CNN-Based Target Detection and Classification When Sparse SAR Image Dataset is Available

Hui Bi, *Member, IEEE*, Jiarui Deng, Tianwen Yang, Jian Wang, and Ling Wang, *Member, IEEE*

***Abstract*—Synthetic aperture radar (SAR) is an earth observation technology that can obtain high-resolution image in all-weather and all-time conditions, and hence, has been widely used in civil and military applications. SAR target detection and classification are the key processes for the detailed feature information extraction of the interested target. Compared with traditional matched filtering (MF) recovered result, sparse SAR image has lower sidelobes, noise, and clutter. Thus, it will theoretically has better performance in target detection and classification. In this article, we propose a novel sparse SAR image based target detection and classification framework. This novel framework first obtains the sparse SAR image dataset by complex approximate message passing (CAMP), which is an L_1-norm regularization sparse imaging method. Different from other regularization recovery algorithms, CAMP can output not only a sparse solution, but also a nonsparse estimation of considered scene that well preserves the statistical characteristic of the image when protruding the target. Then, we detect and classify the targets by using the convolutional neural network based technologies from the sparse SAR image datasets constructed by the sparse and nonsparse solutions of CAMP, respectively. For clarify, these two kinds of sparse SAR image datasets are named as $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$. Experimental results show that under standard operating conditions, the proposed framework can obtain 92.60% and 99.29% mAP on Faster RCNN and YOLOv3 by using the $\mathcal{D}_{\mathrm{Nsp}}$ sparse SAR image dataset. Under extended operating conditions, the mAP value of Faster RCNN and YOLOv3 are 95.69% and 89.91% mAP, respectively. These values based on the $\mathcal{D}_{\mathrm{Nsp}}$ dataset are much higher than the classified result based on the corresponding MF dataset.**

***Index Terms*—Convolutional neural network (CNN), complex approximate message passing (CAMP), sparse synthetic aperture radar (SAR) image, target detection and classification.**

Manuscript received December 1, 2020; revised April 14, 2021; accepted June 27, 2021. Date of publication June 30, 2021; date of current version July 15, 2021. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant NE2020004, in part by the National Natural Science Foundation of China under Grant 61901213, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20190397, in part by the Aeronautical Science Foundation of China under Grant 201920052001, and in part by the Young Science and Technology Talent Support Project of Jiangsu Science and Technology Association. *(Corresponding author: Hui Bi.)*

Hui Bi, Jiarui Deng, Jian Wang, and Ling Wang are with the Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China, and also with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: bihui@nuaa.edu.cn; djr_919@163.com; 1528114249@qq.com; tulip_wling@nuaa.edu.cn).

Tianwen Yang is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 211189, China, and also with the College of Electronics and Information Engineering, Southeast University, Nanjing 211189, China (e-mail: yangtianwen0524@163.com).

Digital Object Identifier 10.1109/JSTARS.2021.3093645

I. Introduction

As a kind of high-resolution earth observation technique, synthetic aperture radar (SAR) has all-time and all-weather surveillance ability, and has been widely used in many military and civilian fields [1], [2]. Target detection and classification are the key fields of SAR applications, which can extract the image feature information, e.g., target position, shadow, and contour, and hence, play an important role in military reconnaissance, social security, and resource exploration [3]–[6].

Traditional matched filtering (MF) based SAR imaging algorithms, such as Range Doppler algorithm [7], [8] and Chirp Scaling algorithm [9]–[11], need the echo data satisfying the Shannon–Nyquist sampling theory in scene recovery [12]. Therefore, the amount of data required to obtain high-resolution image will greatly increase the load of data storage and processing, and dramatically increase the complexity of radar system [13]. In 1990s, sparse signal processing theory was proposed, which uses fewer samples than required by traditional sampling theory to reconstruct the original signal [14]. Then, in 2006, Donoho [15] and Candes [16], [17] proposed the compressive sensing (CS), that is an important development of sparse signal processing. CS breaks the limitation of Shannon–Nyquist theory, which can achieve high-quality recovery of sparse scene with less amount of data. After introducing sparse signal processing into SAR imaging, sparse SAR imaging theory was formed. Compared with traditional SAR, sparse SAR imaging radar can decrease the system complexity and increase the swath width [13], and shows great application potential. However, typical regularization based sparse SAR imaging algorithms, e.g., iterative soft thresholding (IST) [18]–[20] and orthogonal matching pursuit (OMP) [21], [22] could only obtain the sparse estimation of the observed scene with ruined background distribution. Although the sparse image has better performance than MF based result, it will also lose the feature information of the target, which greatly reduces the accuracy of target detection and classification. To solve this problem, complex approximate message passing (CAMP) algorithm was introduced to sparse SAR imaging [23]–[25]. Different from other regularization recovery algorithms, it outputs not only a sparse image, but also a nonsparse estimation of considered scene with complete image statistical characteristics and improved quality [26]. Because of this advantage, CAMP-based sparse SAR imaging method is used to acquire the two kinds of sparse SAR image datasets, constructed by the sparse and nonsparse solutions, respectively. For clarify, these two kinds of sparse SAR image datasets are named as $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$. Compared with $\mathcal{D}_{\mathrm{Sp}}$, $\mathcal{D}_{\mathrm{Nsp}}$ will

provide more feature information for the target detection and classification.

SAR automatic target detection methods are mainly divided into two types, i.e., template-based [27]–[29] and model-based methods [30]–[32]. The core of template-based method performs feature extraction and selection, which requires wide professional knowledge as the basis. Some hidden features may not be used effectively, which limits the detection performance. The core of the model-based method lies in the design of target model, which relies too much on the acquisition of target model information and requires time-consuming high-frequency electromagnetic calculation. Deep learning technique provides a new solution without artificial feature design and object modeling. In 2012, Hinton *et al.* [33] designed a deep convolutional neural network (CNN), named as AlexNet. In the ImageNet Large Scale Visual Recognition Challenge [34], the Top5 error ratio of AlexNet is just 17.0%, which is considerably better than the state-of-the-art then. This makes CNN become the most important tool in the field of target detection and classification. Meanwhile, it also attracts the attention of researchers in the field of radar image processing. CNN-based target detectors are usually divided into two types, one-stage object detector [35]–[37] and two-stage object detector [38]–[41]. Two-stage object detector first generates the target candidate bounding box and then uses the target detection network to classify the candidate bounding box and perform border regression. The most representative two-stage object detector is RCNN series [35]–[37]. The one-stage object detector directly outputs the target coordinates and conditional probabilities of all classes. Its representative models are Single Shot MultiBox Detector (SSD) [41] and YOLO [38]–[40] series. Faster RCNN [37] and Yolov3 [40] have the best performance in the abovementioned two kinds of detectors, respectively, and hence, being selected for the SAR target detection and classification. Nowadays, several researchers have applied CNN-based methods to solve the SAR target detection and classification problems. Dong *et al.* [42] proposed a modified Faster RCNN model and SSD model with data augmentation to address target recognition problem. Kang *et al.* [43] modified Faster RCNN by the traditional constant false alarm rate so as to better detect the SAR target. Wang *et al.* [44] designed a deep framework using multiple CNNs for feature-fused SAR target discrimination. However, all these works are based on the MF recovered SAR image. It is known that compared with MF-based image, sparse SAR image has better quality with lower sidelobes and reduced noise and clutter. Thus, it is meaningful to study CNN-based target detection and classification technique when the sparse SAR image dataset is available.

In this article, we propose a novel sparse SAR image based target detection and classification framework. This framework first obtains the sparse SAR image datasets $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$ by using CAMP based sparse SAR imaging method. Then, it detects the targets by using two conventional CNN-based methods, Faster RCNN and YOLOv3, for the constructed sparse SAR image dataset. Experimental results based on MSTAR data show that compared with MF dataset and $\mathcal{D}_{\mathrm{Sp}}$ composed of sparse SAR images with damaged statistical distribution, $\mathcal{D}_{\mathrm{Nsp}}$ shows better performance in CNN based target detection and classification. When $\mathcal{D}_{\mathrm{Nsp}}$ is available, under extended operating conditions (EOC), the mAP value of Faster RCNN and YOLOv3 are 95.69% and 88.21% mAP, respectively. While under standard operating conditions (SOC), these values even reach 92.60% and 99.29%, which is a good result in practical SAR target detection process.

The rest of this article is organized as follows. Section II introduces the CAMP-based sparse SAR imaging principles for echo data and complex image data, respectively. Target detection and classification models of Faster RCNN and Yolov3 are described in Section III. Section IV shows the experimental results and performance analysis of SAR target detection and classification based on different datasets. Finally, Section V concludes this article.

II. CAMP-Based Sparse SAR Imaging

In this section, the advantages of CAMP algorithm in SAR imaging performance improvement and image statistical characteristics preservation will be discussed. This is the precondition for CNN-based target detection and classification to be presented.

A. Sparse SAR Imaging From Echo Data

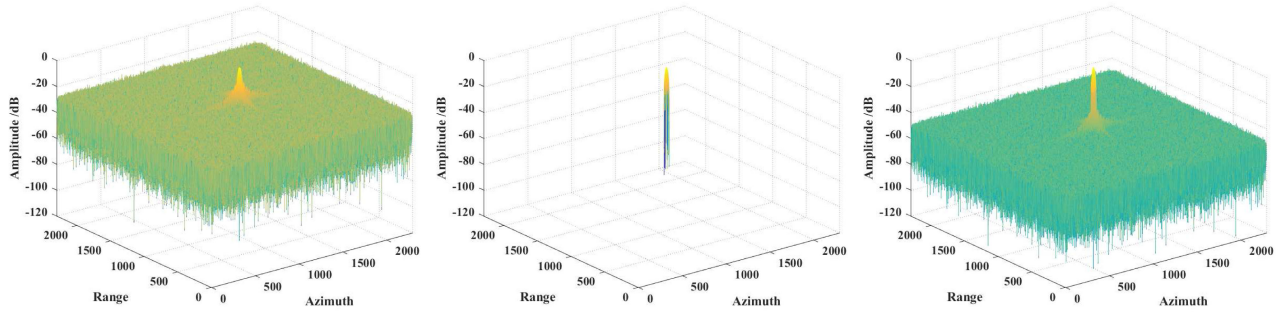
As discussed in [26], one-dimensional (1-D) sparse SAR imaging model can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}_0 \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{M\times 1}$ and $\mathbf{x} \in \mathbb{C}^{N\times 1}$ are the echo data and backscattering coefficient of considered scene, respectively, $\mathbf{n}_0 \in \mathbb{C}^{M\times 1}$ is the noise vector, and $\mathbf{H} \in \mathbb{C}^{M\times N}$ is the system measurement matrix, which represents the transmitted signal and the imaging geometry relationship between radar and surveillance area. According to the CS theory [15], when $\mathbf{x}$ is sparse enough and $\mathbf{H}$ satisfies the RIP condition [17], the sparse scene can be recovered by solving

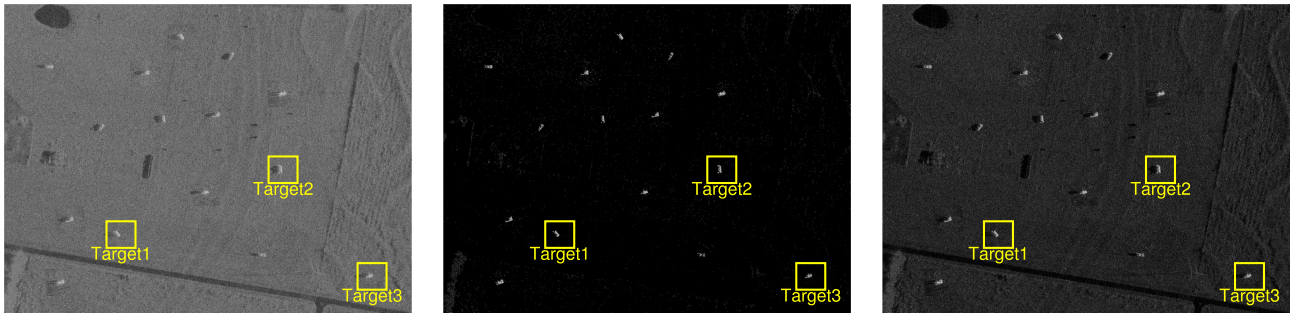
$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \right\} \tag{2}$$

where λ is the regularization parameter. After recovery, the 2-D backscattering coefficient $\hat{\mathbf{X}}$ of considered scene can be obtained by reshaping $\hat{\mathbf{x}}$. For the Lasso problem in (2), CAMP algorithm can be used for the scene recovery. The detailed iterative procedures are listed in [26]. Different from other regularization recovery algorithms, CAMP can obtain not only the traditional sparse image $\hat{\mathbf{x}}$, but also a nonsparse estimation $\tilde{\mathbf{x}}$ of the considered scene, which has an improved image quality and well preserved background statistical distribution. compared with MF based result. It is known that compared with conventional sparse SAR imaging technique via model in (1), MF-based method has better calculation efficiency. However, its recovered image usually suffers from serious noise and sidelobes, which will affect the further application of the image. In addition, compared with original echo, the available data is always the MF recovered SAR complex image, such as the used MSTAR dataset. Therefore, in order to obtain a large number of sparse SAR images for further application, the complex image based



(a) (b) (c)

Fig. 1. Reconstructed images of simulated scene by different methods. (a) MF. (b) Sparse solution $\hat{\mathbf{X}}$ of CAMP-based method. (c) Nonsparse solution $\tilde{\mathbf{X}}$ of CAMP-based method.



(a) (b) (c)

Fig. 2. Reconstructed images of considered scene by different methods. (a) MF. (b) Sparse solution $\hat{\mathbf{X}}$ of CAMP-based method. (c) Nonsparse solution $\tilde{\mathbf{X}}$ of CAMP-based method.

sparse SAR imaging technique is essential to enhance the MF dataset directly.

B. Sparse SAR Imaging From Complex Image Data

According to [45], the complex image based sparse SAR imaging model can be written as

$$\mathbf{X}_{\mathrm{MF}} = \mathbf{X} + \mathbf{N} \tag{3}$$

where $\mathbf{X} \in \mathbb{C}^{N_P(\mathrm{Azimuth}) \times N_Q(\mathrm{Range})}$ is the 2-D backscattering coefficient of the scene, whose (p, q) entry is $x(p, q)$, $\mathbf{X}_{\mathrm{MF}}$ is the known complex-valued MF recovered SAR image, $\mathbf{N} \in \mathbb{C}^{N_P \times N_Q}$ is a complex matrix that denotes the difference between $\mathbf{X}_{\mathrm{MF}}$ and $\mathbf{X}$ including sidelobes, noise, etc. Similar to the sparse imaging from echo data, we can recover the scene of interest by solving the following Lasso problem:

$$\hat{\mathbf{X}} = \min_{\mathbf{X}} \left\{ \frac{1}{2} \|\mathbf{X}_{\mathrm{MF}} - \mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_1 \right\}. \tag{4}$$

CAMP algorithm can also be used to solve the optimization problem in (4). The detailed iterative process is shown in [45]. Similar to the CAMP algorithm via echo data, the complex image based CAMP algorithm can also output the sparse ($\hat{\mathbf{X}}$) and nonsparse ($\tilde{\mathbf{X}}$) images of the considered scene, which have the similar performance to the images recovered from echo data.

TABLE I
TBR VALUES OF IMAGE RECONSTRUCTED BY MF AND CAMP-BASED METHODS

Target	Target 1	Target 2	Target 3
MF	38.20 dB	34.11 dB	34.75 dB
$\hat{\mathbf{X}}$ **of CAMP**	57.30 dB	52.26 dB	53.86 dB
$\tilde{\mathbf{X}}$ **of CAMP**	60.73 dB	55.17 dB	57.53 dB

CAMP-based algorithm introduces a term of "state evolution" to evolve the standard deviation of "noise" as the iteration proceeds, and thus produces the sparse estimation $\hat{\mathbf{X}}$ and nonsparse estimation $\tilde{\mathbf{X}}$ of the scene simultaneously [24]. Therefore, different from other regularization recovery algorithms, such as IST [18]–[20] and OMP [21], [22], CAMP can obtain a nonsparse solution with a similar background statistical distribution to that of the MF-based image.

C. Verification

In the following, experiments based on the MSTAR dataset will be used to verify the CAMP-based algorithm in SAR imaging performance improvement and image statistical distribution preservation. Figs. 1 and 2 show the image recovered by MF- and CAMP-based sparse SAR imaging methods, respectively.

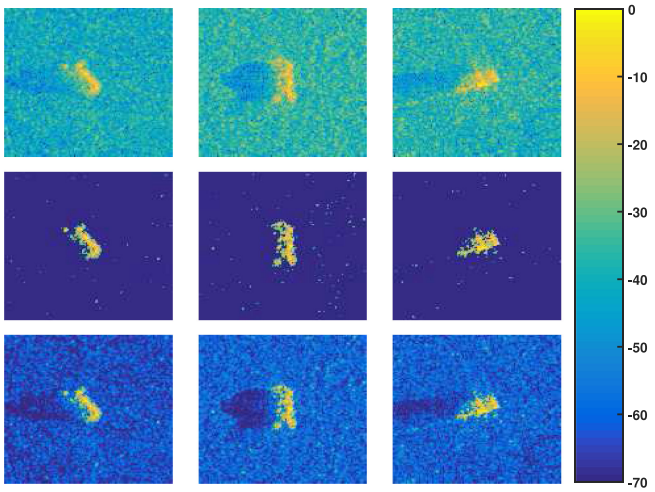


Fig. 3. Zoom-in plots of three selected targets in Fig. 2. From top to bottom rows: MF recovered images, sparse solutions of CAMP-based method, and nonsparse solutions of CAMP-based method. From left to right columns, the focused targets are Target1, Target2, and Target3, respectively.

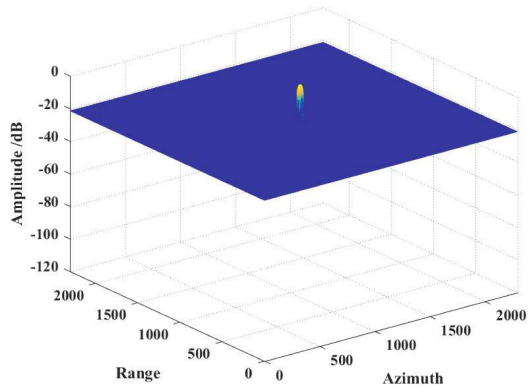


Fig. 4. Amplitude deviation between $\tilde{\mathbf{X}}$ and the MF-based image.

TABLE II
DATA DESCRIPTION FOR SOC

Class	Serial No.	Train set (Depression17°)	Test set (Depression15°)
2S1	B01	299	274
BMP2	SN9563	233	196
BRDM2	E-71	298	274
BTR60	Kloyt7532	256	195
BTR70	C71	233	196
D7	92v13015	299	274
T62	A51	299	273
T72	SN132	232	196
ZIL131	E12	299	274
ZSU234	D08	299	274
Total		2747	2426

Compared with MF-based image, it is seen that both sparse and nonsparse solutions of CAMP have better image quality, which is very helpful for the further SAR image applications. To evaluate the performance improvement quantitatively, target-to-background ratio (TBR) is selected as the judging criterion, defined as [46]

$$\mathrm{TBR}\left(\mathbf{X}\right) \triangleq 20\log_{10}\left(\frac{\max_{(u,v)\in\mathcal{T}}\left|\left(\mathbf{X}\right)_{(u,v)}\right|}{\left(1/N_{\mathcal{B}}\right)\sum_{(u,v)\in\mathcal{B}}\left|\left(\mathbf{X}\right)_{(u,v)}\right|}\right) \tag{5}$$

where $\mathcal{T}$ is the target area which is surrounded by the background region $\mathcal{B}$, and $N_{\mathcal{B}}$ is the number of pixels in the background region. Three targets indicated by the yellow rectangles in Fig. 2 are selected to calculate the TBR values (see Table I). Their zoom-in images are shown in Fig. 3. From Table I and Fig. 3, it is seen that the TBR values of both $\hat{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ of selected targets are all above 50 dB, which show better image quality than MF recovered images. To demonstrate the ability of image statistical distribution preservation by the CAMP-based sparse SAR imaging methods, the difference between MF-based image in Fig. 1(a) and $\tilde{\mathbf{X}}$ in Fig. 1(c) are calculated and shown in Fig. 4. From Fig. 4, it is seen that compared with MF-based result, $\tilde{\mathbf{X}}$ has the similar background statistical distribution. However, it can suppress the amplitude value by about 20 dB in the nontarget area, which means that $\tilde{\mathbf{X}}$ has both better image performance and complete feature information of the target. This is very helpful for further SAR target detection and classification applications.

III. CNN-Based Target Detection and Classification Framework Via Sparse SAR Image Datasat

A. Principle of CNN

CNN is a deep feed forward neural network with excellent feature learning ability, which realizes the receptive field by convolution. There are three major characteristics of CNN, i.e., locality of features based local field, repeatability of features based weight sharing, and pooling operation in the subsampled processing. These operations greatly decrease the used parameters for deep learning, and hence, reduce the complexity of network. This makes CNN achieve better fault tolerance and robustness. Generally, as shown in Fig. 5, the typical structure of CNN is composed of input layer, convolution layer, pooling layer, fully-connected layer, and output layer. The input layer is used to receive the input image data. In this article, the input is a single-channel grayscale SAR image. Thus, the feature of the input layer contains only one feature map, i.e., $x^1 = \{x_1^1\}$.

The convolutional layer simulates the response mechanism of human neurons to visual stimuli. The function of this layer is to perform convolution operation on the input data to extract image maps, and connect the result locally to the next layer. Generally, the more convolutional layers, the stronger ability of the network to express features. The feature map of the convolutional layer can be described as

$$x_j^m = f\left(\sum_{i=1}^{N^{m-1}} \mathbf{G}_{i,j}^m (k_{i,j}^m \otimes x_i^{m-1}) + b_j^m\right) \tag{6}$$

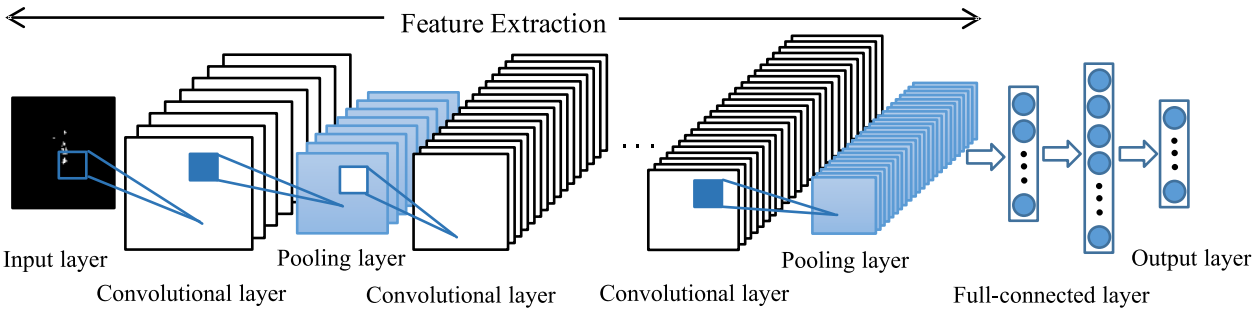


Fig. 5. Typical structure of CNN [33].

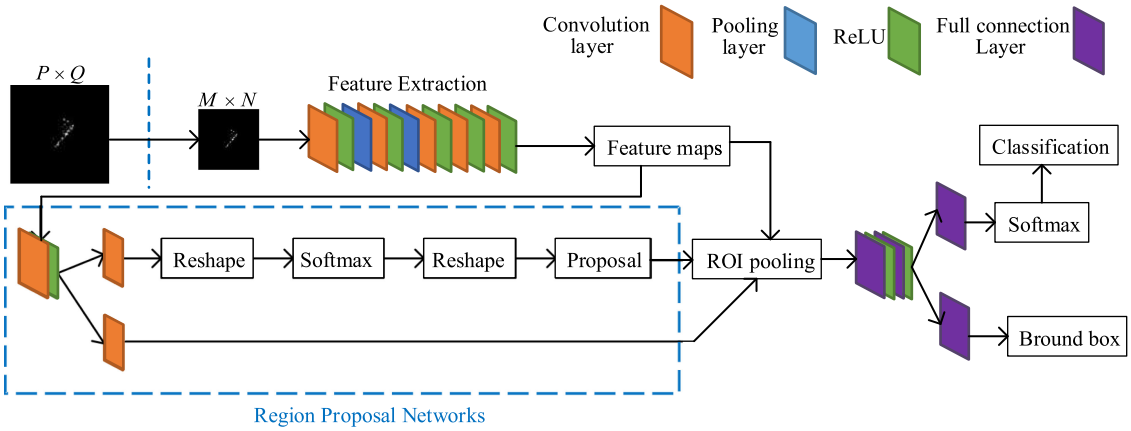


Fig. 6. Architecture of Faster RCNN with ZFNet [37].

TABLE III
COMPARISON OF DIFFERENT DATASETS ON FASTER RCNN UNDER SOC

Dataset	Category										mAP	Time
	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234		
MF	89.77%	81.80%	92.76%	77.65%	87.99%	91.83%	91.49%	90.73%	93.46%	93.84%	**89.13%**	
$\mathcal{D}_{\mathrm{Sp}}$	88.56%	79.37%	89.62%	80.99%	81.62%	90.08%	87.44%	86.36%	85.08%	94.92%	**86.40%**	31.12ms
$\mathcal{D}_{\mathrm{Nsp}}$	93.15%	87.86%	96.69%	86.04%	92.26%	93.94%	94.54%	93.92%	93.90%	93.77%	**92.60%**	

TABLE IV
COMPARISON OF DIFFERENT DATASETS ON YOLOV3 UNDER SOC

Dataset	Category										mAP	Time
	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234		
MF	99.27%	97.90%	99.98%	97.09%	98.19%	99.54%	99.99%	98.97%	99.98%	99.57%	**99.05%**	
$\mathcal{D}_{\mathrm{Sp}}$	99.25%	87.97%	98.54%	88.47%	89.57%	99.59%	99.62%	97.74%	97.96%	99.88%	**95.86%**	15.67ms
$\mathcal{D}_{\mathrm{Nsp}}$	99.59%	98.59%	99.67%	96.92%	98.43%	100.0%	99.99%	99.71%	99.99%	99.62%	**99.29%**	

where x_j^m is the jth feature map in the mth layer, N^{m-1} is the number of feature maps in the $m-1$th layer, $\mathbf{G}_{i,j}^m$ is the connection matrix between x_j^m and x_i^{m-1}, $k_{i,j}^m$, and b_j^m are the convolution kernels and bias, respectively, and $f(\cdot)$ is the nonlinear activate function, which is usually set as tanh, sigmoid, ReLU, or SELU.

The pooling layer, also known as the subsampling layer, is usually located between successive convolutional layers. The main function of pooling layer is to make the features have a certain degree of spatial invariance, and reduce the parameters and computations to avoid overfitting. The feature map of this

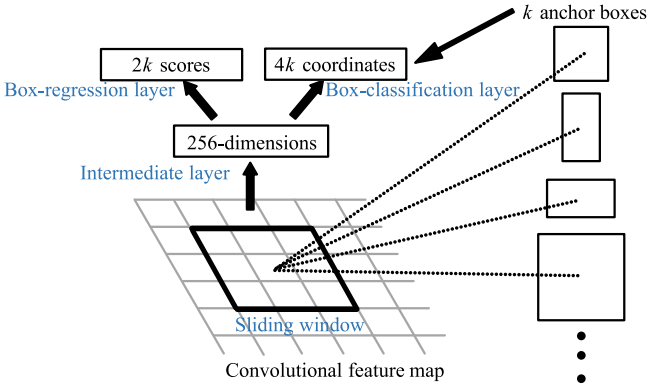


Fig. 7. Architecture of RPN [43].

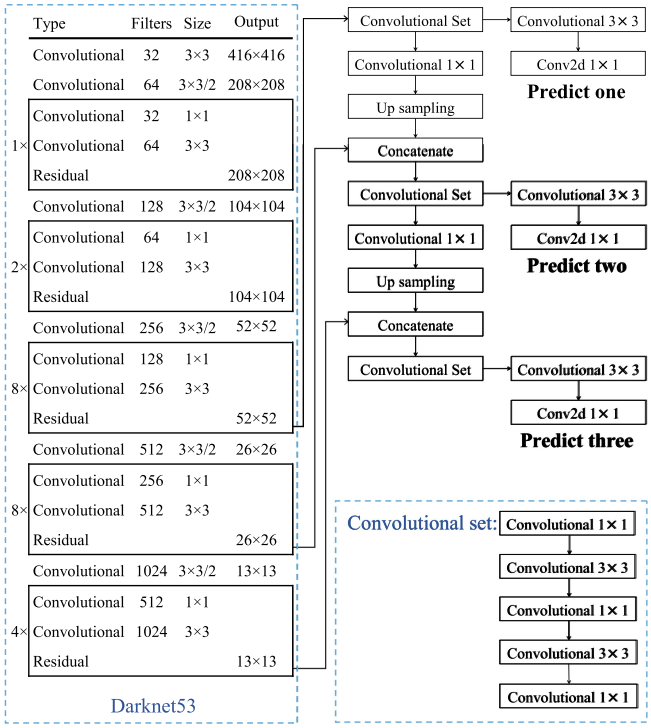


Fig. 8. Network structure of YOLOv3 [40].

layer can be expressed as

$$x_j^m = p(x_j^{m-1}) \tag{7}$$

where $p(\cdot)$ is the pooling function. Common pooling functions are mean-pooling and max-pooling. In recent years, some CNN-based detectors have used convolutional layers with a step size greater than 1 instead of pooling layers, making pooling layers not a necessary part of CNN.

The full-connected layer is set after the feature extraction. Its function is to connect all the neurons in the previous layer with the neurons in the current layer, and then map the features according to the specific task of the output layer. The form of the output layer is determined by the specific task that the network needs to complete. If the convolutional neural network is used as a classifier, the output layer will use the softmax logistic regression model to output prediction vectors of all classes, i.e., $\mathbf{y} = (y_1, y_2, \ldots, y_H)^{\mathrm{T}}$, where H represents the number of classes.

TABLE V
DATA DESCRIPTION FOR EOC

Class	Serial No.	Training set (Depression17°)	Testing set (Depression30°)
2S1	B01	299	288
BRDM2	E-71	298	287
T72	A64	299	288
ZSU234	D08	299	288
Total		1195	1151

B. *Faster RCNN*

Faster RCNN is a region-based two-stage target detection algorithm based on CNN. It first generates the candidate regions, then classifies the candidate regions, and finally refines the locations. Faster RCNN consists of four parts, convolution layer, region proposal networks (RPN), region of interest (ROI) pooling, and classification. The architecture of Faster RCNN with ZFNet is shown in Fig. 6, whose main steps for deep learning can be summarized as follows [37].

1) The image data are input into CNN to obtain the corresponding feature maps.

2) The feature maps are transmitted through two different ways, one to RPN, and the other to forward.

3) The PRN calculates the region proposals through the feature maps, then performs maximum suppression on the region proposals and outputs the score of each region proposals.

4) The Top-N ranked proposal regions in step 3 and feature maps obtained in step 2 are passed to the ROI pooling layer to obtain the features corresponding to the region proposals.

5) The features of region proposals are introduced to the full-connected layer. Then, the result of classification and regression will be outputted.

The main difference between Faster RCNN and other algorithms of RCNN series is that RPN is proposed in Faster RCNN to specifically recommend candidate regions, which realizes an end-to-end target detection framework. RPN shares the full-image convolutional features with the entire network, thus, makes region proposals almost free. The main idea of PRN is to distinguish candidate boxes and optimize the target position according to the feature maps by the network convolution layers. The structure of PRN is shown in Fig. 7 [43]. In Fig. 7, for each sliding-window location, the PRN generates k bounding boxes, known as anchor boxes at multiple scales and aspect ratios. Then, the sliding window will be mapped to a lower dimensional feature, i.e., 256 dimensions for ZFNet, and fed into a box-regression layer and a box-classification layer. For k anchor boxes of each sliding window, box-regression layer, and box-classification layer will output 4 k coordinates and 2 k scores, respectively. These coordinates and scores will be used to estimate the probability of being an object or not for the proposal of each anchor box. Before RPN was proposed, the most popular

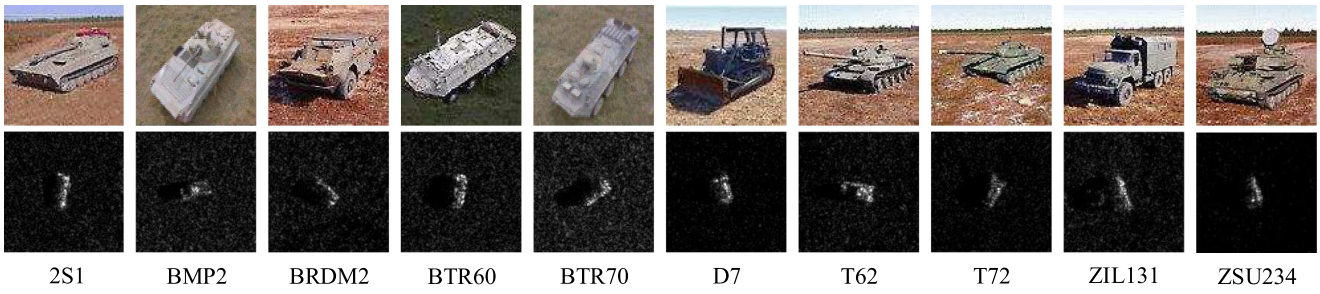


Fig. 9. SAR images of ten classes of targets in MSTAR dataset and their corresponding optical images.

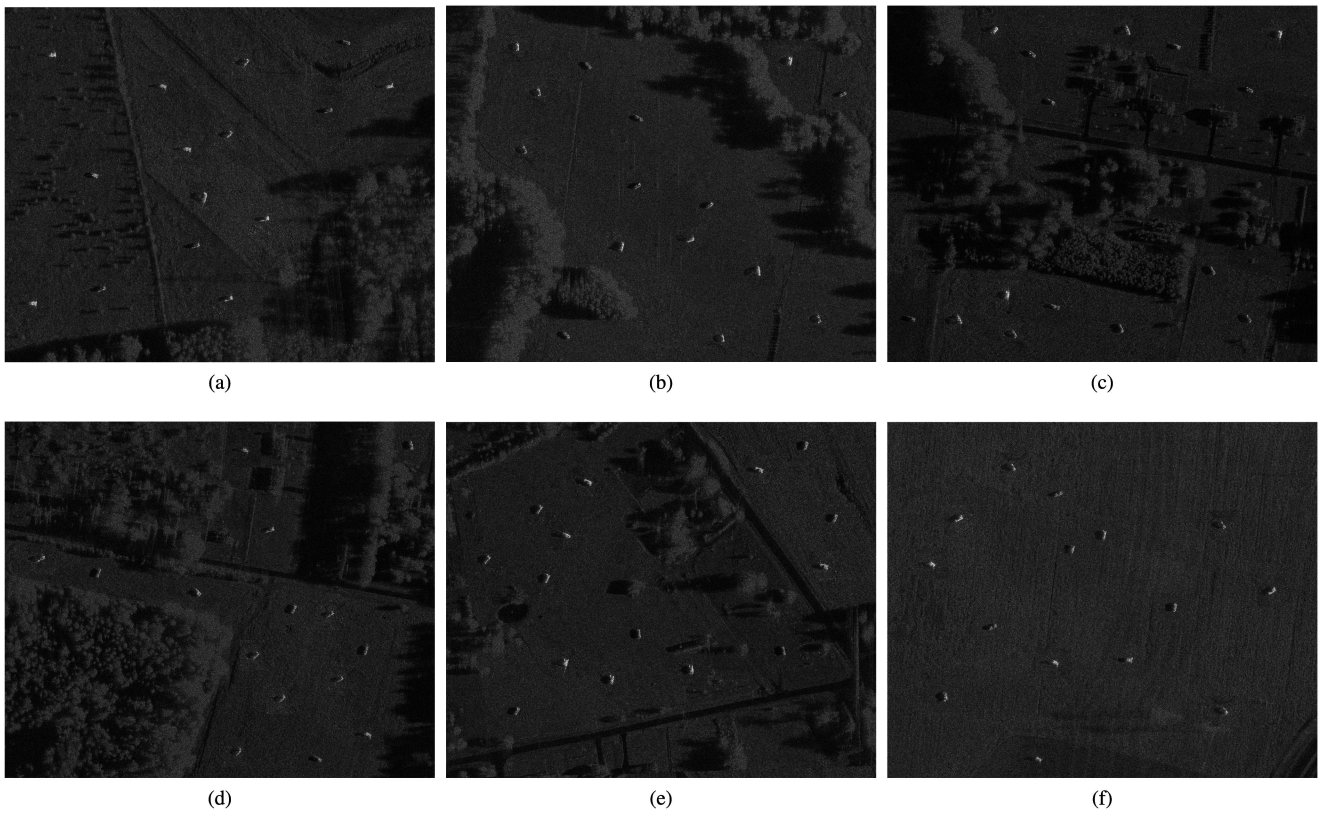
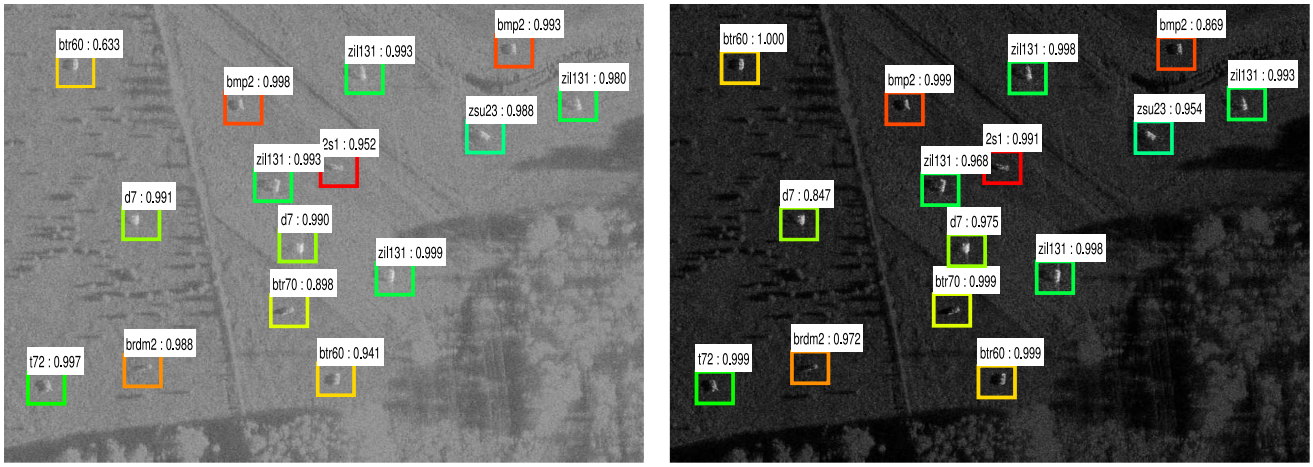


Fig. 10. Examples in the dataset $\mathcal{D}_{\mathrm{Nsp}}$. Each image in this dataset is fused with 15 targets randomly.

TABLE VI
COMPARISON OF DIFFERENT DATASETS ON FASTER RCNN UNDER EOC

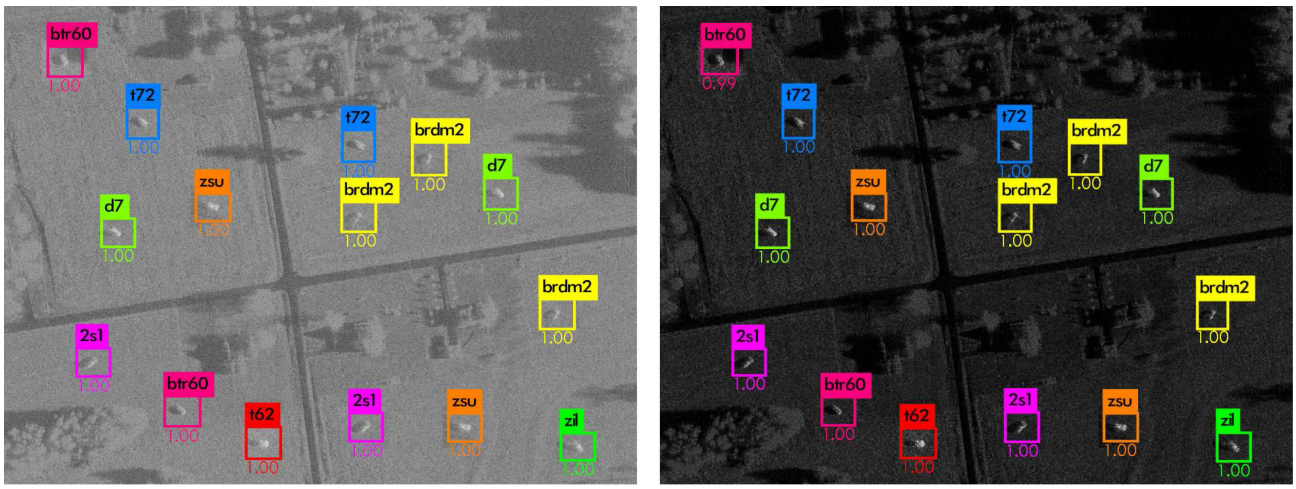
Dataset	Category				mAP	Time
	2S1	BRDM2	T72	ZSU234		
MF	88.44%	90.07%	86.80%	92.26%	**89.39%**	
$\mathcal{D}_{\mathrm{Sp}}$	93.73%	93.80%	92.86%	95.08%	**93.87%**	31.12ms
$\mathcal{D}_{\mathrm{Nsp}}$	95.64%	98.43%	94.55%	94.12%	**95.69%**	

region proposal approach was selective search method, which is computationally expensive and time-consuming. In Faster RCNN, PRN replaces the selective search method. It optimizes the structure of the proposed region for efficient and accurate region proposal generation, and shares the convolution features with convolutional layers for dramatically reducing the computational cost. RPN improves the quality and efficiency of region proposal, thereby improving the accuracy and speed of target detection and classification in Faster RCNN.



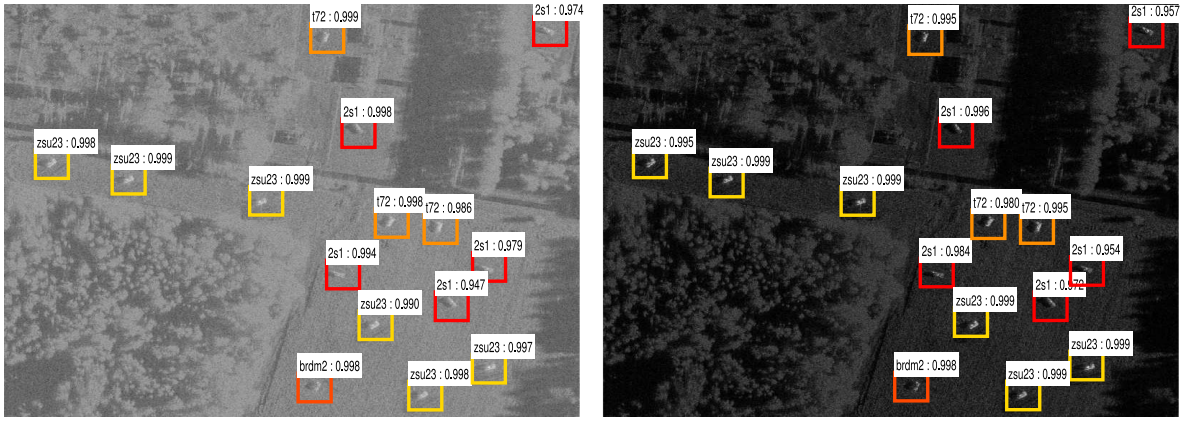
(a) (b)

Fig. 11. Target detection and classification results of Faster RCNN under SOC. (a) Faster RCNN with MF dataset. (b) Faster RCNN with $\mathcal{D}_{\mathrm{Nsp}}$.



(a) (b)

Fig. 12. Target detection and classification results of YOLOv3 under SOC. (a) YOLOv3 with MF dataset. (b) YOLOv3 with $\mathcal{D}_{\mathrm{Nsp}}$.



(a) (b)

Fig. 13. Target detection and classification results of Faster RCNN under EOC. (a) Faster RCNN with MF dataset. (b) Faster RCNN with $\mathcal{D}_{\mathrm{Nsp}}$.

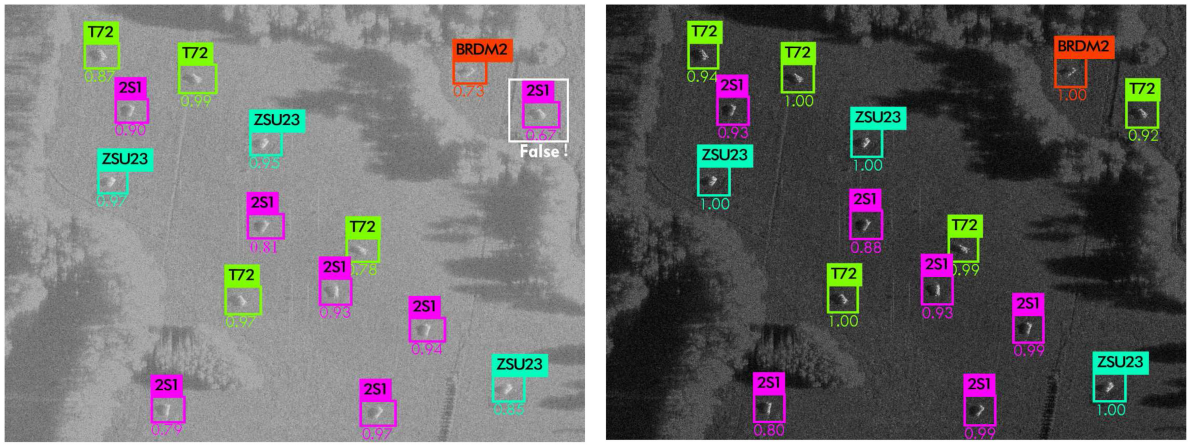


Fig. 14. Target detection and classification results of YOLOv3 under EOC. (a) YOLOv3 with MF dataset. (b) YOLOv3 with $\mathcal{D}_{\mathrm{Nsp}}$. White box represents the false classification.

TABLE VII
COMPARISON OF DIFFERENT DATASETS ON YOLOV3 UNDER EOC

Dataset	Category				mAP	Time
	2S1	BRDM2	T72	ZSU234		
MF	83.79%	83.41%	79.91%	92.17%	**84.82%**	
$\mathcal{D}_{\mathrm{Sp}}$	88.06%	96.46%	80.37%	91.85%	**89.19%**	15.67ms
$\mathcal{D}_{\mathrm{Nsp}}$	89.14%	97.73%	84.46%	88.31%	**89.91%**	

C. YOLOv3

YOLOv3 is a regression-based one-stage target detection algorithm based on CNN, which directly outputs target coordinates and conditional class probabilities. The main idea of YOLOv3 can be summarized as follows. First, feature extraction network, named as Darknet53, is used to extract features from the input image to obtain a feature map of a certain size. The input image is divided into $S \times S$ grids, which has three scales in YOLOv3, i.e., 13, 26, and 52. The selection of scale is determined by the size of feature map. Each grid is responsible for predicting three bounding boxes of objects. The center coordinates of these objects are located in the grid. Each bounding box corresponds to five parameters, i.e., four coordinates and one objective prediction. Only the bounding box that mostly overlaps the ground truth box of the object is selected to predict the target. The other bounding boxes are used only to calculate the confidence. Then, YOLOv3 will output a result of $S \times S \times (5 + C)$ dimension for each scale, where C is the number of classes.

The network structure of YOLOv3 is shown in Fig. 8. Compared with other algorithms in YOLO series, a major improvement of YOLOv3 is that it introduces a deeper network, called darknet53, to achieve the better feature extraction [40]. Darknet53 has 53 convolutional layers including one connected layer. It directly discards the pooling layers and uses the convolutional layers with a step size of 2 for subsampling. To solve the problems of increasing training difficulty and decreasing accuracy in the network, according to ResNet [47], darknet53 uses several residual modules, and then designs a novel one with more layers and higher accuracy. Nowadays, darknet53 is still one of the most advanced feature extraction networks. In terms of target prediction, with the help of feature pyramid networks (FPN) [48], YOLOv3 uses multiscale prediction to predicts bound boxes with three different scales. For the input image with the size of 416×416, as shown in Fig. 8, YOLOv3 predicts the scales of 13×13, 26×26, 52×52, respectively. The multiscale prediction enables YOLOv3 can be used to detect the targets with different receptive field sizes, which significantly improves the detected ability of small target [40].

IV. SPARSE SAR IMAGE BASED TARGET DETECTION AND CLASSIFICATION

A. Dataset

In this section, the MSTAR dataset constructed by the MF recovered images will be used to verify the proposed framework. MSTAR is a public dataset containing enough SAR target samples with resolution being $0.3 \times 0.3\,\mathrm{m}$, and is widely used in the field of SAR target detection and classification. In this dataset, there are ten different classes of military vehicle targets. The aspect angle of each target class ranges from 0° to 360°. The SAR images of ten classes of targets in MSTAR and their corresponding optical images are shown in Fig. 9. In the experiments of this paper, SAR images with 15 different scenes are selected

for background fusion. Each scene is fused with 15 targets. The class of each target in the scene is completely random, which means that the fused image may contain all categories or only one category. This randomness of dataset makes the experimental results more reliable and applicable. In the proposed framework, we first reconstruct the fused images of MSTAR data by using the CAMP-based sparse SAR imaging method, and obtain the novel sparse SAR image dataset $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$, respectively. Examples in the $\mathcal{D}_{\mathrm{Nsp}}$ dataset are shown in Fig. 10. In this article, the results of SAR target detection and classification based on the original MF dataset (MSTAR), the CAMP's sparse solution dataset $\mathcal{D}_{\mathrm{Sp}}$, and the CAMP's nonsparse solution dataset $\mathcal{D}_{\mathrm{Nsp}}$ will be compared. All experiments are conducted on the hardware platform with NVIDIA RTX2080Ti GPU and Intel Xeon CPU. The results of target detection and classification will be compared by the evaluation indexes, mAP. Average precision (AP) is an index combining precision and recall rate, which can comprehensively evaluate the recognition performance of the model. In general, the model performance is proportional to AP. mAP is the average of APs on multiple validation sets.

B. Comparison Under SOC

SOC is suitable for the situation where the target categories and serial numbers in the test set are the same as those in the train set, but with the different depression angles. In the train set, the targets are acquired at a 17° depression angle. While in the test set, the targets are collected at a 15° depression angle. The serial number, depression angle, and number of targets per class in train set and test set are shown in Table II. All target slices are randomly merged into 15 different scenes. Each fused image contains 15 targets. Due to random fusion, very few targets are repeatedly fused in different images. After abovementioned fusion, the train set contains 220 images and the test set contains 200 images.

1) Faster RCNN. The network with learning rate of 0.001 and batch size of 16 is trained first. Quantitative experimental results of Faster RCNN based on different datasets are listed in Table III. Fig. 11 shows the examples of target detection and classification result based on MF and $\mathcal{D}_{\mathrm{Nsp}}$ datasets, respectively. From Table III, it is seen that by using the CAMP's nonsparse solution dataset $\mathcal{D}_{\mathrm{Nsp}}$, Faster RCNN can obtain 92.60% mAP in the test set, which is much higher than 89.13% mAP of MF dataset, and 86.40% mAP of CAMP's sparse solution dataset $\mathcal{D}_{\mathrm{Sp}}$. The main reason is that traditional sparse SAR image, such as the sparse solution of CAMP, destroys the background statistical distribution and details of target features, thus, greatly reduces the accuracy of large-class target classification. Compared with the sparse SAR image, the nonsparse solution of CAMP algorithm can well preserve the feature information of the target, e.g., shadow, details of target with low amplitude, which is very helpful for the high accuracy target classification of SAR targets. In addition, compared with MF-based image, the nonsparse solution of CAMP has better image quality, which also will improve the accuracy of classification.

2) YOLOv3. In the following, the comparison based on YOLOv3 will be performed for the three kinds of datasets discussed previously. The network with initial learning rate of 0.001 and batch size of 16 is trained in this experiment. Quantitative results of YOLOv3 are listed in Table IV. Fig. 12 shows the target detection and classification results based on MF and $\mathcal{D}_{\mathrm{Nsp}}$ datasets, respectively. Similar to the result of Faster RCNN, $\mathcal{D}_{\mathrm{Nsp}}$ also shows the best classification performance with 99.29% mAP, which outperforms MF dataset and $\mathcal{D}_{\mathrm{Sp}}$ by 0.14% and 3.43%, respectively. In addition, it should be noted that since the mAP of YOLOv3 via MF dataset has reached 99.01%, there is no room for improvement. Therefore, an increase of 0.14% is meaningful for target detection and classification. After comparing the results of YOLOv3 and Faster RCNN under SOC, it can be seen that YOLOv3 has obvious advantages in both classification accuracy and detection time. It outperforms Faster RCNN by 6.69% mAP when $\mathcal{D}_{\mathrm{Nsp}}$ is available. In terms of detection time, YOLOv3 only needs 15.67 ms per image, which is much faster than Faster RCNN.

C. Comparison Under EOC

EOC is consist of EOC-1 and EOC-2. EOC-1 is suitable for the situation where the target in the train set and the test set has a big change in depression angle. Without loss of generality, in this article, we use EOC-1 as the example to validate the proposed framework. Similar result will be obtained under EOC-2. EOC-1 contains four categories for training and testing, whose serial number, depression angle, and number of targets per class are shown in Table V. In the training set, the targets are acquired at the depression angle of 17°. In the test set, the targets are collected at 30° depression angle. Because SAR image is extremely sensitive to the change of depression angles, the 13° depression angle change in EOC-1 will increase the difficulty of target detection and classification. After background fusion, the train set contains 90 images and the test set contains 95 images. The framework and network parameters in this experiment are the same as those in SOC.

1) Faster RCNN. Experimental results of Faster RCNN under EOC are shown in Table VI. It is seen that Faster RCNN still gets improved results with 95.69% mAP by using $\mathcal{D}_{\mathrm{Nsp}}$ dataset. Compared with the result based on MF dataset, the mAP via $\mathcal{D}_{\mathrm{Nsp}}$ is significantly increased by 6.30%, which is an important improvement of target classification performance. From Table VI, it also can be seen that compared with MF dataset, $\mathcal{D}_{\mathrm{Sp}}$ is still helpful in improving the recognition performance, but not as good as $\mathcal{D}_{\mathrm{Nsp}}$. Examples of the target classification results under EOC are shown in Fig. 13.

2) YOLOv3. As shown in Table VII, when the CAMP's nonsparse solution dataset $\mathcal{D}_{\mathrm{Nsp}}$ is used for the target classification, YOLOv3 will achieve the optimal result with 89.91% mAP under EOC. It outperforms MF dataset and CAMP's sparse solution dataset $\mathcal{D}_{\mathrm{Sp}}$ by 5.09% and 0.72% mAP, respectively. Examples of the classification results are shown in Fig. 14. Compared with Faster RCNN, YOLOv3 has less accuracy, but has faster detection speed.

In this section, we conduct four groups of twelve comparisons. From the experimental results, it is found that by using both Faster RCNN and YOLOv3 methods, whether under SOC or

EOC, the CAMP's nonsparse solution dataset $\mathcal{D}_{\mathrm{Nsp}}$ has shown optimal performance in target detection and classification. In addition, when using $\mathcal{D}_{\mathrm{Nsp}}$, it can be seen that Faster RCNN has better performance than YOLOv3 under EOC. In contrast, YOLOv3 works better under SOC, especially in the detection speed, which is desirable to the real-time processing.

V. Conclusion

In this article, we propose a novel target detection and classification framework based on sparse SAR image dataset. First, a novel CAMP-based sparse imaging method is used to obtain the sparse SAR image datasets $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$. Then, two conventional CNN methods, Faster RCNN and YOLOv3 are used for the target detection and classification based on $\mathcal{D}_{\mathrm{Sp}}$ and $\mathcal{D}_{\mathrm{Nsp}}$. Experimental results show that $\mathcal{D}_{\mathrm{Nsp}}$ has optimal performance in CNN-based target detection and classification. Under EOC, the mAP value of Faster RCNN and YOLOv3 are 95.69% and 88.21% mAP, respectively, which is higher than the other two kinds of datasets, MF dataset and $\mathcal{D}_{\mathrm{Sp}}$. These two mAP values even reach 92.60% and 99.29% under SOC, which means that the novel sparse SAR image dataset has much better performance in SAR target detection and classification, and shows a huge application potential for military battlefield in the future.

References

[1] J. C. Curlander and R. N. Mcdonough, *Synthetic Aperture Radar: Systems and Signal Processing*. New York, NY, USA: Wiley, 1991.

[2] F. M. Henderson and A. J. Lewis, "*Principle and Application of Imaging Radar*. New York, NY, USA: Wiley, 1998.

[3] C. Clemente and J. J. Soraghan, "Vibrating target micro-doppler signature in bistatic SAR with a fixed receiver," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3219–3227, Aug. 2012.

[4] D. Cerutti-Maori *et al.*, "Precision SAR processing using chirp scaling," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 3019–3030, Jul. 2008.

[5] S. Singha, T. J. Bellerby, and O. Trieschmann, "Satellite oil spill detection using artificial neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2355–2363, Dec. 2013.

[6] M. Neumann, L. Ferro-Famil, and A. Reigber, "Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1086–1104, Mar. 2010.

[7] R. Bamler, "A comparsion of range-doppler and wavenamber domain SAR focusing algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 706–713, Jul. 1992.

[8] Y. L. Neo, F. H. Wong, and I. G. Cumming, "Processing of azimuth-invariant bistatic SAR data using the range doppler algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 14–21, Jan. 2008.

[9] R. K. Raney, H. Runge, R. Bamler, I. G. Cumming, and F. H. Wong, "Precision SAR processing using chirp scaling," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 786–799, Jul. 1994.

[10] J. Mittermayer, R. Lord, and E. Borner, "Sliding spotlight SAR processing for TerraSAR-X using a new formulation of the extended chirp scaling algorithm," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Toulouse, France, 2003, pp. 1462–1464.

[11] F. H. Wong, I. G. Cumming, and Y. L. Neo, "Focusing bistatic SAR data using the nonlinear chirp scaling algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2493–2505, Sep. 2008.

[12] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Norwood, MA, USA: Artech House, 2004.

[13] B. Zhang, W. Hong, and Y. Wu, "Sparse microwave imaging: Principles and applications," *Sci. China Inf. Sci.*, vol. 55, no. 8, pp. 1–33, 2012.

[14] R. G. Baraniuk *et al.*, "Applications of sparse representation and compressive sensing," *Proc. IEEE.*, vol. 98, no. 6, pp. 906–909, Jun. 2010.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] E. T. Candes, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[17] E. T. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure. Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[18] I. Daubechies, M. Defriese, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[19] H. Bi and G. Bi, "A novel iterative soft thresholding algorithm for L_1 regularization based SAR image enhancement," *Sci. China Inf. Sci.*, vol. 62, no. 4, pp. 1–3, 2019.

[20] H. Bi and G. Bi, "Performance analysis of iterative soft thresholding algorithm for L_1 regularization based sparse SAR imaging," in *Proc. IEEE Radar Conf.*, Boston, MA, USA, 2019, pp. 1–6.

[21] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Proc. Asilomar Conf. Signal Sys. Comput.*, Pacific Grove, CA, USA, 1993, pp. 40–44.

[22] D. L. Donoho, Y. Tsaig, I. Drori, and J. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.

[23] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, 2009.

[24] A. Maleki, L. Anitori, Z. Yang, and R. G. Baraniuk, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4290–4308, Jul. 2013.

[25] L. Anitori, A. Maleki, M. Otten, R. G. Baraniuk, and P. Hoogeboom, "Design and analysis of compressed sensing radar detectors," *IEEE Trans. Signal process.*, vol. 61, no. 4, pp. 813–827, Feb. 2013.

[26] H. Bi, B. Zhang, X. Zhu, W. Hong, J. Sun, and Y. Wu, "L_1 regularization based SAR imaging and CFAR detection via complex approximated message passing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3426–3440, Jun. 2017.

[27] C. Shan *et al.*, "Gesture recognition using temporal template based trajectories," in *Proc. 17th Int. Conf. Pattern Recognit.*, Cambridge, Britain, 2004, pp. 954–957.

[28] R. Ahmmed and M. F. Hossain, "Tumor detection in brain MRI image using template based k-means and fuzzy c-means clustering algorithm," in *Proc. IEEE Int. Conf. Comput. Commun. Inform.*, Coimbatore, India, 2016, pp. 1–6.

[29] J. Zhu, X. Qiu, Z. Pan, Y. Zhang, and B. Lei, "Projection shape template-based ship target recognition in TerraSAR-X images," *IEEE Trans. Geosci. Remote Sens.*, vol. 14, no. 2, pp. 222–226, Feb. 2017.

[30] G. Magna *et al.*, "Adaptive classification model based on artificial immune system for breast cancer detection," in *Proc. IEEE AISEM Annu. Conf.*, Trento, Italy, 2015, pp. 1–4.

[31] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag.*, vol. 8, no. 6, pp. 3–11, Dec. 1988.

[32] A. Sheikhi, A. Zamani, and Y. Norouzi, "Model-based adaptive target detection in clutter using MIMO radar," in *Proc. CIE Inter. Conf. Radar*, Shanghai, China, 2006, pp. 1–4.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, Trento, Italy, 2015, pp. 1–4.

[34] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.

[35] R. Girshick, "Fast R-CNN," in *Proc. IEEE 5th Int. Conf. Comput. Vis.*, Santiago, Spain, 2015, pp. 1440–1448.

[36] R. Girshick *et al.*, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, America, 2014, pp. 580–587.

[37] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[38] R. Girshick *et al.*, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, America, 2016, pp. 779–788.

[39] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, America, 2017, pp. 6517–6525.

[40] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," Tech. Rep., pp. 1–6, 2018.
[41] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comp. Vis.*, EurAmsterdam, Netherlands, 2016, pp. 21–37.
[42] M. Dong *et al.*, "End-to-end target detection and classification with data augmentation in SAR images," in *Proc. IEEE Int. Conf. Comput. Electromagnetics*, Shanghai, China, 2019, pp. 1–3.
[43] M. Kang *et al.*, "A modified faster R-CNN based on CFAR algorithm for SAR ship detection," in *Proc. Int. Workshop Remote Sens. Intell. Process*, Shanghai, China, 2017, pp. 1–4.
[44] N. Wang, Y. Wang, H. Liu, Q. Zuo, and J. He, "Feature-fused SAR target discrimination using multiple convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1695–1699, Oct. 2017.
[45] H. Bi, G. Bi, B. Zhang, and W. Hong, "Complex-image-based sparse SAR imaigng and its equivalence,"*IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5006–5014, Sep. 2018.
[46] M. Çetin, W. C. Karl, and D. A. Castanon, "Feature enhancement and ATR performance using nonquadratic optimization-based SAR imaging," *IEEE Trans. Aero. Elec. Sys.*, vol. 39, no. 4, pp. 1375–1395, Oct. 2003.
[47] K. He *et al.*, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, America, 2016, pp. 770–778.
[48] T. Lin *et al.*, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, America, 2017, pp. 936–944.

Hui Bi (Member, IEEE) was born in Shandong, China, in 1991. He received the bachelor's degree in electronics and information engineering from YanTai University, Yantai, China, in 2012, and the Ph.D. degree in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2017.

From 2012 to 2017, he was with the Science and Technology on Microwave Imaging Laboratory, Institute of Electronics, Chinese Academy of Sciences, China. He was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2017 to 2018. Since 2018, he has been working in the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China as an Associate Professor. His main research interests include sparse microwave imaging with compressive sensing, synthetic aperture radar data processing and application, sparse signal processing, and tomographic SAR imaging.

Jiarui Deng was born in Jiangsu, China, in 1997. She received the bachelor's degree in electronics and information engineering from the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2020. She is currently working toward the master's degree in signal and information processing with Nanjing University of Aeronautics and Astronautics, Nanjing.

Her research interest includes sparse SAR image processing and application.

Tianwen Yang was born in Jiangsu, China, in 1998. She received the bachelor's degree in electronics and information engineering from the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2020. She is currently working toward the master's degree in signal and information processing with Southeast University, Nanjing.

Her research interests include communication and information system.

Jian Wang was born in Anhui, China, in 1999. He received the bachelor's degree in electronics and information engineering from the College of Automation, Chongqing University, Chongqing, China. He is currently working toward the master's degree in signal and information processing with Nanjing University of Aeronautics and Astronautics, Nanjing, China.

His research interest includes sparse SAR image processing and application.

Ling Wang (Member, IEEE) received the B.S. degree in electrical engineering, and the M.S. and Ph.D. degrees in information acquirement and processing from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2000, 2003, and 2006, respectively.

Since 2003, she has been with the Nanjing University of Aeronautics and Astronautics, where she is currently a Professor with the Department of Information and Communication Engineering. From February 2008 to May 2009, she was a Postdoctoral Research Associate with the Department of Mathematical Sciences and the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. She has authored and coauthored more than 100 publications. Her current research interests include radar, imaging problems, image processing in vision-based navigation, and image-based target reconstruction and recognition.

Miss Wang was the recipient of the Alexander Humboldt Fellowship for Experienced Researchers in 2014 and worked with Professor Otmar Loffeld in University of Siegen from 2015 to 2016.