial remote sensing images have scale diversity. The shooting height ranges from several hundred meters to nearly 10,000 meters. The ground objects are different in size even if they are of the same type. Most of the objects in the aerial remote sensing images are small objects, containing dozens or even several pixels. Therefore, most of the aerial remote sensing image object detection datasets are small object detection datasets. The datasets commonly used are as follows: DOTA [35] is a dataset for remote sensing image object detection, which contains 15 categories, a total of 2806 images, and the size of each image is not more than 4000×4000 pixels. The UCAS-AOD dataset [36] only contains two categories of vehicles and aircraft, including 7,482 aircraft samples and 7,114 vehicle samples. NWPU VHR-10 [37] is an aerospace remote sensing object detection dataset annotated by Northwestern Polytechnical University, including 10 categories such as aircraft, ships, vehicles. The dataset has a total of 800 images, including 650 images of objects and 150 images of background. The RSOD-Dataset dataset [38] contains four types of objects: aircraft, playground, overpass, and oil bucket. There are 446 images for aircraft, 189 images for playground, 176 images for overpass, and 165 images for oil buckets.

4. Challenges and research trends of small object detection

4.1 Challenges of small object detection

Most of the object detection algorithms based on deep learning are aimed at medium and large objects with a certain size or proportion, which is difficult to apply to small object detection in complex backgrounds. Specifically, small object detection has the following challenges:

(a) The underlying features lack semantic information. In the existing object detection models, the underlying features of backbone network are generally used to detect small objects, but the underlying features lack semantic information, which brings certain difficulties to the detection of small objects.

(b) There are few datasets for algorithm research, and the

object scale distribution is uneven. At present, the dataset (PASCAL VOC, COCO) widely used in the mainstream object detection algorithm has less training samples for small objects, which makes the small objects cannot be fully learned in the process of model training. In the COCO dataset, there are a large number of small objects, but the distribution is very uneven. Only 52.3% of the images contain small objects, which makes the trained model more suitable for medium and large objects. The generalization ability of small objects is weak and the detection accuracy is low.

(c) Related algorithm research lacks versatility, which is difficult to directly migrate to other fields. The main network of the existing object detection models are all trained on the classification datasets, but the scale distribution of the objects in the classification datasets is different from that in the detection datasets.

4.2 Research trends

At present, the core problem of the research on small object detection based on deep learning is how to improve the feature expression of small objects to make it contain rich semantic information, which is also the key to improve the performance of small object detection. Small object detection has important application value in real life, such as in the field of criminal investigation, small packages on the table, small pedestrians in the corner of surveillance videos, small signs on clothes are clues to solve the case; In agriculture, the identification of insect pests and diseases is convenient for real-time detection of crop health status, and early treatment can be achieved in time, so as to reduce the decline of crop yield and quality caused by diseases and insect pests, and contribute to the development of precision agriculture and intelligent agriculture. Therefore, small object detection has important research significance.

Compared with the detection performance of large and medium objects, there is still a big gap in the performance of small object detection. In the future, we can explore better and more comprehensive multi-scale feature fusion strategies to further improve the detection performance of small objects. Of course, there are many problems to be solved. This section points out the future research directions for the difficulties of small object research.

(a) The bottom layer of the networks is fully used to detect multi-scale. Because the underlying features of convolutional neural networks have weak semantic information and high resolution, and high-level features have strong semantic information and low resolution, the low-

level features are very important for detecting small objects. Through feature fusion, all scales of the convolutional neural network can have strong semantic information, and a more complete feature pyramid network can be established [39].

(b) Increase the amount of training data for small objects. Data expansion is a key component of training deep learning model. An optimized data expansion strategy can effectively improve the detection effect of small objects [40] [41]. At present, the research of small objects is mainly aimed at common object datasets, and the network cannot learn the characteristics of small objects in depth. Subsequent work can establish datasets for small object detection, which is convenient for fully learning the characteristics of small objects, and is more conducive to small object detection.

(c) Object detection in the complex environment is one of the core issues in the field of computer vision. Complex backgrounds include object ambiguity, overlapping of objects and object occlusion, etc. While the detection of small objects in the complex background is even more difficult, Zheng et al. [42] proposed a modified multi-category single-order detector (SSD) algorithm for the object occlusion, which is tested on the VOC and COCO datasets. The experimental results show that the detection effect of the improved method has been greatly improved. Subsequent work can be carried out in this aspect to ensure the accuracy and robustness of the detection in the complex environment.

References

- (1) KISANTAL M, WOJNA Z, MURAWSKI J, et al. Augmentation for Small Object Detection [DB/OL]. arXiv: 1902.07296v1, 2019.
- (2) Wq604887956. 小目标检测文章总结 [EB/OL]. <https://blog.csdn.net/wq604887956/article/details/83053927>, 2018-10-16/2019-09-02.
- (3) LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[A]. The 13th European Conference on Computer Vision (ECCV) [C]. Cham: Springer, 2014.740-755.
- (4) BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS-Improving object detection with one line of code[A]. 2017 IEEE International Conference on Computer Vision [C]. New York, USA: IEEE, 2017. 5561-5569.
- (5) LI Y H, CHEN Y T, WANG N Y, et al. Scale-aware Trident Networks for Object Detection [DB/OL]. arXiv: 1901.01892, 2019.
- (6) 王海屹. 基于特征融合的小物体目标检测研究与应用[D]. 北京:北京邮电大学, 2018.
- (7) EGGERT C, BREHM S, WINSCHER A, et al. A closer look: Small object detection in faster R-CNN [A]. 2017 IEEE International Conference on Multimedia and Expo (ICME) [C]. Hong Kong, China: IEEE Computer Society, 2017. 421-426.
- (8) 刘晓楠, 王正平, 贺云涛等. 基于深度学习的小目标检测研究综述[J]. 战术导弹技术, 2019,193 (01) : 106-113.
- (9) 于进勇, 丁鹏程, 王超. 卷积神经网络在目标检测中的应用综述[J]. 计算机科学, 2018, 45(S2): 27-36.
- (10) HARIHARAN B, Arbelaez P, GIRSHICK R, et al. Hyper-columns for object segmentation and fine-grained localization[A]. 2014 IEEE Conference on Computer Vision and Pattern Recognition[C]. Boston, MA, USA : IEEE, 2014. 447-456.
- (11) 李旭冬, 叶茂, 李涛. 基于卷积神经网络的目标检测研究综述[J]. 计算机应用研究, 2017,34(10):2881-2886.
- (12) Uijlings J R R, Van De Sande K E A, Gevers T, et al. Selective search for object recognition[J]. International journal of computer vision, 2013,104(2) : 154-171.
- (13) J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE conference on

- computer vision and pattern recognition, 2016, pp. 779-788.
- (14) J. Redmon and A. Farhadi, "Yolo9000 : better, faster, stronger," arXiv preprint, 2017.
 - (15) —, "Yolov3 : An incremental improvement," arXiv preprint arXiv: 1804.02767, 2018.
 - (16) W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in European conference on computer vision. Springer, 2016, pp. 21-37.
 - (17) T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," IEEE transactions on pattern analysis and machine intelligence, 2018.
 - (18) Fung G, Mangasarian O L. Incremental support vector machine classification[C]// Proceedings of the 2002 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2002 : 247-260.
 - (19) Neubeck A, Van Gool L. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR'06). IEEE, 2006,3:850-855.
 - (20) K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in European conference on computer vision. Springer, 2014, pp. 346-361.
 - (21) Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]// Advances in neural information processing systems. 2012: 1097-1105.
 - (22) R. Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440-1448.
 - (23) R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580-587.
 - (24) S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91-99.
 - (25) S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis & Machine Intelligence, no.6, pp. 1137-1149, 2017.
 - (26) J. Dai, Y. Li, K. He, and J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," in Advances in neural information processing systems, 2016, pp. 379-387.
 - (27) Z. Li, C. Peng, G. Yu, X. Zhang, Y. Deng, and J. Sun, "Light-head r-cnn: In defense of two-stage object detection," arXiv Preprint arXiv: 1711.07264, 2017.
 - (28) Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE conference on computer vision and pattern recognition.2017: 2117-2125.
 - (29) H K M, Gkioxari G, Dollar P, et al. Mask R-CNN[C] // Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017:2980-2988.
 - (30) 杨浩琪, 姚红革, 王诚, 等. 复杂背景下遥感船舶的检测与分割定位[J]. 计算机辅助设计与图形学学报, 2020, 32(3):9-14.

- (31) HU P Y, RAMANAN D. Finding tiny faces[A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition[C]. Honolulu, HI: IEEE, 2017. 1522-1530.
- (32) EVERINGHAM M, WINN J. The PASCAL visual object classes challenge 2007 (VOC2007) development kit[J]. International Journal of Computer Vision , 2006,111(1): 98-136.
- (33) YU X, GHONG Y, JIANG N, et al. Scale match for tiny person detection[C]//Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2020:1246-1254.
- (34) ZHU Z, LIANG D, ZHANG S H, et al. Traffic-sign detection and classification in the wild[A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]. Las Vegas, NV, USA: IEEE Computer Society, 2016. 2110-2118.
- (35) XIA G S, BAI X, DING J, et al. DOTA: A Large-scale Dataset for Object Detection in Aerial Images[DB/OL]. arXiv: 1711. 10398, 2017.
- (36) ZHU H G, CHEN X G, DAI W Q, et al. Orientation robust object detection in aerial images using deep convolutional neural network[A]. 2015 IEEE International Conference on Image Processing (ICIP)[C]. Quebec City, QC, Canada: IEEE, 2015. 3735-3739.
- (37) CHENG G, HAN J W, ZHOU P C, et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 98(1): 119-132.
- (38) XIAO Z F, LIU Q, TANG G F, et al. Elliptic fourier transformation-based histograms of oriented gradients for rotationally invariant object detection in remote-sensing images[J]. International Journal of Remote Sensing, 2015, 36(2): 618-644.
- (39) SHEN Z Q, SHI H H, YU J H, et al. Learning Object Detectors from Scratch with Gated Recurrent Feature Pyramid [DB/OL]. arXiv: 1712. 00886, 2017.
- (40) ZOPH B, CUBUK E D, GHIASI G, et al. Learning Data Augmentation Strategies for Object Detection [DB/OL]. arXiv: 1906. 11172, 2019.
- (41) KHOREVA A, BENENSON R, ILG E, et al. Lucid data dreaming for video object segmentation[J]. International Journal of Computer Vision, 2019,127(9): 1175-1197.
- (42) 郑浦, 白宏阳, 李伟, 等. 复杂背景下的小目标检测算法[J]. 浙江大学学报, 2020, 54(9): 5-8.
- (43) 刘颖, 刘红艳, 范九伦, 等. 基于深度学习的小目标检测研究与应用综述[J]. 电子学报, 2020, 48(3): 590-601.