

PAPER • OPEN ACCESS

YOLO-RD: A lightweight object detection network for range doppler radar images

To cite this article: Long Zhou *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **563** 042027

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

YOLO-RD: A lightweight object detection network for range doppler radar images

ZHOU Long^{1,2}, WEI Suyuan¹, CUI Zhongma², DING Wei^{2,3}

¹The Rocket Force University of Engineering, Xi'an 710025, China;

²Beijing Institute of Remote Sensing Equipment, Beijing 100854, China;

³University of Electronic Science and Technology of China, Chengdu 611731, China

Email:1792563879@qq.com

Abstract. Under the condition of limited memory and computing power of radar aircraft equipment, large-scale object detection network based deep learning can not be deployed. Based on the darknet framework, Our paper proposes a lightweight object detection network for range doppler(RD) radar images: YOLO-RD, and builds a lightweight RD dataset: mini-RD, for efficient network training. Firstly, YOLO-RD extracts features from the input image through a series of small convolutional. Secondly, the dense block connection module is used to design the backbone extraction network. Finally, the prediction layer is combined with multi-scale features for prediction. Experiments show that YOLO-RD has achieved good results on the mini-RD dataset with a smaller memory budget, with a detection accuracy of 97.54%.

1. Introduction

In the field of military reconnaissance and target strike, the detection of sea ship has always been an important research direction. Traditional detection algorithms such as the classical Cfar have always had low detection accuracy. In recent years, with the rise of deep learning, the method based on convolutional neural network(CNN) extraction features has attracted much attention in the field of target detection with its high detection accuracy. However, on some embedded devices, large target detection models cannot be run due to limited memory and computing power. Based on this, some lightweight networks applied to embedded devices have attracted the attention of researchers. Some networks proposed such as SqueezeNet[1], MobileNet[2], and ShuffleNet[3], which aim to achieve the same effect as deep networks through simple and efficient network design. The lightweight CNN is mounted on the target detection framework such as ssd[4] and YOLO[5] to form a lightweight target detection network such as MobileNet-SSD[6] and tiny-YOLOv3[7] for deployment on the mobile side. Compared with the mainstream model, this kind of network takes up less memory and has faster detection speed, but the detection accuracy is lower. For specific engineering applications, the impact of speed and accuracy needs to be weighed against mission requirements.

The RD image detected in this paper is the imaging of the radar in the doppler domain. The image of the radar carried on the missile carrier is the RD image. The RD image has been widely studied in the field of radar image detection. For the RD radar image target detection, the author has achieved good results based on the improved YOLOv3 method in the literature[8]. However, because the weight file occupies a large memory, YOLOv3 cannot be applied to deploying aircraft equipment. In this regard, our paper proposes a lightweight object detection network for RD radar images:



YOLO-RD, which draws on the DenseNet [9] network structure and uses a dense connection module design network. The model occupies less memory and the detection accuracy and speed meet the application requirements. For the redundancy problem of RD dataset, a lightweight RD dataset was constructed: mini-RD. Finally, the high performance of YOLO-RD was verified by a large number of experimental comparisons.

2. Related work

2.1 YOLO

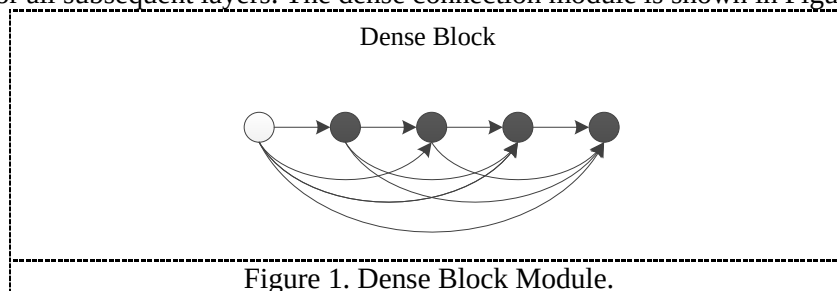
The YOLO series network has achieved end-to-end training. The latest YOLOv3 achieves 57.9% mAP on the COCO dataset, which is one of the best object detection models. The backbone of YOLOv3 uses darknet-53, which uses the residual network structure to effectively reduce the disappearance of the gradient, and uses a large number of 3*3 and 1*1 convolution kernel extraction features. YOLOv3's loss function integrates coordinate error, confidence error and classification error by means of mean square and error. The mathematical abstract expression is as follows:

$$\text{loss} = \sum_{i=0}^{s^2} \text{coord_Err} + \text{conf_Err} + \text{class_Err} \quad (1)$$

By introducing a confidence error term, YOLOv3 discusses the case whether the grid contains the target, which solves the problem of unstable training when the coordinate error and the classification error weight are consistent.

2.2 DenseNet

DenseNet's paper obtained the oral of CVPR2017, which proposed a new convolutional neural network connection: dense connections. It connects each layer to each of the other layers in a feedforward manner. A traditional convolutional network with an L layer has L connections between each layer and its subsequent layers, while a DenseNet network has $L(L+1)/2$ straight connections. For each layer, the feature maps of all the previous layers are used as input, and its own feature map is used as input for all subsequent layers. The dense connection module is shown in Figure 1:



The DenseNet formula is as follows. $[x_0, x_1, \dots, x_{l-1}]$ indicates that the output feature map of the 0 to l-1 layer is concatenation. Concatenation is the merging of channels, just like Inception. The former resnet is the sum of the values, and the number of channels is constant. H_l includes convolutions of BN, ReLU and 3*3.

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]) \quad (2)$$

Thanks to the design of the dense block module, the DenseNet network is narrower and has fewer parameters. Through dense connections, the use of feature maps is more efficient, and at the same time, the problem of gradient disappearance is alleviated.

3. Model

3.1 Mini-RD

Aiming at the redundancy problem of RD dataset in Document[9], This paper builds a new lightweight dataset: mini-RD ,which includes three types of target: target1, target2, target3.The train set contains 1282 images and the test set contains 92 images , covering 25 types of target scenarios, to meet the needs of this task.

3.2 YOLO-RD structure

● **Stem block** In the stem block of DenseNet networks, a 7×7 convolution kernel is applied to extract the input image. However, the large convolution kernel extraction feature is relatively rough, and some details are ignored in the early stage. We propose a new Stem Block that allows the input image to pass through a series of small-scale convolution kernels to further improve the feature extraction capabilities of the input image information.

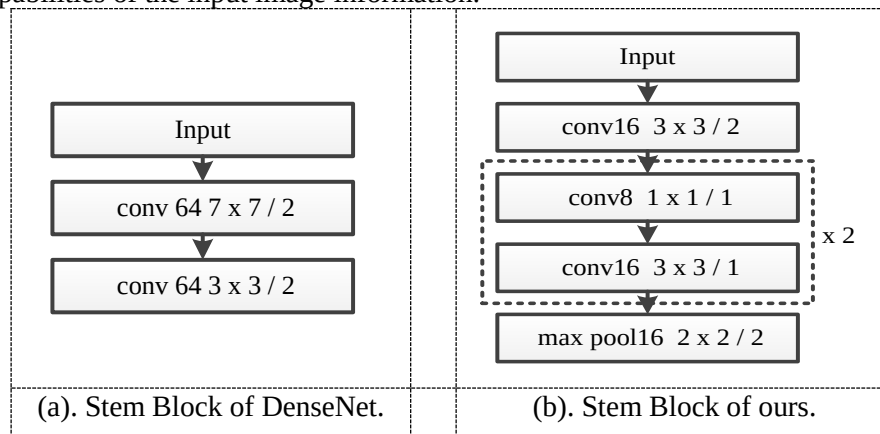
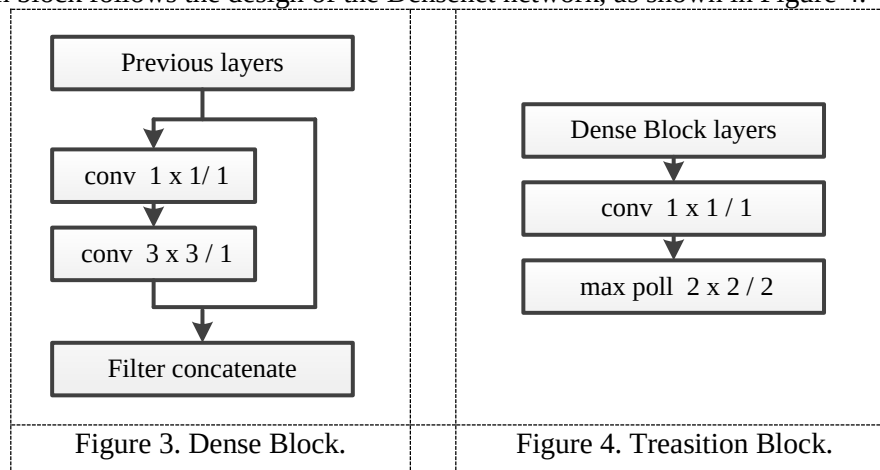


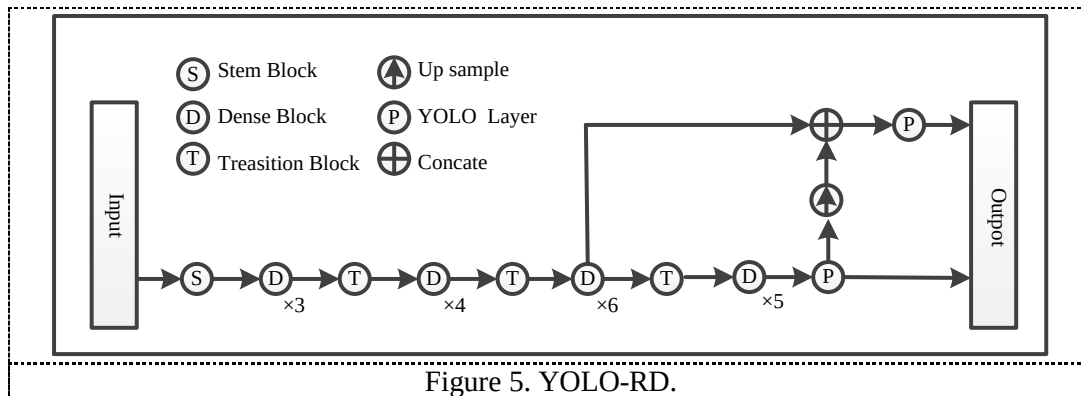
Figure 2.Stem Block

● **Dense block** According to the receptive field criterion, we believe that for the RD radar image, the main expression layer of the feature is concentrated in the 3rd and 4rd dense blocks. Through experimental comparison, the number of the dense block of our network is finally taken as {3,4,6,5}. The treasion block follows the design of the Densenet network, as shown in Figure 4.



● **Dense Prediction Structure** In the prediction module, we use the two-layer prediction structure of Tiny-YOLOv3, and the prediction layer splicing the feature extraction layers of different scales by upsampling.

In summary, the structure of the YOLO-RD model proposed in this paper is as follows:



4. Experiment

4.1 Experimental environment

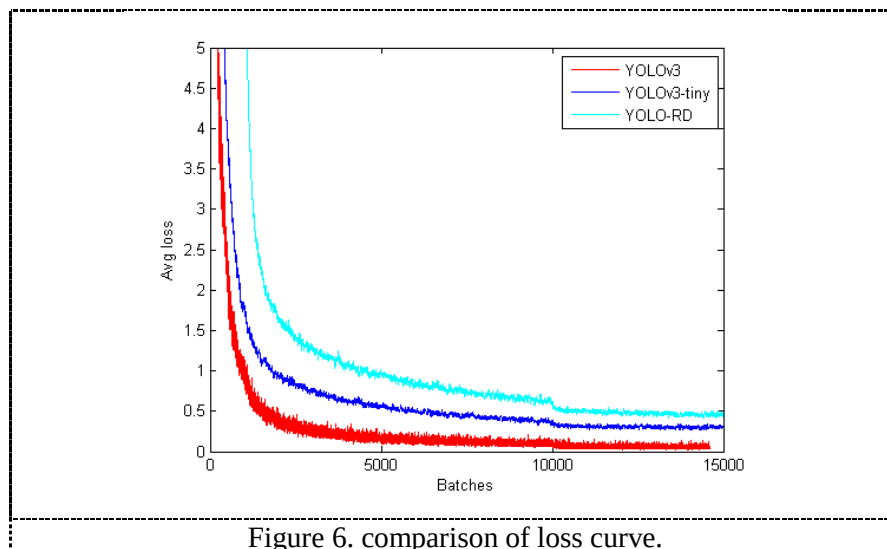
The experimental environment configuration in this paper is shown in Table 1.

Table 1. Experimental environment configuration

System	Ubuntu18.04
CPU	Intel Core i5-8400 CPU
	@2.80GHz×6
GPU	GeForce GTX 1070
Hardware	CUDA9.0;CUDNN7.5
Acceleration	

4.2 Comparison experiment with other models

To verify the experimental results, based on the Darknet framework, our paper compares the results of YOLOv3, Tiny-YOLOv3 and YOLO-RD on the mini-RD dataset. Apply different learning rate adjustment strategies to converge to their respective optimal values. The loss reduction curve is as follows:



As can be seen from the figure, the best convergence of YOLOv3, loss can be reduced to 0.029; The loss of YOLO-RD and Tiny-YOLOv3 is very close, YOLO-RD can be reduced to 0.43. It can be seen that YOLO-RD has a good learning ability.

The comparison of the model indicators is as follows:

Table 2. Comparison of model indicators.

	YOLOv3	Tiny-YOLOv3	YOLO-RD
Weight/MB	246.3	34.7	7.1
MAP/%	99.33	96.80	97.54
IOU/%	89.74	83.43	81.16
BFLOPS	65.304	5.451	6.444
Time/ms	3.272	0.919	1.603

It can be seen from the above table that YOLOv3 has achieved better training results in the case of ignore the weight. However, in the case of limited embedded memory and computing power, YOLOv3 takes up too much memory, and loading the network model consumes too much BFLOPS. In comparison, YOLO-RD achieved similar results with Tiny-YOLOv3 under the condition of occupying small memory, just 7.1MB. and it is easier to deploy on embedded devices. The performance of YOLO-RD was verified by the above comparative experiments.

4.3 Prediction results display

The following figure shows some of the results of YOLO-RD on the mini-RD dataset. It can be seen that YOLO-RD has good class prediction and border regression.

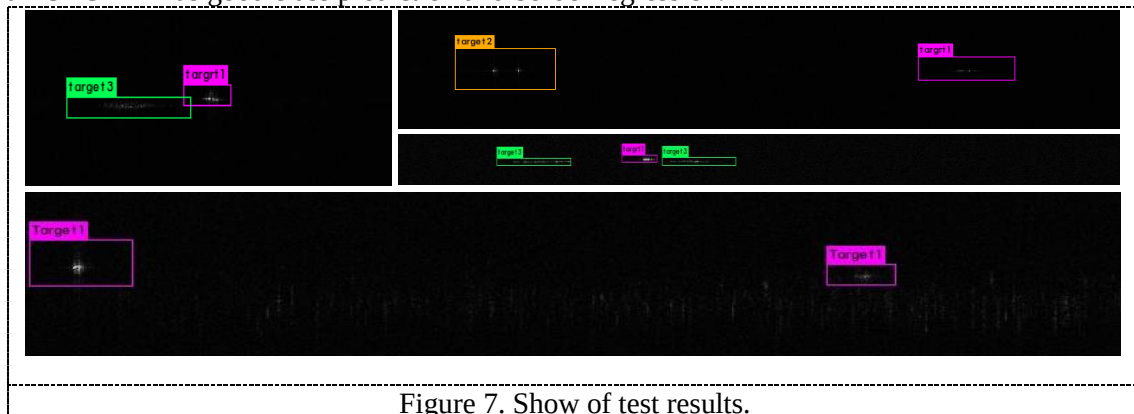


Figure 7. Show of test results.

5. Conclusion

We proposed a lightweight object detection model YOLO-RD for RD radar images. By applying effective stem block and dense block modules, our YOLO-RD is designed to be simple and efficient, with a small footprint on the mini-RD dataset for better accuracy and easy to mount on embedded devices.

References

- [1] Iandola N, Han S, Moskewicz W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[EB] . eprint arXiv: 160207360, 2016.
- [2] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [EB]. eprint arXiv: 1704. 04861, 2017
- [3] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [EB] . eprint arXiv: 1707. 01083, 2017.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD; Single shot multibox detector [J]. 2015, arXiv: 15 12. 02325.

- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once; Unified, real-time object detection [C]//Computer Vision&Pattern Recognition. 2016.
- [6] MobileNets: Efficient convolutional neural networks for mobile vision applications. Howard AG, Zhu ML, Chen B, et al.
- [7] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [8] ZHOU Long, WEI Suyuan, CUI Zhongma, et al. Multi-objective detection of complex background radar image based on deep learning [J/OL]//Systems Engineering and Electronics. <http://kns.cnki.net/kcms/detail/11.2422.TN.20190228.1544.006.htm1>.
- [9] Huang G, Liu Z, Weinberger K Q, et al. Densely connected convolutinnal networks [J/OL].2016.arXiv: 1608.06993[cs.CV].