



Article

Complex-Valued Sparse SAR-Image-Based Target Detection and Classification

Chen Song ¹, Jiarui Deng ^{2,*}, Zehao Liu ², Bingnan Wang ¹, Yirong Wu ¹ and Hui Bi ²

¹ National Key Laboratory of Microwave Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China

² The Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

* Correspondence: deng_jr@nuaa.edu.cn

Abstract: It is known that synthetic aperture radar (SAR) images obtained by typical matched filtering (MF)-based algorithms always suffer from serious noise, sidelobes and clutter. However, the improvement in image quality means that the complexity of SAR systems will increase, which affects the applications of SAR images. The introduction of sparse signal processing technologies into SAR imaging proposes a new way to solve this problem. Sparse SAR images obtained by sparse recovery algorithms show better image performance than typical complex SAR images with lower sidelobes and higher signal-to-noise ratios (SNR). As the most widely applied fields of SAR images, target detection and target classification rely on SAR images with high quality. Therefore, in this paper, a target detection framework based on sparse images recovered by complex approximate message passing (CAMP) algorithm and a novel classification network via sparse images reconstructed by the new iterative soft thresholding (BiST) algorithm are proposed. Experimental results show that sparse SAR images have better performance whether for target classification or for target detection than the images recovered by MF-based algorithms, which validates the huge application potentials of sparse images.



Citation: Song, C.; Deng, J.; Liu, Z.; Wang, B.; Wu, Y.; Bi, H. Complex-Valued Sparse SAR-Image-Based Target Detection and Classification. *Remote Sens.* **2022**, *14*, 4366. <https://doi.org/10.3390/rs14174366>

Academic Editor: Timo Balz

Received: 28 June 2022

Accepted: 30 August 2022

Published: 2 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: sparse synthetic aperture radar (SAR); convolutional neural network (CNN); complex approximate message passing (CAMP); target classification; target detection

1. Introduction

Synthetic aperture radar (SAR) is a microwave remote sensing observation system which can obtain high-resolution images due to its ability of working all day and under all weather conditions. Thus, it plays an irreplaceable and important role in both the military and civilian fields [1,2]. As important fields of SAR applications, target detection and target classification have developed slowly due to the problems of hardware equipment and manual feature extraction [3–5]. In 2006, Hinton et al. proposed the concept of deep learning, noting that the multi-layer convolutional neural network (CNN) has the ability of automatic feature extraction, which is of great research significance for target classification [6]. In 2012, Krizhevsky et al. proposed the first CNN model used for image classification, with its top-five error ratio only being 17.0%, making deep learning popular in the field of image classification [7,8]. Chen et al. [9] used an unsupervised sparse self-encoder in place of typical back-propagation in 2014, so as to apply CNN to SAR image recognition. Since then, more attention has been paid to radar image processing based on CNN [10,11]. After CNN began being widely used in the field of image processing, scholars started to carry out considerable research on whether CNN can show comparable performance in the field of target detection. At present, CNN-based target detection frameworks can be mainly divided into two types: one-stage frameworks [12–14] and two-stage frameworks [15–18]. The first two-stage framework, called a regional convolutional neural network (RCNN), was introduced by Girshick et al. in 2014 [15]. Then, the improved RCNN series such as

Fast RCNN [16], Faster RCNN [17] and Mask RCNN [18] were also proposed. Compared with two-stage frameworks, one-stage frameworks can directly output the locations and categories of different targets, which improve the efficiency of target detection. The algorithms represented by YOLO (You Only Look Once) series and single-shot multibox detectors (SSD) [12] are widely used in target detection. In 2015, Redmon et al. [13] proposed a YOLO algorithm based on a single neural network, which turns target detection into a regression problem and uses the entire information of images to predict the target locations and categories. Then, in 2020, Bochkovskiy et al. developed a target detection model called YOLOv4 [14]. It integrates several advanced target detection techniques, which facilitates real-time processing and improves the target detection ability of YOLO algorithms.

When it comes to target classification, different from optical images, SAR images have both amplitude and phase information. However, traditional CNNs only use amplitude information for SAR target classification, similar to the target classification of optical images. Many research methods on how to use the phase information of SAR images have been proposed in recent years, which show better performance than amplitude-based methods [19–21]. Zhang et al. proposed a novel classification network called complex-valued CNN (CV-CNN) in 2017 to extract the phase information of SAR images [19]. Compared with typical CNNs that only use amplitude information, CV-CNN achieves lower classification error rate in experiments based on polarimetric SAR datasets. In 2018, Coman et al. adopted the method of amplitude–real–imaginary three-layer data to form the input of the network and achieved about 90% accuracy in the experiments based on the MSTAR dataset, which alleviated the over-fitting problem caused by the lack of training data [20]. In 2020, Yu et al. proposed a new framework on the basis of CV-CNN, named the complex-valued fully convolutional neural network (CV-FCNN), in which the pooling layers and fully connected layers are replaced by convolutional layers to avoid complex pooling operation and over-fitting [21]. Experiments on MSTAR demonstrated that CV-FCNN improves the classification accuracy, showing better performance over CV-CNN.

In the last decade or so, sparse signal processing technologies were widely used in SAR imaging. The limitation of conventional Shannon–Nyquist sampling theory is broken by sparse SAR imaging algorithms. Benefiting from this, the sparse SAR imaging algorithms can achieve the high-quality recovery of sparse scenes with less data, reducing the complexity of radar systems [22,23]. In sparse SAR imaging, typical sparse recovery algorithms, such as orthogonal matching pursuit (OMP) [24,25] and iterative soft thresholding (IST) [26,27], could only improve the image quality of observed scenes with ruined background statistical characteristics and phase information. This will lose the feature information of focused targets and further hinder the development of sparse SAR image processing. The introduction of complex approximate message passing (CAMP) [28,29] and a novel iterative soft thresholding algorithm (BiIST) [30,31] into sparse SAR imaging solves these problems. Both CAMP and BiIST-based sparse SAR imaging methods can acquire two kinds of sparse SAR images, i.e., sparse solution and non-sparse solution of the interested scenes. The sparse solution of BiIST and CAMP is similar to the results of typical sparse reconstruction algorithms. However, non-sparse solutions of CAMP can retain similar background statistical distributions to matched filtering (MF)-based images and the non-sparse solution of BiIST can retain phase information, which will offer more feature information for SAR target classification and detection, so as to theoretically improve the performance of proposed methods.

In this paper, experiments will be carried out based on the sparse SAR images recovered by the CAMP and BiIST algorithms. In the case of the CAMP-based sparse SAR imaging method, a target detection framework based on sparse SAR images is introduced. It firstly constructs the sparse SAR image dataset using the results of CAMP. Then, YOLOv4 is introduced into target detection on the basis of reconstructed datasets. When it turns to BiIST-based sparse SAR imaging method, a classification network based on the amplitude and phase information of SAR images is used to classify targets in a sparse SAR image dataset composed of the non-sparse solution of BiIST.

The rest of this paper is organized as follows. Sparse SAR imaging methods based on complex image data are introduced in Section 2. Section 3 describes the models of YOLOv4 and the amplitude–real–imaginary classification network. Experimental results on the basis of the MSTAR dataset are shown in Section 4, and performance analysis under different situations is discussed in Section 5. Finally, Section 6 concludes our work.

2. Sparse SAR Image Recovery Algorithm

In order to protect the copyright of echo data and the system confidentiality of SAR, compared with original data, complex SAR image data recovered by MF-based methods is more convenient to obtain. Therefore, to get abundant sparse SAR images, sparse SAR imaging methods based on complex image data via regularization-based algorithms are introduced, i.e., CAMP [28,29] and BiIST [30,31]. The complex-image-based sparse SAR imaging model can be write as:

$$\mathbf{X}_{MF} = \mathbf{X} + \mathbf{N} \quad (1)$$

where $\mathbf{X} \in \mathbb{C}^{N_p(\text{Azimuth}) \times N_Q(\text{Range})}$ is the back-scattering coefficient of interested scenes, \mathbf{X}_{MF} is the known complex-valued SAR image data and $\mathbf{N} \in \mathbb{C}^{N_p \times N_Q}$ is a complex matrix representing the difference between \mathbf{X}_{MF} and \mathbf{X} .

2.1. CAMP Based Sparse SAR Recovery Algorithm

Based on the model shown in (1), \mathbf{X} can be recovered from the following regularization problem with regularization parameter λ , i.e.,

$$\hat{\mathbf{X}}_{CAMP} = \min_{\mathbf{X}} \left\{ \frac{1}{2} \|\mathbf{X}_{MF} - \mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_1 \right\} \quad (2)$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix. Then CAMP is introduced to deal with the optimization problem in (2), and the specific iterative process is listed in [28,29]. Compared with other recovery algorithms, CAMP introduces a term of “state evolution” to evolve the standard deviation of “noise” during the iteration process, so that it can output the sparse $\hat{\mathbf{X}}_C$ and non-sparse $\tilde{\mathbf{X}}_C$ estimations of considered scenes, simultaneously. Different from typical sparse images $\hat{\mathbf{X}}_C$, $\tilde{\mathbf{X}}_C$ not only has an improved image quality compared with MF-based results, but it can also preserve the statistical characteristics in the image background.

To support our viewpoints, experiments based on the MSTAR dataset are used to validate the CAMP-based algorithm in image performance improvement and the preservation of image backgrounds’ statistical distributions. Figure 1 shows the image of five point targets recovered by MF- and CAMP-based sparse SAR imaging methods, respectively. Compared with MF-based result, it should be noted that $\hat{\mathbf{X}}_C$ highlights the target with ruined background statistical characteristics, as shown in Figure 1b. However, $\tilde{\mathbf{X}}_C$ (Figure 1c) has a relatively close background distribution to that of the MF-based images (see in Figure 2d) with improved image quality. The target-to-background ration (TBR) is selected to evaluate the improvement of image quality quantitatively, which can be defined as:

$$\text{TBR}(\mathbf{X}) \triangleq 20 \log_{10} \left(\frac{\max_{(u,v) \in \mathcal{T}} |(\mathbf{X})_{(u,v)}|}{(1/N_B) \sum_{(u,v) \in \mathcal{B}} |(\mathbf{X})_{(u,v)}|} \right) \quad (3)$$

where \mathcal{T} is the target area, and \mathcal{B} represents the background region. Then, N_B denotes the number of pixels in the background region. The TBR of five targets in Figure 1 is shown in Table 1. It is seen that both $\hat{\mathbf{X}}_C$ and $\tilde{\mathbf{X}}_C$ of the five selected targets can obtain a higher TBR, with all of them reaching 50 dB, which shows better image performance than MF-based images.

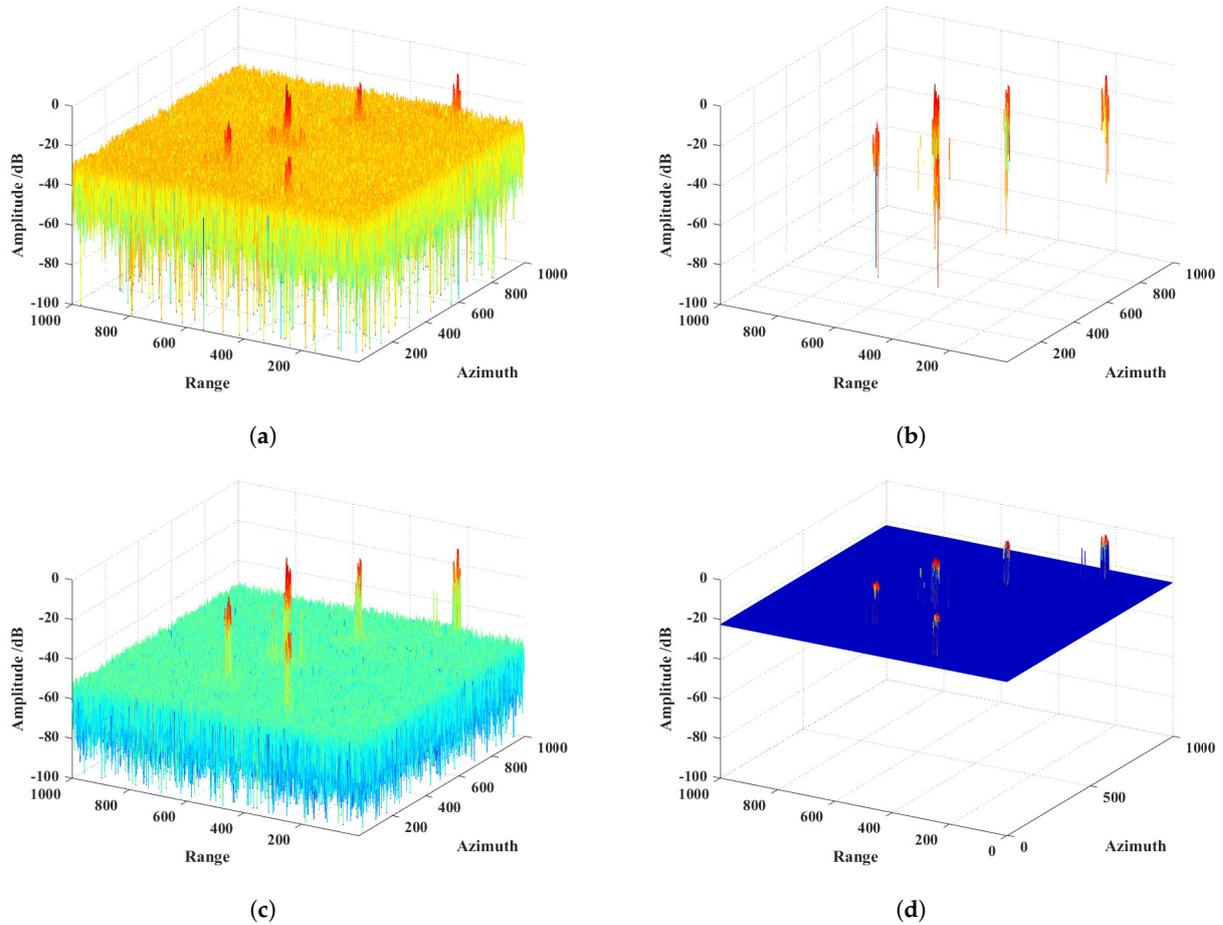


Figure 1. Images restored by different methods: (a) MF-based; (b) \hat{X}_C of CAMP; (c) \tilde{X}_C of CAMP; (d) Amplitude deviation between (c) and (a).

Table 1. TBR of images reconstructed by MF and CAMP-based methods.

Target	Target 1	Target 2	Target 3	Target 4	Target 5
MF	38.20 dB	35.71 dB	34.11 dB	33.30 dB	34.75 dB
\hat{X}_C of CAMP	57.30 dB	55.04 dB	52.26 dB	53.39 dB	53.86 dB
\tilde{X}_C of CAMP	60.73 dB	54.54 dB	55.17 dB	52.77 dB	57.53 dB

2.2. BiIST-Based Sparse SAR Recovery Algorithm

When it comes to the BiIST-based algorithm, the following optimization problem based on the model in (1) is to be solved by:

$$\hat{X}_{BiIST} = \min_X \left\{ \|X_{MF} - X\|_F^2 + \beta \|X\|_1 \right\} \quad (4)$$

where β is the regularization parameter. The detailed iterative process of BiIST used to deal with the above optimization problem is shown in [30,31]. Complex SAR images recovered by the BiIST-based algorithm can also obtain two kinds of results of the observed scenes, named sparse solutions (\hat{X}_B) and non-sparse solutions (\tilde{X}_B). As shown in Figure 2, the sparse solution \hat{X}_B and non-sparse solution \tilde{X}_B are similar to the result of the CAMP-based sparse recovery algorithm. However, it should be noted that the phase information of \hat{X}_C recovered by CAMP is ruined, facing the same problem that phase information cannot be introduced into the applications of SAR images. Different from the CAMP-based method, \tilde{X}_B of BiIST-based algorithm can retain the phase information, which will provide more effective information for the SAR image applications.

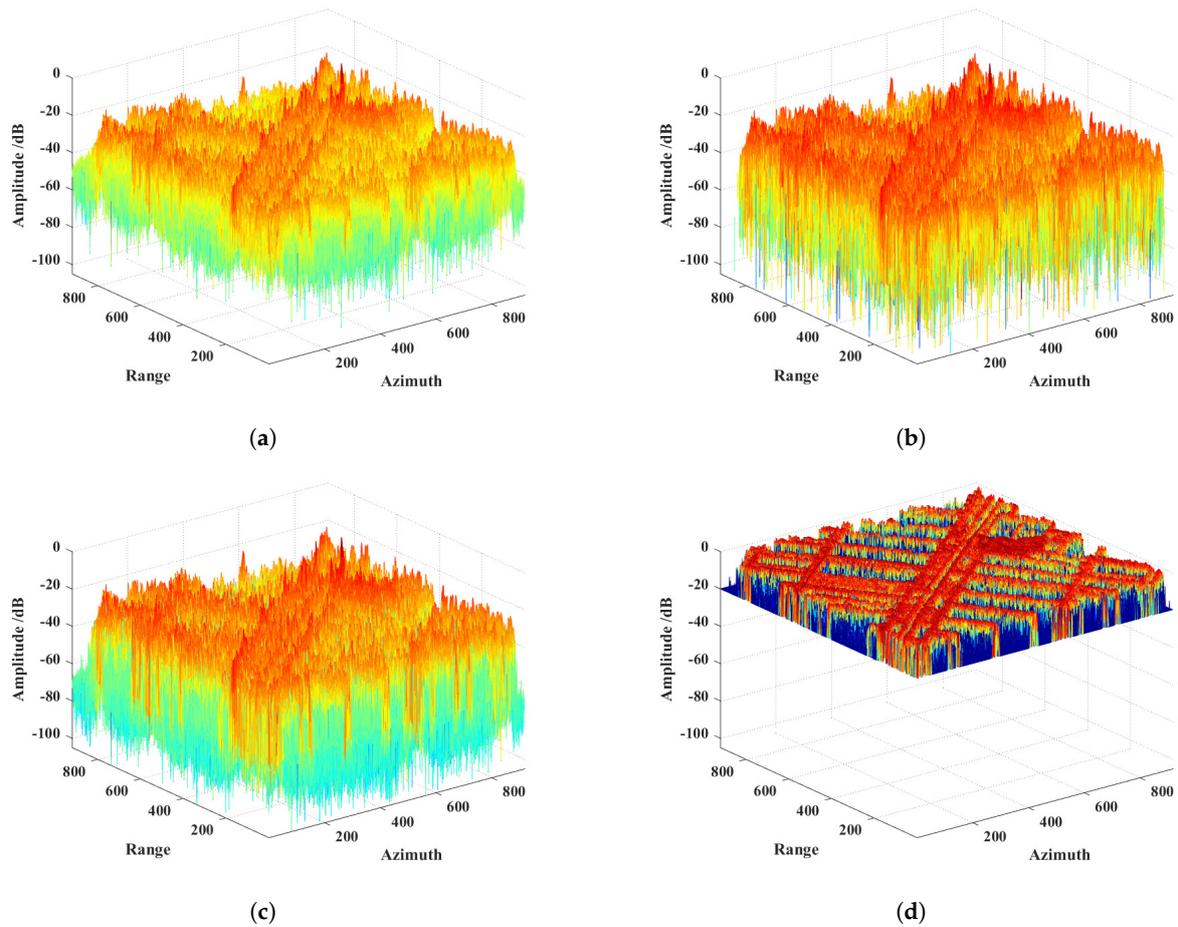


Figure 2. Images restored by different methods: (a) MF; (b) \hat{X}_B of BiIST; (c) \tilde{X}_B of BiIST; (d) Amplitude deviation between (c) and (a).

Take the real scene as an example. The phase difference between an MF-based image and the recovered sparse image is shown in Figure 3. As represented in Figure 3b, the phase difference between the MF-based image and the non-sparse solution is zero, different from that between the MF-based image and the sparse solution (see Figure 3a). Therefore, \tilde{X}_B can retain the phase information of MF-based images, making phase-based SAR image applications such as SAR interferometry (InSAR), constant false alarm rate (CFAR) detection and the proposed classification network possible.

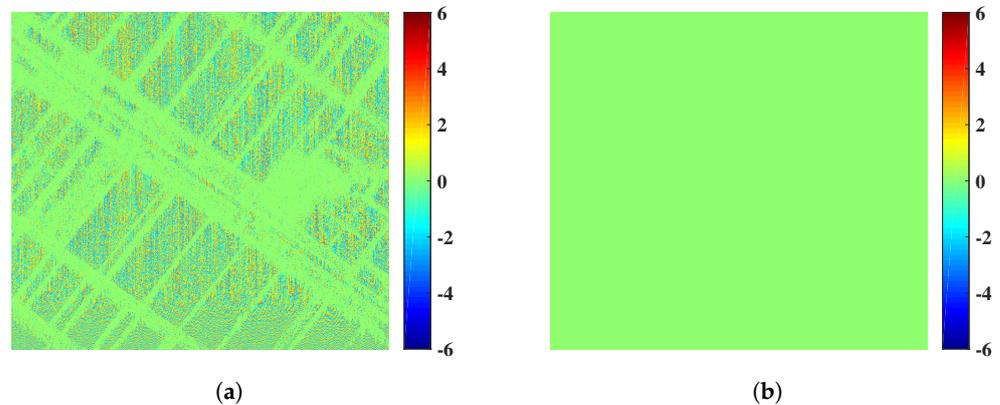


Figure 3. The phase difference between an MF-based image and (a) sparse solution \hat{X}_B ; (b) non-sparse solution \tilde{X}_B .

3. SAR Image Application Based on Sparse SAR Images

3.1. Target Detection by YOLOv4

YOLOv4 integrates several advanced target detection techniques to improve the accuracy and accelerate the training of CNNs, whose architecture is shown in Figure 4. The low requirement of hardware equipment makes YOLOv4 able to be widely used. YOLOv4 is made up of four parts. The main methods and tricks in each part are shown as follows [13]:

- **Input:** Mosaic data augmentation, cross mini-batch normalization (CmBN) and self-adversarial training (SAT);
- **Backbone:** Cross-stage partial connections Darknet53 (CSPDarknet53), mesh-activation and dropblock regularization;
- **Neck:** SPP, modified feature pyramid network (FPN) and path aggregation network (PAN);
- **Prediction:** Modified complete IOU (C-IOU) loss and distance IOU (D-IOU) nms.

Parts of these methods are able to obviously improve the network's performance. Mosaic data augmentation randomly utilizes four images with clipping and scaling and then splices them, so as to enrich the dataset, reducing the size of the mini-batch and the requirement for GPU computing. SAT changes original images to create the deception of an undesired target. Methods in the input part are used to modify images which can strengthen the robustness of the network for training in the next part. Cross-stage partial connections in the second part of YOLOv4 divide the gradient flow, propagating it into different paths to reduce the computational complexity and accelerate the detection speed of the network. In addition, some modules such as SPP, PAN and FPN improve the feature extraction ability of YOLOv4. Loss functions used in the network, such as modified C-IOU loss and D-IOU nms make YOLOv4 achieve better performance.

3.2. Target Classification by Amplitude–Real–Imaginary CNN

Different from traditional amplitude-based classification networks, the three-channel CNN is a novel framework based on the amplitude and phase information of SAR images for target classification. As shown in Figure 5, the architecture of the amplitude–real–imaginary CNN is made up of several convolutional layers, average pooling layers and fully connected layers, which is similar to the typical structure of CNNs. The difference between the amplitude–real–imaginary CNN and amplitude-based CNNs is mainly reflected in the input layer. When using amplitude-based CNNs, input data are one-channel images that only contain amplitude information of SAR images. However, the input layer of proposed framework is a three-channel image with the amplitude layer, real layer and imaginary layer (see in Figure 6), so as to utilize both amplitude and phase information during the training process. Batch normalization is processed after the convolution operation to accelerate the convergence speed of network and improve the stability of training process. The activation function used in this framework is ReLU, which can relieve the problems of gradient vanishing and gradient explosion.

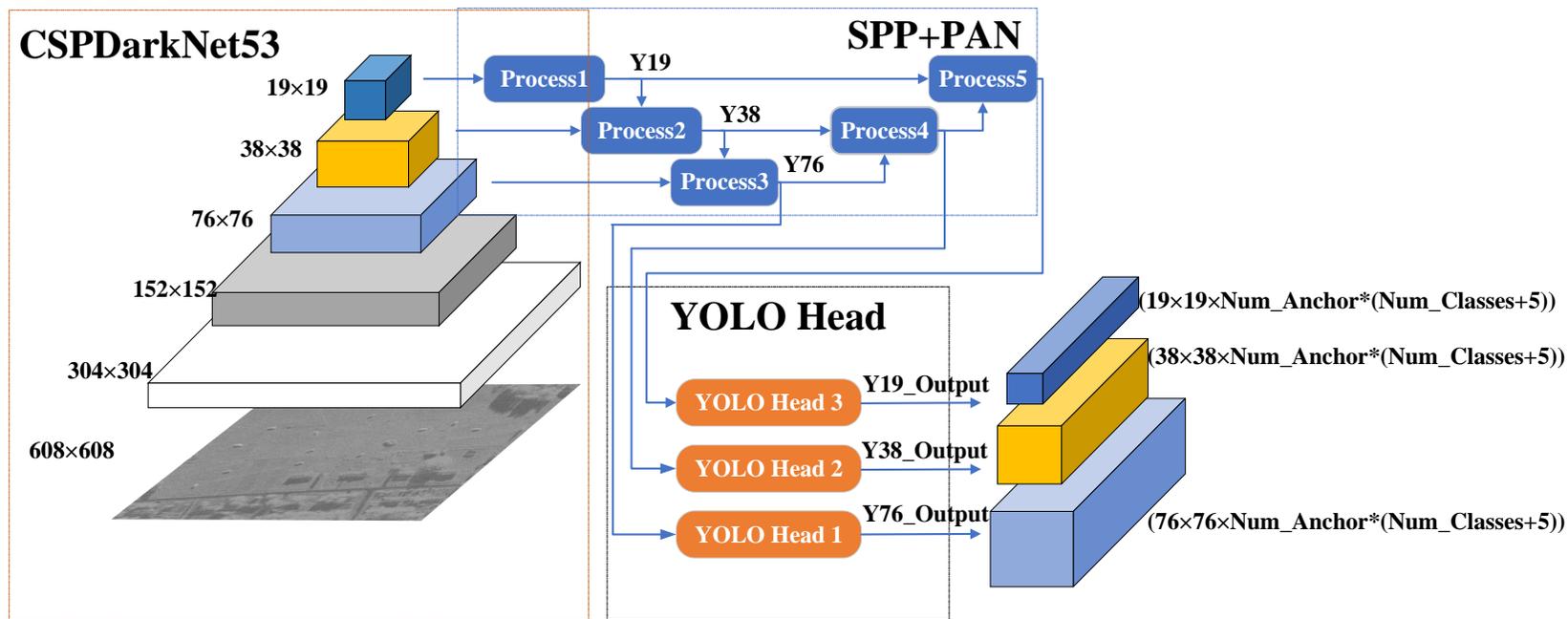


Figure 4. The architecture of YOLOv4 [13].

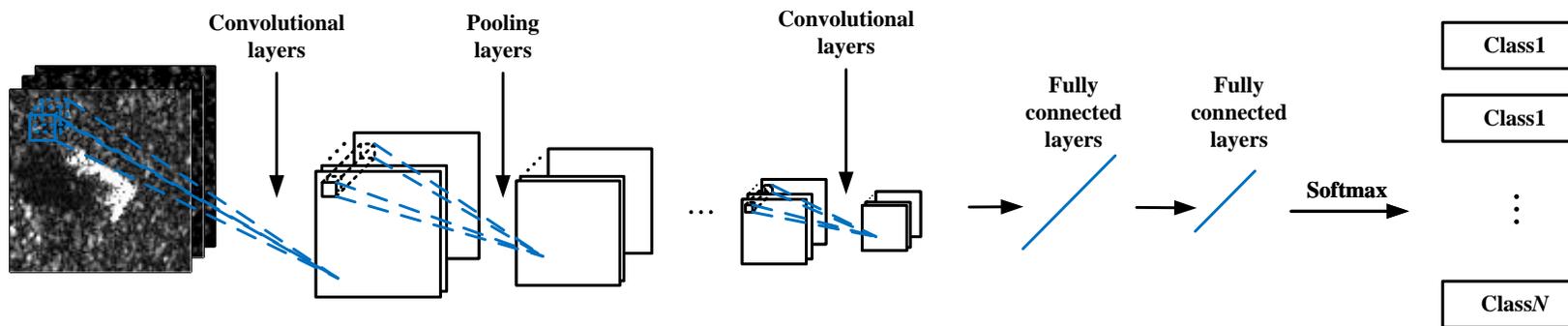


Figure 5. Architecture of amplitude-real-imaginary CNN.

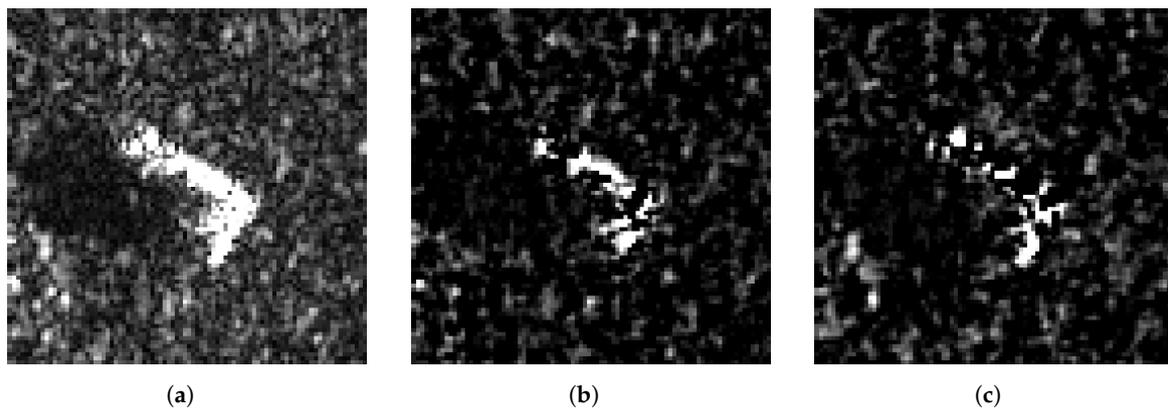


Figure 6. Input data of the amplitude-real-imaginary CNN: (a) Amplitude layer; (b) Real layer; (c) Imaginary layer.

4. Experiments Based on MSTAR Dataset

The dataset used in experiments is constructed based on the military vehicle samples in MSTAR, which consists of ten kinds of targets and can be divided into standard operating conditions (SOC) and extended operating conditions (EOC). The optical images and corresponding SAR images of the ten category vehicles are shown in Figure 7. In this section, all experiments are carried out both under SOC and EOC.

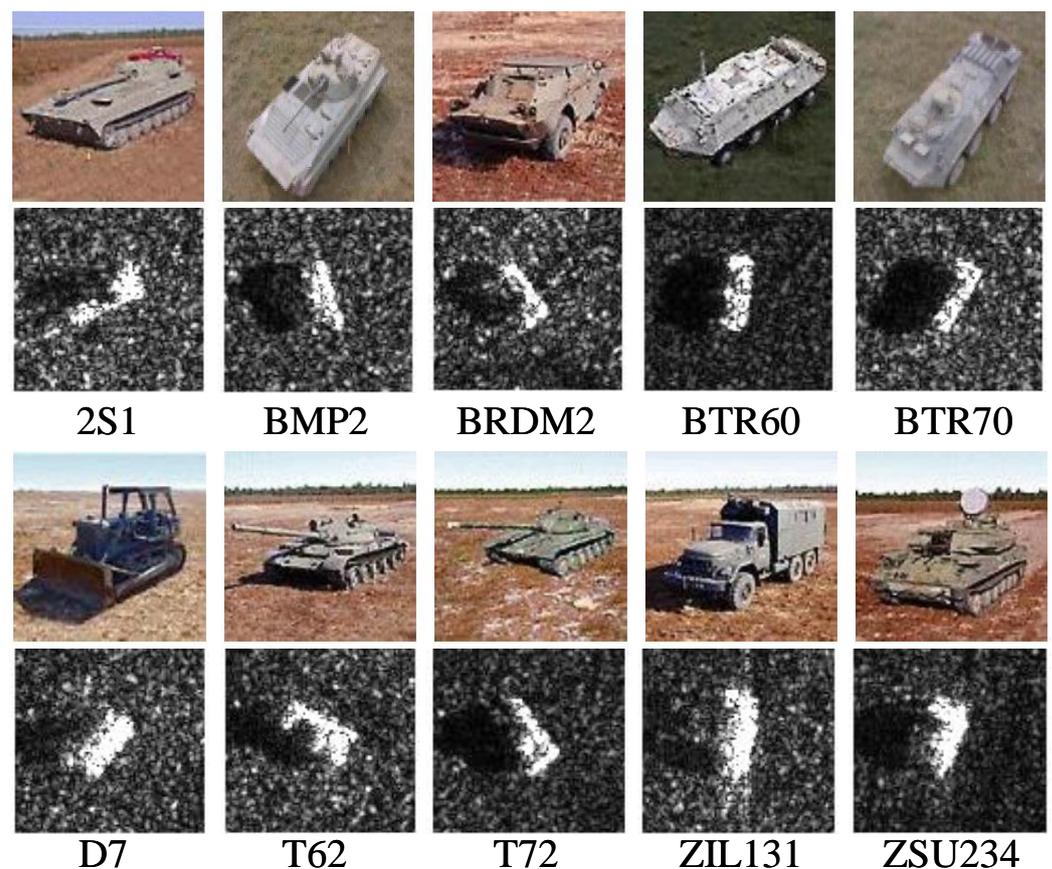


Figure 7. Optical images and corresponding SAR images of vehicles in the MSTAR dataset.

4.1. Experiments under SOC

The data description for SOC is shown in Table 2, with serial numbers, depression angles and numbers of targets in the dataset. From the table, it can be seen that the depression angle of the target in the training set is 17° , and targets for testing are collected at a 15° depression angle.

Table 2. Data description for SOC.

Class	Serial No.	Training Set (17°)	Testing Set (15°)	Scene Set
2S1	B01	299	274	15
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	

4.1.1. Target Detection Based on YOLOv4

CAMP-based sparse SAR image recovery algorithm is utilized in this part to form the input data of YOLOv4 framework for target detection. For background fusion, all of the target slices are randomly merged into 15 different scenes, with 15 targets drawn from various classes in each scene. To support our viewpoints, experiments are conducted according to the metrics of intersection over union (IOU), mean average precision (mAP) and rendered frames per second (FPS). Average precision (AP) is usually used to evaluate the recognition performance of models proportional to the models' performance, and then mAP is the average AP over several validation sets. In this part, the learning rate is set to 0.0013. Figure 8 presents the target detection results based on datasets recovered by MF- and CAMP-based algorithms, respectively. In addition, quantitative experimental results are presented in Table 3. It can be seen from Table 3 that the \tilde{X}_C dataset with a 99.78% mAP and an 89.34% IOU performs almost as well as the MF one. However, the \tilde{X}_C dataset underperforms compared to the MF dataset by 2.56% mAP and 11.35% IOU.

Table 3. Target detection under SOC.

Dataset	Category										mAP	IOU	FPS
	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234			
MF	100.0%	99.72%	100.0%	99.49%	99.71%	100.0%	100.0%	99.69%	100.0%	99.97%	99.88%	88.43%	53
\tilde{X}_C	98.98%	92.98%	99.09%	93.49%	93.11%	99.24%	99.27%	97.87%	99.19%	99.96%	97.32%	77.08%	
\tilde{X}_C	99.99%	99.73%	100.0%	98.87%	99.35%	99.97%	99.98%	99.88%	100.0%	99.99%	99.78%	89.34%	

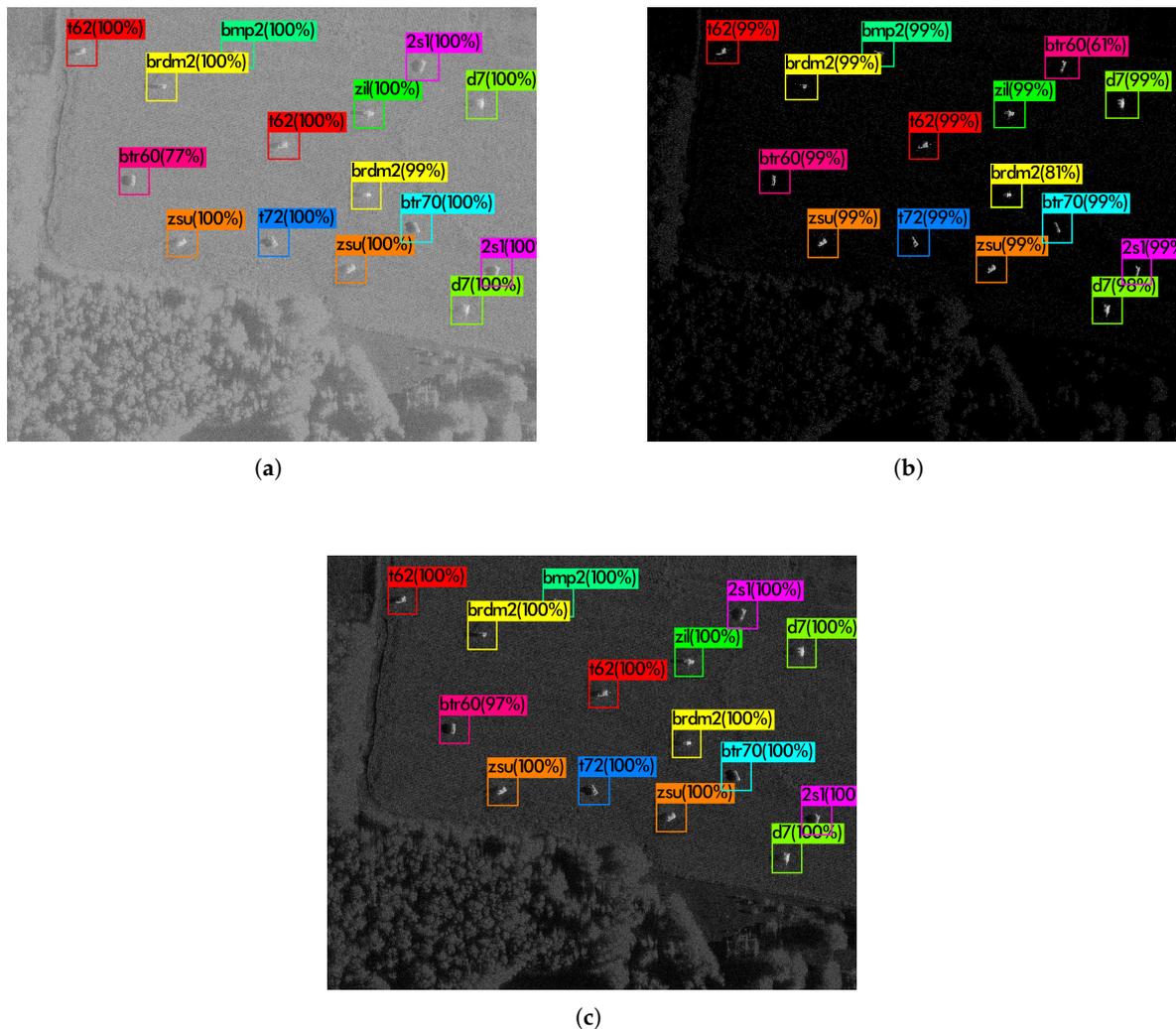


Figure 8. The presentation of target detection under SOC by using different datasets. (a) MF dataset. (b) \hat{X}_C dataset. (c) \tilde{X}_C dataset.

4.1.2. Target Classification Based on Amplitude–Real–Imaginary CNN

In this part, the non-sparse solution of BiST is used to compose the constructed sparse datasets. Several experiments under different numbers of training samples are going on to verify the classification performance of the three-channel CNN. In order to evaluate the proposed framework in detail, a comparison based on classification networks and input datasets are carried out.

1. Comparison based on networks

In order to validate the target classification performance of the proposed amplitude–real–imaginary CNN, experiments based on an amplitude-based CNN and a three-channel CNN are carried out. The number of training samples ranges from 100 to 200 per class. Experimental results are shown in Figure 9, and corresponding classification accuracies are listed in Table 4. From Figure 9, it can be seen that the proposed framework achieves higher accuracy than the amplitude-based CNN based on a sparse SAR image dataset no matter the number of training samples. When the training samples reduce to 1000, which means 100 of each category, the classification accuracy based on the amplitude–real–imaginary CNN can still reach 95.47% (see in Table 4).

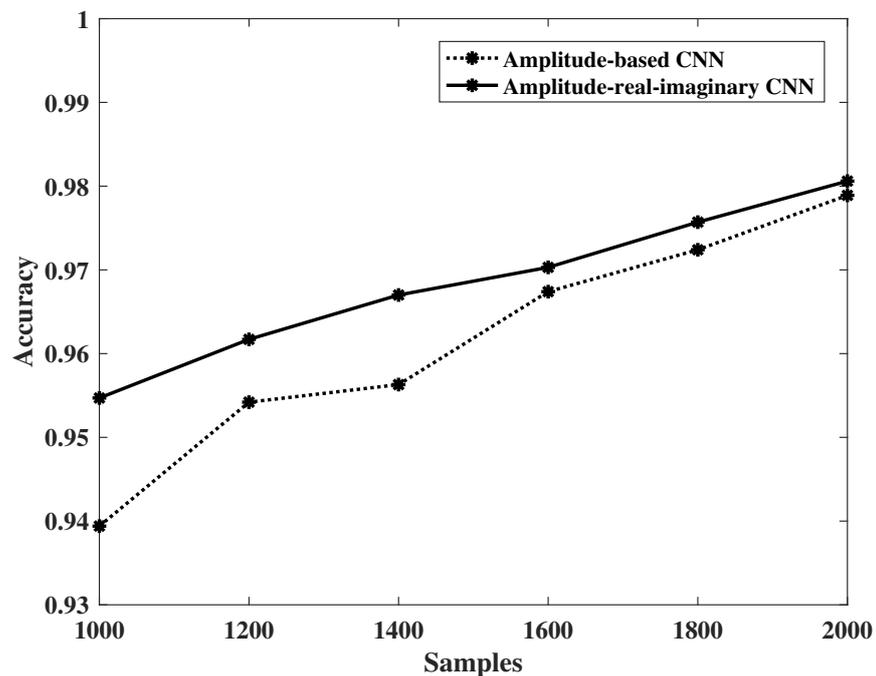


Figure 9. Experimental results based on different networks under SOC.

Table 4. Classification accuracy based on different networks under SOC.

Samples	1000	1200	1400	1600	1800	2000
Amplitude-based CNN	93.94%	95.42%	95.63%	96.74%	97.24%	97.89%
Amplitude-real-imaginary CNN	95.47%	96.17%	96.70%	97.03%	97.57%	98.06%

2. Comparison based on input datasets

Experiments based on the MF-based image dataset and sparse SAR image dataset are carried out as well. In this part, the amplitude–real–imaginary CNN is used to classify military vehicle targets in the datasets. The classification accuracy is listed in Table 5, and Figure 10 visually shows the experimental results. It can be seen from Figure 10 that sparse SAR images recovered by the BiIST-based algorithm show better performance than MF-based images for target classification under different numbers of training samples. When the number of training sample comes to 1000, shown in Table 5, sparse SAR images improve the accuracy by 1.45% compared with MF-based recovery images.

Table 5. Classification accuracy based on different input datasets under SOC.

Samples	1000	1200	1400	1600	1800	2000
bfMF	94.02%	95.42%	96.49%	96.94%	96.95%	97.36%
\tilde{X}_B	95.47%	96.17%	96.70%	97.03%	97.57%	98.06%

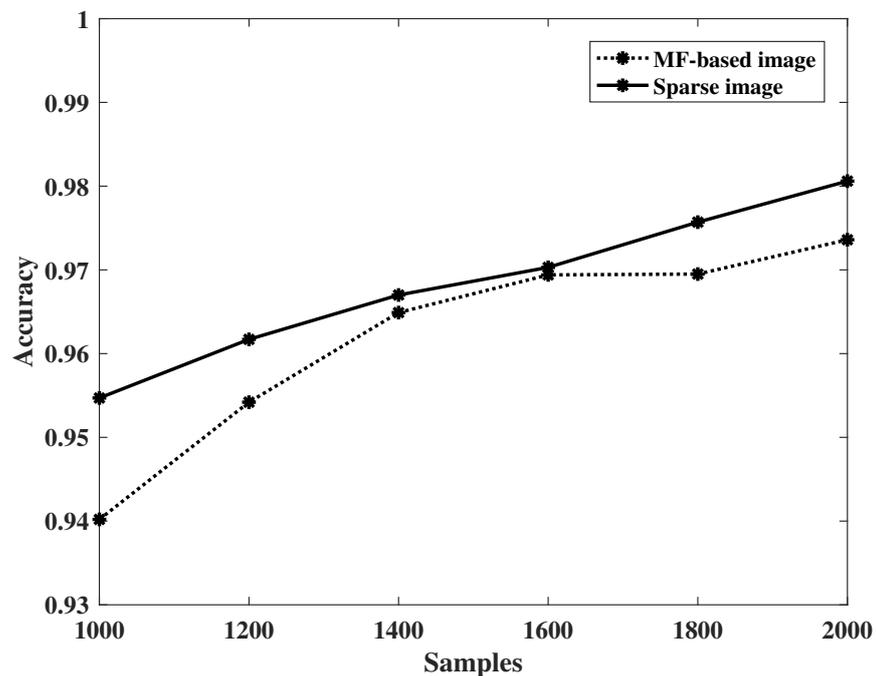


Figure 10. Experimental results based on different input datasets under SOC.

3. Analysis

As shown in the above experiments, it can be found that the combination of sparse SAR images and the proposed three-channel CNN shows optimal performance in contrast to others. The classification accuracy with 2000 samples based on the proposed network and a sparse dataset can reach 98% under SOC, 0.15% and 0.7% higher than the amplitude-based CNN with a sparse dataset and the three-channel CNN with an MF-based dataset, respectively. The confusion matrix in Table 6 presents the accuracy of each category, which shows that even if the accuracy of 2S1 is the lowest among the ten classes, it can still reach 93.79%.

Table 6. Confusion matrix of the amplitude–real–imaginary CNN based on sparse images under SOC.

Class	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234	Total
2S1	257	0	0	1	0	0	2	0	2	0	
BMP2	7	195	1	0	0	0	1	0	0	0	
BRDM2	2	0	262	0	0	0	0	0	0	1	
BTR60	0	0	3	193	1	0	0	0	0	0	
BTR70	2	0	0	1	195	0	0	0	0	0	
D7	0	0	0	0	0	272	1	0	1	0	
T62	2	0	0	0	0	0	266	0	0	1	
T72	2	1	2	0	0	0	0	196	0	0	
ZIL131	1	0	6	0	0	2	3	0	271	0	
ZSU234	1	0	0	0	0	0	0	0	0	272	
Accuracy (%)	93.79	99.49	95.62	98.97	99.49	99.27	97.44	100.00	98.91	99.27	98.06

4.2. Experiments under EOC

To demonstrate the superiority of the proposed frameworks, similar experiments are carried out under EOC. Unlike SOC, samples in the training and testing sets under EOC differ a lot in the aspect of the depression angle. In addition, the serial numbers of several targets are also different in the training set and testing set, which will increase the difficulty of target detection and classification. The data description of the samples under EOC is shown in Table 7.

Table 7. Data description for EOC.

Class	Serial No.	Training Set (17°)	Testing Set (30°)
2S1	B01	299	288
BRDM2	E-71	298	287
T72	SN132/A64	299	288
ZSU234	D08	299	288
Total		1195	1151

4.2.1. Target Detection Based on YOLOv4

Similar to the experiments under SOC, Figure 11 shows the detection results of an MF-based image and a sparse SAR image under EOC, with quantitative experimental results being listed in Table 8. The learning rate in this experiment is set to 0.0012. It can be seen that the \hat{X}_C and \tilde{X}_C datasets both show excellent performance in SAR target detection. The \tilde{X}_C dataset can obtain 92.00% mAP and 65.21% IOU, which is higher than the 90.80% mAP and 55.36% IOU of the MF one. The \hat{X}_C dataset achieves the best performance with a 96.32% mAP and a 70.85% IOU, which outperforms the MF-based dataset by 5.52% mAP and 15.49% IOU. According to above results, it is seen that both non-sparse and sparse solutions of CAMP have superior detection performance in the case of large angle differences under EOC.

Table 8. Target detection under EOC.

Dataset	Category				mAP	IOU	FPS
	2S1	BRDM2	T72	ZSU234			
MF	91.85%	92.45%	85.31%	93.60%	90.80%	55.36%	
\hat{X}_C	96.48%	99.09%	93.45%	96.24%	96.32%	70.85%	48
\tilde{X}_C	92.20%	97.99%	87.46%	90.34%	92.00%	65.21%	

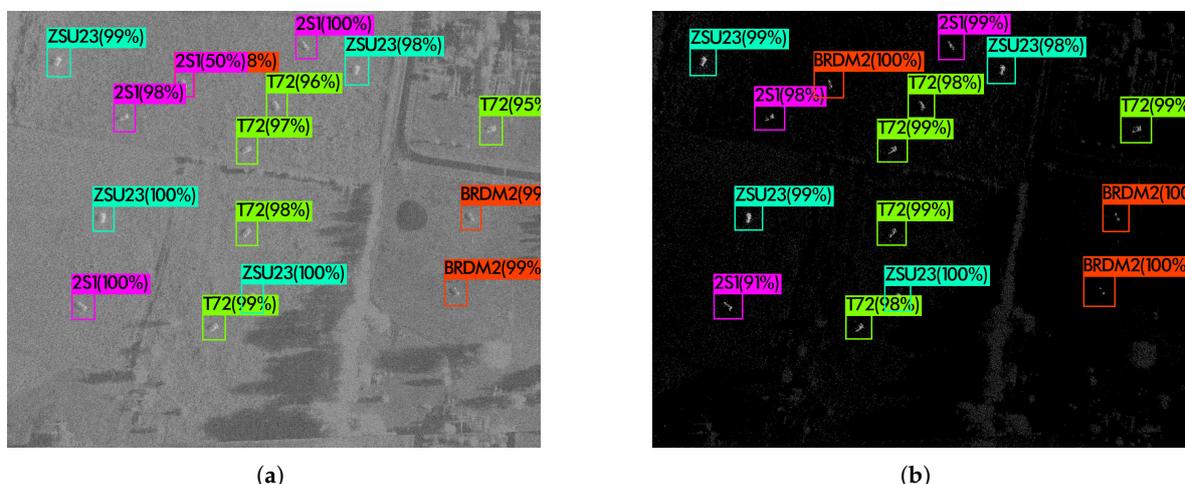


Figure 11. Cont.



(c)

Figure 11. The presentation of target detection of YOLOv4 under EOC by using different datasets: (a) MF dataset; (b) \tilde{X}_C dataset; (c) \tilde{X}_C dataset.

4.2.2. Target Classification Based on Amplitude–Real–Imaginary CNN

1. Comparison based on networks

The number of training samples ranges from 800 to 1000, with a total of four kinds of military vehicles under EOC. Table 9 and Figure 12 show a comparison of classification accuracy between the amplitude-based CNN and the amplitude–real–imaginary CNN on the basis of the sparse SAR image \tilde{X}_B dataset. It can be found that similar to results under SOC, the \tilde{X}_B dataset performs better than the MF-based dataset at any number of training samples.

Table 9. Classification accuracy based on different networks under EOC.

Samples	800	840	880	920	960	1000
Amplitude-based CNN	84.27%	86.44%	86.19%	87.23%	88.62%	92.79%
Amplitude-real-imaginary CNN	89.23%	90.10%	90.01%	90.96%	90.96%	93.83%

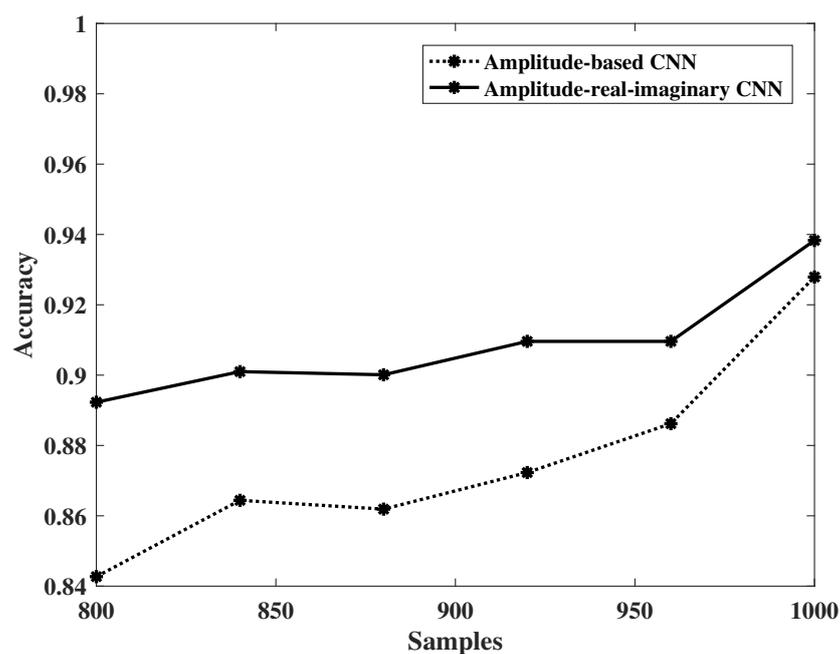


Figure 12. Experimental results based on different networks under EOC.

2. Comparison based on input datasets

Experimental results based on different input datasets are shown in Table 10 and Figure 13. We can conclude from the experimental results based on input datasets that the proposed target classification framework always has better performance no matter the number of training samples. It should be noted that although targets under EOC show great challenges in target classification, when the training samples reduce to 800, that is, 200 of each category, the accuracy of the proposed method can still reach 89%, showing a great potential in practical application when the number of samples is limited.

Table 10. Classification accuracy based on different input datasets under EOC.

Samples	800	840	880	920	960	1000
MF	85.66%	85.49%	87.14%	88.10%	89.49%	92.35%
\tilde{X}_B	89.23%	90.10%	90.01%	90.96%	90.96%	93.83%

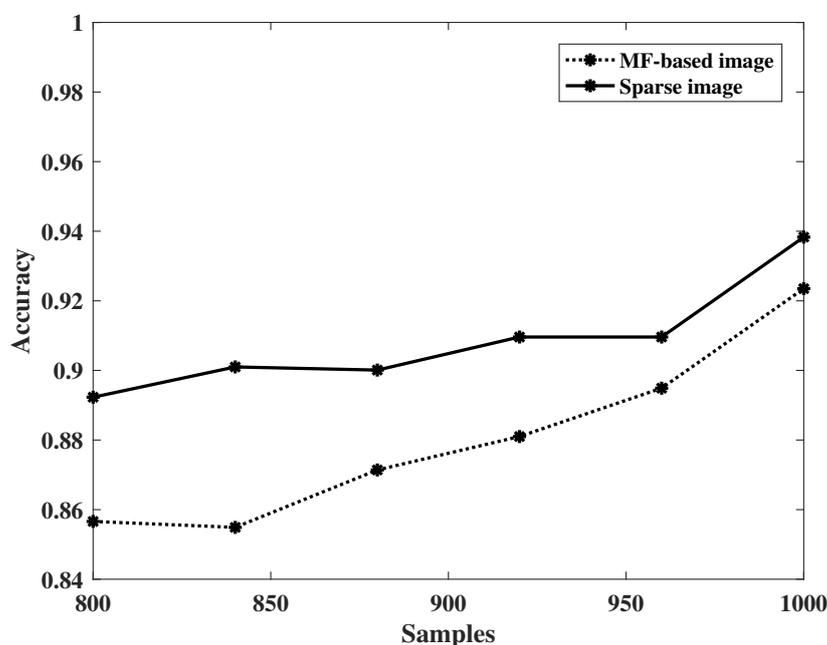


Figure 13. Experimental results based on different input datasets under EOC.

3. Analysis

The same conclusion can be drawn under EOC with the accuracy being 93.83% when the number of training samples is only 250 per class and the confusion matrix is listed in Table 11. From the table, it can be found that the accuracy of T72 is only 81.25%, which is much lower than the other classes. Because the types of images in the training set and testing set are different (see in Table 7), it increases the difficulty of target classification.

Table 11. Confusion matrix of the amplitude–real–imaginary CNN based on sparse images under EOC.

Class	2S1	BRDM2	T72	ZSU234	
2S1	263	1	1	12	
BRDM2	21	284	2	1	
T72	4	0	260	2	
ZSU234	0	2	25	273	
Accuracy (%)	91.32	98.95	90.28	94.79	93.83

5. Experimental Analysis

Experimental results on SAR target detection can be divided into two scenarios. In SOC, the \tilde{X}_C dataset shows similar performance to the MF-based dataset in the aspect of mAP, but it outperforms MF by 0.91% in the aspect of IOU. However, the \hat{X}_C dataset underperforms compared to the MF-based dataset by 2.56% mAP and 11.35% IOU. In EOC, both the \tilde{X}_C and \hat{X}_C datasets present excellent performance in target detection. \tilde{X}_C can obtain a 92.00% mAP and a 65.21% IOU, which is higher than the 90.80% mAP and 55.36% IOU of the MF-recovered images. The \hat{X}_C dataset achieves the best performance with a 96.32% mAP and a 70.85% IOU, which outperforms the MF-based dataset by 5.52% mAP and 15.49% IOU. The different results under SOC and EOC can be explained by the sparse estimations recovered by the CAMP algorithm highlighting the main features with lots of detailed features being ruined, and the non-sparse estimations highlight the main features while retaining detailed features simultaneously. In SOC, the difference between the testing set and training set is small, so that almost all the features learned by the network can match the features of the actual testing set. Therefore, the non-sparse estimations retain more detailed characteristics, showing better performance in target detection. However, in EOC, the difference between the training set and the testing set is huge. Some features learned from the training set may be changed or these features might not even be in testing set, so the detection performance of non-sparse estimations will be worse and the sparse estimations retaining main features will be better at target detection in this case. In summary, the non-sparse estimations of CAMP can be used for routine target detection tasks, and when it comes to more complex target detection tasks, the sparse estimations will be more suitable.

In addition, results of SAR target classification show that the amplitude–real–imaginary CNN performs better than typical CNNs based on sparse SAR datasets. Compared with the MF-based dataset, a higher classification accuracy can be achieved with sparse SAR images as the input data. Moreover, it should be noted that the accuracy of the proposed classification framework is 95.47% under SOC when the training samples decrease to 100 per class, outperforming the amplitude-based CNN with a sparse dataset and the amplitude–real–imaginary CNN with an MF-based dataset by 1.53% and 1.45%, respectively. Similar conclusions can be obtained when the number of training samples come to 800 under EOC, which shows great application potential of the proposed classification network in the situation of limited samples.

6. Conclusions

In this paper, sparse SAR images are reconstructed by CAMP and BiIST-based recovery algorithms for target classification and detection. Different from traditional MF-based images, the CAMP-based recovery algorithm can obtain an image-quality-improved sparse solution and a background-distribution-preserved non-sparse solution. Similarly, the non-sparse solutions of the BiIST-based algorithm can retain the same phase information as MF-based images with improved image quality. Experiments based on the reconstructed sparse SAR image datasets and MF-recovered dataset demonstrate that in the classification task, compared with MF-based images, sparse SAR images show better performance no matter the number of training samples. When it comes to the detection task, the combination of protruding target features and retaining SAR image background distributions can obtain better results.

Author Contributions: Conceptualization, Y.W.; methodology, J.D. and Z.L.; validation, C.S. and H.B.; writing—original draft preparation, J.D. and Z.L.; writing—review and editing, C.S., B.W. and H.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61901213, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020B1515120060, 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,

in part by the Fundamental Research Funds for the Central Universities under Grant NE2020004, University Joint Innovation Fund Project of CALT under Grant CALT2021-11, and in part by the Science and Technology Innovation Project for Overseas Researchers in Nanjing.

Data Availability Statement: The complex SAR image data used in this paper are from the MSTAR dataset, which can be found at <https://www.sdms.af.mil/index.php?collection=mstar>.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this paper:

SAR	Synthetic Aperture Radar
MF	Matched Filtering
TBR	Target-to-background Ration
SNR	Signal-to-noise Ratio
IST	Iterative Soft Thresholding
OMP	Orthogonal Matching Pursuit
CNN	Convolutional Neural Network
RCNN	Regional Convolutional Neural Network
SSD	Single-Shot Multibox Detector
CV-CNN	Complex-valued Convolutional Neural Network
CVFCNN	Complex-valued Fully Convolutional Neural Network
SAT	Self-adversarial Training
FPN	Feature Pyramid Network
PAN	Path Aggregation Network
MSTAR	Moving and Stationary Target Acquisition and Recognition
CAMP	Complex Approximate Message Passing
InSAR	SAR Interferometry
CFAR	Constant False Alarm Rate
SOC	Standard Operating Conditions
EOC	Extended Operating Conditions
IOU	Intersection Over Union
mAP	Mean Average Precision
AP	Average Precision
FPS	Frames Per Second

References

1. Curlander, J.C.; McDonough, R.N. *Synthetic Aperture Radar: Systems and Signal Processing*; Wiley: New York, NY, USA, 1991.
2. Henderson, F.M.; Lewis, A.J. *Principle and Application of Imaging Radar*; John Wiley and Sons: New York, NY, USA, 1998.
3. Eryildirim, A.; Cetin, A.E. Man-made object classification in SAR images using 2-D cepstrum. In Proceedings of the 2009 IEEE Radar Conference, Pasadena, CA, USA, 4–8 May 2009.
4. Zhu, J.; Qiu, X.; Pan, Z.; Zhang, Y.; Lei, B. Projection shape template-based ship target recognition in TerraSAR-X images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *14*, 222–226. [[CrossRef](#)]
5. Magna, G.; Jayaraman, S.V.; Casti, P.; Mencattini, A.; Natale, C.D.; Martinelli, E. Adaptive classification model based on artificial immune system for breast cancer detection. In Proceedings of the 2015 XVIII AISEM Annual Conference, Trento, Italy, 3–5 February 2015.
6. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
7. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2012**, *60*, 84–90. [[CrossRef](#)]
8. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
9. Chen, S.; Wang, H. SAR Target recognition based on deep learning. In Proceedings of the 2014 International Conference on Data Science and Advanced Analytics (DSAA), Shanghai, China, 30 October–1 November 2014.
10. Yue, Z.; Gao, F.; Xiong, Q.; Wang, J.; Huang, T.; Yang, E.; Zhou, H. A novel semi-supervised convolutional neural network method for synthetic aperture radar image recognition. *Cogn. Comput.* **2021**, *13*, 795–806. [[CrossRef](#)]
11. Ma, F.; Zhang, F.; Xiang, D.; Yin, Q.; Zhou, Y. Fast task-specific region merging for SAR image segmentation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [[CrossRef](#)]

12. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
13. Bochkovskiy, A.; Wang, C.Y.; Liao, H. YOLOv4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
14. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016.
15. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
16. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.
17. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Machine Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
18. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Machine Intell.* **2020**, *42*, 86–397. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, Z.; Wang, H.; Xu, F.; Jin, Y.-Q. Complex-valued convolutional neural network and its application in polarimetric SAR image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 7177–7188. [[CrossRef](#)]
20. Coman, C.; Thaens, R. A deep learning SAR target classification experiment on MSTAR dataset. In Proceedings of the 19th International Radar Symposium (IRS), Bonn, Germany, 20–22 June 2018.
21. Yu, L.; Hu, Y.; Xie, X.; Lin, Y.; Hong, W. Complex-valued full convolutional neural network for SAR target classification. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1752–1756. [[CrossRef](#)]
22. Zhang, B.; Hong, W.; Wu, Y. Sparse microwave imaging: Principles and applications. *Sci. China Inf. Sci.* **2012**, *55*, 1–33. [[CrossRef](#)]
23. Bi, H.; Bi, G.; Zhang, B.; Hong, W.; Wu, Y. From theory to application: Real-time sparse SAR imaging. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2928–2936. [[CrossRef](#)]
24. Donoho, D.-L.; TSAIG, Y.; Drori, I.; Starck, J.-L. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2012**, *58*, 1094–1121. [[CrossRef](#)]
25. Pati, Y.C.; Rezaifar, R.; Krishnaprasad, P.S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 1–3 November 1993.
26. Bi, H.; Bi, G. Performance analysis of iterative soft thresholding algorithm for L_1 regularization based sparse SAR imaging. In Proceedings of the 2019 IEEE Radar Conference (RadarConf), Boston, MA, USA, 22–26 April 2019.
27. Daubechies, I.; Defriese, M.; De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **2004**, *57*, 1413–1457. [[CrossRef](#)]
28. Bi, H.; Zhang, B.; Zhu, X.; Hong, W.; Sun, J.; Wu, Y. L_1 -regularization-based SAR imaging and CFAR detection via complex approximated message passing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3426–3440. [[CrossRef](#)]
29. Bi, H.; Deng, J.; Yang, T.; Wang, J.; Wang, L. CNN-based target detection and classification when sparse SAR image dataset is available. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6815–6826. [[CrossRef](#)]
30. Bi, H.; Bi, G.; Zhang, B.; Hong, W. A novel iterative thresholding algorithm for complex image based sparse SAR imaging. In Proceedings of the 12th European Conference on Synthetic Aperture Radar, Aachen, Germany, 4–7 June 2018.
31. Bi, H.; Bi, G. A novel iterative soft thresholding algorithm for L_1 regularization based SAR image enhancement. *Sci. China Inf. Sci.* **2019**, *62*, 049303. [[CrossRef](#)]