A Small Target Detection Research Based on Dynamic Convolution Neural Network

Chenshuai Bai

Lanzhou Jiaotong University

Kaijun Wu (✉ wkj@mail.lzjtu.cn)

Lanzhou Jiaotong University

Dicong Wang

Lanzhou Jiaotong University

Hong Li

Xianyang Normal University

Mingjun Yan

Lanzhou Jiaotong University

Tao Huang

Lanzhou Jiaotong University

Research Article

Keywords: small Target detection, Dynamic convolution neural network, optimization parameters, RSOD remote sensing image, TGRS-HRRSD remote sensing image

Posted Date: September 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-916747/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A small target detection research based on dynamic convolution neural network

Chenshuai Bai¹, Kaijun Wu^{1*}, Dicong Wang^{1,2}, Hong Li³, Mingjun Yan¹, Tao Huang¹

¹School of Electronic and Information Engineering of Lanzhou Jiaotong University, China

²College of Intelligence and Computing of Tianjin University, China

³School of computer science of Xianyang Normal University, China

*Corresponding author : Kaijun Wu(E-mail: wkj@mail.lzjtu.cn)

ABSTRACT

Because the detection effect of EfficientNet-YOLOv3 target detection algorithm is not very good, this paper proposes a small target detection research based on dynamic convolution neural network. Firstly, the dynamic convolutional neural network is introduced to replace the traditional convolutional neural network, which makes the algorithm model more robust; Secondly, in the training process, the optimization parameters are continuously adjusted to further strengthen the model structure; Finally, in order to prevent over fitting, the Learning Rate and Batch Size parameters are modified during the training process. remote sensing image The results of the proposed algorithm on RSOD remote sensing image data sets show that compared with the original EfficientNet-YOLOv3 algorithm, the (Average Precision, AP) value is increased by 1.93% and the (Log Average Miss Rate ,LAMR) value is reduced by 0.0500; The results of the proposed algorithm on TGRS-HRRSD remote sensing image data set show that compared with the original EfficientNet-YOLOv3 algorithm, the mAP value is increased by 0.07% and the mLAMR value is reduced by 0.0007.

Keywords: small Target detection; Dynamic convolution neural network; optimization parameters; RSOD remote sensing image; TGRS-HRRSD remote sensing image

1.Introduction

The emergence of deep learning technology promotes the development of computer vision. Target detection based on remote sensing image is one of the important directions in the field of computer vision. Target detection based on remote sensing image can be applied to urban planning, traffic safety and environmental monitoring ^[1-5]. The target detection algorithm of remote sensing image comes from the target detection algorithm in the general field. The target detection algorithm of remote sensing image is a series of algorithms combined with the characteristics of multi-scale, strong density and shape difference of remote sensing image. The traditional remote sensing image target detection algorithm mainly adopts multi-step strategy, such as preprocessing, image segmentation, extracting regions of interest and target detection ^[6-7]. Each step is independent of each other, and each is the corresponding algorithm designed for specific problems, which has the problems of insufficient process solidification, automation and intelligence^[8]. Therefore, due to the strong feature representation ability and end-to-end learning ability, the performance of target detection algorithm in remote sensing image has been greatly improved.

Nowadays, remote sensing image is a data set with small target and complex background. There are huge problems in the research of remote sensing image target detection. Although many target detection algorithms have relatively good results on ordinary data sets, the detection accuracy of target detection in remote sensing images is not very good. Chen et al. ^[9] proposed a detection algorithm based on multi classification learning, which realizes the accurate detection of ships by finding the minimum circumscribed rectangle of the target; Ding et al. ^[10] proposed the region of interest deformer (ROI trans) to solve the remote sensing target detection with significant directivity; Xu et al. ^[11] proposed the gliding vertex to characterize the directed bounding box by first detecting the horizontal bounding box and then learning the offset of the four corners of the horizontal bounding box; Yang et al. ^[12] proposed dense rotating small target detector. By designing sampling fusion network, multi-layer features are fused into effective anchor sampling to improve the sensitivity to small targets; At the same time, by suppressing noise and highlighting object characteristics, a supervised pixel attention network and channel attention network are designed; Finally, the (Intersection Over Union ,IOU) factor is added to smooth L1 loss to solve the problem of inaccurate boundary of rotating bounding box; Aiming at the problems of large aspect ratio and category imbalance of target instances in remote sensing images; Lin et al. ^[13] proposed a fine tuned single-stage detector, which realizes feature reconstruction and alignment by designing feature refinement module, and realizes target detection by stepwise regression from coarse-grained to fine-grained; Yao et al. ^[14] proposed a target detection algorithm of remote

sensing image based on multi-resolution feature fusion; Wang et al. [15] proposed the method of combining multi-level features for regional proposal; Yao et al. [16] proposed a multi-scale convolutional neural network remote sensing target detection framework; Li et al. [17] proposed a target detection algorithm for visual perception of high-resolution remote sensing images; Chen et al. [18] proposed a fast ship target detection method in large scene remote sensing images based on cascaded convolutional neural network; Zhao et al. [19] aiming at the problems such as poor target performance and of and slow inference speed of remote sensing images with dense permutation and significant direction detected by RoI Trans, adopted the method of progressively enhancing the positioning accuracy of rotating candidate frame and non-local enhancement of features to improve the target detection performance of remote sensing images; Cai et al. [20] proposed a new progressive shrinkage algorithm, a generalized pruning method, which reduces the model size in more dimensions than pruning (depth, width, core size and resolution); Chen et al. [21] proposed an algorithm of dynamic convolution neural network; Liu et al. [22] proposed a new strategy to assemble multiple identical backbones through the composite connection between adjacent backbones to form a more powerful backbone, which is called composite backbone network; Hua et al. [23] proposed a cascade convolutional neural network real-time target detection framework combining visual perception and convolutional memory network reasoning; Li et al. [24] proposed an end-to-end depth network based on the shape of network structure, which detects salient objects in optical images in a pure data-driven manner; Zhang et al. [25] proposed a scale adaptive network scheme to improve the accuracy of multi-target detection; Zhu et al. [26] proposed a new target detection method based on multi-scale Selu densenet and (Dynamic Anchor Assignment, DAA) strategy; Yao et al. [27] proposed an online latent semantic scatter method to solve this problem; Pang [28] proposed a unified self enhancement network R^2 -CNN: convolutional neural network based on remote sensing region, which is composed of backbone microgrid, intermediate global attention block, final classifier and detector; Chen et al. [29] proposed a new target heat mAP network to solve these problems; Yao et al. [30] proposed a new calculation model for airport detection using optical (Remote Sensing Image, RSI); Yu et al. [31] proposed a method based on image decomposition and distance regularization (Chan Vese, CV) model; Li et al. [32] proposed the remaining problems and future development trend of ship detection and classification methods based on optical remote sensing images; Youme et al. [33] proposed to divide the image into four regions, mark the image through the region of interest, and extract image features by single-stage target detection algorithm SSD to improve the accuracy of small target detection.

This paper improves the Efficientnet-YOLOv3 algorithm [34]. The backbone feature extraction network used by Efficientnet-YOLOv3 algorithm is EfficientNet [35]. Firstly, based on the idea of dynamic convolution neural network proposed by Chen et al. [21], the dynamic convolution neural network is used to replace the two-dimensional convolution neural network of the backbone feature extraction network, Dynamic convolution kernel dynamically aggregates multiple parallel convolution kernels, which are related to input, improves the backbone feature extraction network and strengthens the feature extraction ability of the backbone network for remote sensing images; Secondly, adjust and optimize the parameters to improve the feature extraction ability of small and medium-sized targets in remote sensing images; Finally, during the training, the learning rate and batch size were modified to prevent overfitting. The validity of the proposed algorithm is verified by using RSOD and TGRS-HRRSD remote sensing image data sets.

2. EfficientNet-YOLOv3 algorithm model

2.1. EfficientNet backbone feature extraction network

EfficientNet backbone feature extraction network is a research under the same computing power constraints. The influence of network depth, width and resolution on the same operation type network is explored to find the optimal configuration proportion parameters. EfficientNet algorithm process of backbone feature extraction network: firstly, the image needs to be transformed into the input dimension required by the (Mobile inverted Bottleneck Convolution, MBConv) module through the first Conv 3×3 convolution layer, and the input remote sensing image needs to be extracted from the feature map through a series of MBConv modules; Secondly, the parameters of each MBConv module are finely adjusted to adapt to the current use environment. The combined scale optimization method can make the network obtain a better receptive field; Finally, using the feature map adaptive connection method based on (Full Convolutional Neural Network, FCNN), the Conv 1×1 convolution network will be able to adapt to the feature maps of different sizes and unify them into the dimensions required by the algorithm. Finally, the image classification, recognition and detection are completed through the output feature map. The structure of EfficientNet backbone characteristic network model is shown in Figure 1:

the increase in the output dimension of each layer is within a controllable range. Traditional convolutional neural network is easily replaced by dynamic convolutional neural network. The traditional convolution neural network is improved. In the traditional convolution neural network, the parameters of any input are fixed, and the dynamic convolution makes the convolution parameters of each layer change with the input during reasoning. For a feature map generated in the convolution process, first operate the feature map to generate K parameters π with a sum of 1. Then the K convolution kernel parameters are linearly summed. In this way, the convolution changes with the change of input.

Figure 3.Dynamic convolutional neural network

3.2. Improved YOLOv3 model structure

As shown in Figure 4, firstly, the input image size of the backbone feature extraction network EfficientNet is modified from $416 \times 416 \times 3$ to $800 \times 800 \times 3$, the residual network is convoluted once with a convolution kernel size of 3×3 and a step size of 2, and the width and height of the input feature layer are compressed to obtain a feature layer; Secondly, carry out a 1×1 convolution and a 3×3 convolution on the feature layer, and then add this result to the feature layer to form a residual structure; Finally, the network is greatly deepened through the superposition of 1×1 convolution, 3×3 convolution and residual edge. The two-dimensional convolution neural network Conv2D uses the Dynamic convolution neural network Dynamic_Conv2D.

Figure4. Improved YOLOv3 algorithm model

4. Experimental results and analysis

In order to verify the effectiveness of the improved EfficientNet-YOLOv3 algorithm, the RSOD remote sensing image data set labeled by Wuhan University and the TGRS-HRRSD remote sensing image data set labeled by Xi'an Institute of Optics and precision machinery, Chinese Academy of sciences are used to verify the effect. The experiment uses five comparison algorithms: target detection algorithm based on Faster-RCNN, target detection algorithm based on RetinaNet, single-stage target detection algorithm based on SSD, target detection algorithm based on YOLOv4 and target detection algorithm based on the original EfficientNet-YOLOv3. Objectively analyze the performance of the target detection algorithm proposed in this paper, and use the (Average Precision , AP) and (Log Average Missed detection Rate , LAMR)

as the evaluation indexes. The larger the mAP value, the better, and the smaller the mLAMR value of the better.

4.1. Experimental environment configuration

The target detection algorithm proposed in this paper uses the built experimental environment platform model: the computer is configured with i5-8250CPU, 8GRAM, 64 bit Windows10 operating system, the server is configured with BSCC-N22, and the GPU queue is configured as follows: each machine is configured with 8 GPUs with NVIDIA Tesla V100-sxm2 32GB video memory, each GPU card is allocated 8CPU core and 36GB memory by default, and the memory ratio is 1GPU card, 8CPU core 36GB memory.

4.2.Experimental results of RSOD remote sensing image data set

4.2.1.RSOD remote sensing image data set

As shown in Figure 5, the RSOD remote sensing image data set is an open data set, which is applied to target detection of remote sensing images. The sample data set diagrams include: Figure 5(a) aircraft, Figure 5(b) oilbank, Figure 5(c) bypass and Figure 5(d) playground. The dataset is formatted using PASCAL_VOC format, the data set includes four files, each of which is an object. 4993 aircraft in 446 images, 191 amusement parks in 189 images, 180 overpasses in 176 images and 1586 oil barrels in 165 images.

Figure 5.Sample of RSOD remote sensing image data

4.2.2. mAP value as evaluation index

For the target detection algorithm of remote sensing image, the detection accuracy of the algorithm is very important. This time, the (Average Precision , mAP) is selected as the evaluation index. mAP is the average after adding the AP values of all classes.

	Efficientnet-YOLOv3 method	Faster-RCNN method	RetinaNet method	SSD method	YOLOv4 method	the method of this paper
aircraft(AP value)	0.85	0.33	0.72	0.70	0.97	0.95
overpass(AP value)	0.93	0.83	1.00	0.96	0.75	0.92
playground(AP	1.00	0.99	1.00	1.00	0.93	1.00

value)						
oiltank(AP	1.00	0.86	1.00	0.99	1.00	0.99
value)						
mAP value	0.9461	0.7536	0.9304	0.9122	0.9233	0.9654

Table 1. Experimental results of map value of RSOD remote sensing image data set

As shown in Table 1, the experimental results AP values of numerical indexes of each class obtained by all algorithms in this paper are shown, and the last row is the mAP values of all classes in all methods. It can be seen from the mAP value in the last line that the mAP value obtained by the proposed algorithm is the highest compared with the five compared algorithms. Among them, the mAP value of the proposed algorithm is 1.93% higher than that of the original algorithm EfficientNet-YOLOv3, the mAP value of the proposed algorithm is 21.18% higher than that of the Faster-RCNN algorithm, and the proposed algorithm is 21.18% higher than that of the RetinaNet algorithm, The mAP value is increased by 3.50%. Compared with SSD algorithm, the mAP value of this method is increased by 5.32%. Compared with YOLOv4 algorithm, the mAP value of this algorithm is increased by 4.21%, which shows that the algorithm proposed in this paper has achieved good results in the target detection process of remote sensing image.

4.2.3. mLAMR value is used as evaluation index

The target detection algorithm for remote sensing images generally uses (Miss Rate , MR). The logarithmic mean is used as the evaluation standard of data, which is called LAMR for short.

	Efficientnet-YOLOv3 method	Faster-RCNN method	RetinaNet method	SSD method	YOLOv4 method	the method of this paper
Aircraft(LAMR value)	0.52	0.88	0.51	0.72	0.16	0.25
overpass(LAMR value)	0.11	0.37	0.00	0.04	0.32	0.17
playground(LAMR value)	0.00	0.00	0.00	0.00	0.07	0.00
oiltank(LAMR value)	0.00	0.20	0.01	0.13	0.00	0.01
mLAMR value	0.1575	0.3625	0.1300	0.2225	0.1375	0.1075

Table 2. Experimental results of mLAMR value of RSOD remote sensing image data set

Table 2 shows the experimental results of the numerical index LAMR of each class obtained by all algorithms in this paper. The last line is the mLAMR value obtained by adding and averaging the LAMR values of all classes in all algorithms. It can be seen from the mLAMR value that the mLAMR value obtained by the algorithm proposed in this paper is the lowest compared with the other five algorithms. Among them, the mLAMR value of the algorithm proposed in this paper is 0.0500 lower than that of the original algorithm EfficientNet-YOLOv3, 0.2550 lower than that of Faster-RCNN target detection algorithm, and 0.0225 lower than that of RetinaNet target detection algorithm, Compared with SSD target detection algorithm, mLAMR value is reduced by 0.1150, and compared with YOLOv4 target detection algorithm, mLAMR value is reduced by 0.0300. Again, it shows that the target detection algorithm proposed in this paper has achieved good results in the process of target detection.

4.3.Experimental results of TGRS-HRRSD remote sensing image data set

4.3.1. TGRS-HRRSD remote sensing image data set

As shown in Figure 6, TGRS-HRRSD remote sensing image data set is a remote sensing image data set produced by the optical image analysis and learning center of Xi'an Institute of Optics and precision machinery, Chinese Academy of Sciences for studying high-resolution remote sensing image target detection. The format of the data set is made into PASCAL_VOC format application, the sample images of remote sensing image data include: Figure 6(a) airplane, Figure 6(b) bridge, Figure 6(c) crossroad, Figure 6(d) ship, Figure 6(e) vehicle, Figure 6(f) harbor, Figure 6(g) ground track field, Figure 6(h) storage tank, Figure 6(i) basketball court, Figure 6(j) parking lot, Figure 6(k) tennis court Figure 6(l) baseball diamond and Figure 6(m) T junction have 13 categories, 21000 diagrams and 40000 target objects in total.

Figure 6. TGRS-HRRSD remote sensing image data sample diagram

4.3.2. mAP value as evaluation index

For the target detection algorithm of remote sensing image, the effectiveness of the algorithm is closely related to the detection accuracy of the algorithm. This time, the average accuracy (mAP) is selected as the evaluation index. mAP is the average of all class AP values.

	Efficientnet- YOLOv3 method	Faster-RCNN method	RetinaNet method	SSD method	YOLOv4 method	the method of this paper
Ground track field(AP value)	0.99	0.94	1.00	0.99	0.97	0.98
crossroad(AP value)	0.98	0.87	1.00	0.98	0.93	0.98
airplane(AP value)	1.00	0.87	0.98	0.99	0.98	0.99
Baseball	0.99	0.77	0.98	0.95	0.87	0.85

diamond(AP value)						
harbor(AP value)	0.99	0.79	0.98	0.97	0.77	0.99
Storage tank(AP value)	0.98	0.86	0.98	0.95	0.60	0.98
T junction(AP value)	0.96	0.30	0.97	0.92	0.81	0.92
bridge(AP value)	0.98	0.75	0.95	0.97	0.88	1.00
Tennis court(AP value)	0.97	0.61	0.95	0.93	0.97	0.95
ship(AP value)	0.98	0.74	0.94	0.93	0.94	0.99
Basketball court(AP value)	0.99	0.17	0.85	0.75	0.74	0.85
vhicle(AP value)	0.96	0.26	0.80	0.32	0.96	0.99
Parking lot(AP value)	0.85	0.38	0.77	0.70	0.71	0.85
mAP value	0.9569	0.6381	0.9340	0.8722	0.8571	0.9576

Table 3. Experimental results of map value of TGRS-HRRSD remote sensing image data set

As shown in Table 3, the numerical indexes of all classes obtained by all algorithms in this paper are the AP values obtained from the experimental results. The last line is the mAP value obtained by adding and averaging the AP values of all classes in all methods. It can be seen from the mAP value in the last line that the mAP value obtained by the algorithm in this paper is the highest compared with the five algorithms. Among them, the mAP value of the algorithm proposed in this paper is 0.07% higher than that of the original EfficientNet-YOLOv3 target detection algorithm, 31.95% higher than that of Faster-RCNN target detection algorithm, and 2.36% higher than that of RetinaNet target detection algorithm, Compared with SSD target detection algorithm, the mAP value is increased by 8.54%, and compared with YOLOv4 target detection algorithm, the mAP value is increased by 10.05%, which further shows that the algorithm proposed in this paper has a good effect in the target detection process of remote sensing images.

4.3.3. mLAMR value is used as evaluation index

The target detection algorithm of remote sensing image generally uses the (Logarithmic Average of Miss Rate , LAMR) as the evaluation standard of data, which is called LAMR for short.

	Efficientnet-YOLOv3 method	Faster-RCNN method	RetinaNet method	SSD method	YOLOv4 method	the method of this paper
Ground track field(LAMR value)	0.01	0.10	0.01	0.02	0.03	0.02
crossroad(LAMR value)	0.03	0.29	0.01	0.06	0.11	0.04
airplane(LAMR value)	0.01	0.30	0.03	0.05	0.03	0.01
Baseball diamond(LAMR value)	0.03	0.48	0.04	0.43	0.24	0.03
harbor(LAMR value)	0.02	0.52	0.09	0.14	0.31	0.04
Storage tank(AP value)	0.10	0.29	0.11	0.28	0.59	0.08
T junction(LAMR value)	0.12	0.88	0.09	0.21	0.38	0.23
bridge(LAMR value)	0.02	0.38	0.07	0.05	0.23	0.00
Tennis court(LAMR value)	0.11	0.54	0.12	0.22	0.07	0.11
ship(A LAMR	0.04	0.42	0.13	0.16	0.11	0.04

value)						
Basketball						
court(LAMR	0.39	0.89	0.28	0.43	0.40	0.39
value)						
vhicle(LAMR	0.18	0.91	0.38	0.89	0.12	0.06
value)						
Parking lot(LAMR	0.41	0.80	0.49	0.61	0.53	0.41
value)						
m LAMR value	0.1130	0.5231	0.1423	0.2731	0.2423	0.1123

Table 4. Experimental results of mLAMR value of TGRS-HRRSD remote sensing image data set

Table 4 shows the experimental results of the numerical index LAMR of all classes obtained by all algorithms in this paper. The last line is to add and average the LAMR values of all classes in all algorithms to obtain the mLAMR value. It can be seen from the mLAMR value that the mLAMR value obtained by the target detection algorithm proposed in this paper is the lowest compared with the five compared algorithms. Among them, the mLAMR value of the target detection algorithm proposed in this paper is reduced by 0.0007 compared with the original EfficientNet-YOLOv3 algorithm, and the mLAMR value is reduced by 0.4108 compared with the Faster-RCNN target detection algorithm, The mLAMR value is reduced by 0.0300. Compared with SSD target detection algorithm, the mLAMR value is reduced by 0.1608. Compared with YOLOv4 target detection algorithm, the mLAMR value is reduced by 0.1300. Again, it shows that the algorithm proposed in this paper obtains good results in the process of target detection.

5. Conclusion and Prospect

This paper presents the research on small target detection based on dynamic convolution neural network, and compares the target detection algorithm with EfficientNet-YOLOv3 algorithm. Without adding any model parameters, it not only improves the mAP value of target detection in remote sensing image, but also reduces the mLAMR value of target detection in remote sensing image. It improves the accuracy of the detection process, reduces the missed detection rate of the algorithm, and enhances the stability of the algorithm. The target detection algorithm proposed in this paper has achieved good results in RSOD and TGRS-HRRSD remote sensing image data sets, but the algorithm has two disadvantages in the face of target detection in remote sensing images. Firstly, the algorithm in this paper is modified on the basis of EfficientNet-YOLOv3 target detection algorithm, and uses dynamic convolutional neural network to replace the traditional convolutional neural network, which improves the complexity of the network model; Secondly, the algorithm proposed in this paper is only for target detection in remote sensing images, which is not necessarily very good when using other data sets and has weak universality. It is hoped that in further work, the use of lightweight backbone feature extraction network will be considered to reduce the complexity of network model, minimize the parameters of model and improve the detection performance of target detection algorithm in remote sensing image.

Reference

- [1] Han J W, Zhang D W, Cheng G, et al. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning[J].IEEE Transactions on Geoscience and Remote Sensing,2015,53(6):3325-3337.
- [2] Cheng G, Han J W. A survey on object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2016,117:11-28.
- [3] 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,2014,98:119-132.
- [4] Deng Z P, Sun H, Zhou S L, et al. Multi-scale object detection in remote sensing imagery with convolutional neural networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2018,145:3-22.
- [5] Zhong Y F, Han X B, Zhang L P. Multi-class geospatial object detection based on a position sensitive balancing framework for high spatial resolution remote sensing imagery[J].ISPRS Journal of Photogrammetry and Remote Sensing,2018,138:281-294.
- [6] Liu J K, Li J B, Ma L, et al. Radar target detection method of aircraft wake vortices based on matrix information geometry[J]. Journal of Radars, 2017, 6(6): 699-708. doi: 10.12000/JR17058.
- [7] Chen X L, Guan J, Huang Y, et al. Radar refined processing and its applications for low-observable moving target[J]. Science & Technology Review, 2017, 35(20): 19-27.
- [8] Su N Y, Chen X L, Guan J, et al. Detection and classification of maritime target with micromotion based on CNNs[J].Journal of Radars, 2018, 7(5):565-574.doi:10.12000/JR18077.
- [9] Chen L Q, Shi W X, Fan C E, et al. Ship detection in optical remote sensing images based on multi-class learning[J]. Journal of Huazhong University of science and Technology, 2019,47(5):62-67.
- [10] Ding J, Nan X, Yang L, et al. Learning RoI trans-former for oriented object detection in aerial images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles:IEEE, 2019:2849-2858.

- [11] Xu Y, Fu M, Wang Q, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 41(4):1452-1459.
- [12] Yang X, Yang J, Yan J, et al. ScrDet: Towards more robust detection for small, cluttered and rotated objects[C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul: IEEE, 2019:8232-8241.
- [13] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice:IEEE, 2017:2980-2988.
- [14] Yao Y Q, Cheng G, Xie X X, et al. Optical remote sensing image object detection based on multi-resolution feature fusion[J]. Journal of remote sensing, 2021,25(05):1124-1137.
- [15] Wang L W, Feng Y Q, Zhang Y B. Target detection method for optical remote sensing imagery[J]. Systems engineering and electronic technology, 2019,41(10):2163-2169.
- [16] Yao Q L, Hu X, Lei H. Object Detection in Remote Sensing Images Using Multiscale Convolutional Neural Networks[J]. Journal of Optics, 2019,39(11):346-353.
- [17] Li C, Zhang Y C, Lan T, et al. An Object Detection Algorithm with Visual Perception for High-Resolution Remote Sensing Images[J]. Journal of Xi'an Jiaotong University, 2018,52(06):9-16.
- [18] Chen H Y, Liu Z Y, Guo W W, et al. Fast Detection of Ship Targets for Large-scale Remote Sensing Image Based on a Cascade Convolutional Neural Network[J]. Journal of radar, 2019,8(03):413-424.
- [19] Zhao T Z, Yang C W, Liu W. Remote sensing image object detection method based on non-local feature enhancement[J]. Journal of Huazhong University of science and Technology, 2021:1-6.
- [20] Cai H, Gan C, Wang T Z, et al. Once-for-all: Train one network and specialize it for efficient deployment[J]. computer Science, 2020, pp:1-15.
- [21] Chen Y P, Dai X Y, Liu M C, et al. Dynamic Convolution: Attention over Convolution Kernels[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp:11030-11039.
- [22] Liu, Y, Wang, Y, Wang, S, et al. CBNNet: A Novel Composite Backbone Network Architecture for Object Detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020,34(07), 11653-11660.
- [23] Hua X, Wang X Q, Rui T, et al. A fast self-attention cascaded network for object detection in large scene remote sensing images[J]. Applied Soft Computing Journal, 2020, pp:1-20.
- [24] Li C, Cong R, Hou J, et al. Nested Network With Two-Stream Pyramid for Salient Object Detection in Optical Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019,57(11), pp: 9156-9166.
- [25] Zhang S, He G, Chen H, et al. Scale Adaptive Proposal Network for Object Detection in Remote Sensing Images[J]. IEEE Geoscience and Remote Sensing Letters, 2019,16(6), pp:864-868.
- [26] Zhu H M, Zhang P, Wang L B, et al. A multiscale object detection approach for remote sensing images based on MSE-DenseNet and the dynamic anchor assignment[J]. Remote Sensing Letters, 2019,10(10), pp:959-967.
- [27] Yao T, Wang G, Yan L S, et al. Online latent semantic hashing for cross-media retrieval[J]. Pattern Recognition, 2019,89, pp:1-11.
- [28] Pang J, Li C, Shi J, et al. R²-CNN: Fast Tiny Object Detection in Large-Scale Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019,57(8), pp:5512-5524.
- [29] Chen H, Zhang L B, Ma J, et al. Target heat-map network: An end-to-end deep network for target detection in remote sensing images[J]. Neuro Computing, 2019,331, pp:375-387.
- [30] Yao X W, Han J W, Guo L, et al. A coarse-to-fine model for airport detection from remote sensing images using target-oriented visual saliency and CRF[J]. Neuro Computing, 2015,164, pp:162-172.
- [31] Song Y, Wu Y Q, Dai Y M. Automatic river target detection from remote sensing images based on image decomposition and distance regularized CV model[J]. Computers and Electrical Engineering, 2016, pp:1-11.
- [32] Li B, Xie X Y, Wei X X, et al. Ship detection and classification from optical remote sensing images: A survey[J]. 2021,26(3), pp:145-163.
- [33] Youme O, Bayet T, Dembele J M, et al. Deep learning and remote sensing : detection of dumping waste using UAV[J]. Procedia Computer Science, 2021,185, pp:361-369.
- [34] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. Computer Science, 2018, arXiv preprint arXiv:1804.02767, 2018,1-6.
- [35] Tian M X, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[C]// Proceedings of the 36th International Conference on Machine Learning, 2019,6105-6114.