

Recent Advances in Deep Learning based SAR Image Targets Detection and Recognition

Ping Lang, *Member, IEEE*, Xiongjun Fu, *Senior Member, IEEE*, Jian Dong, Huizhang Yang, *Member, IEEE*, Junjun Yin, *Senior Member, IEEE*, Jian Yang, *Senior Member, IEEE*, and Marco Martorella, *Fellow, IEEE*

Abstract—SAR image targets detection and recognition (SAR-TDR) have become research hotspots in the remote sensing application. These targets include ships, vehicles, aircrafts, oil tanks, bridges, and so on. However, with the rapid development of SAR technology and increasingly complex electromagnetic environment, complex characteristics of SAR images bring severe challenges to the accurate SAR-TDR via traditional physical models or manually features extraction based machine learning methods. In recent years, deep learning (DL), as a powerful automatic features extraction algorithm, has been widely used in the computer vision domain. More specifically, DL has also been introduced into the SAR-TDR tasks and effectively achieved good performances in terms of accuracy, real-time, etc. With the rapid development of DL, SAR images processing, and practical requirements of SAR-TDR in civilian and military domains, it is very necessary to make a systematic survey on the SAR-TDR in past few years. In this survey paper, we mainly conduct a systematic overview of DL-based SAR-TDR literatures on two tasks, i.e., targets recognition (e.g., ground vehicles, ships, and aircrafts) and targets detection (e.g., ships, aircrafts, , change detection, searsurface oil spill, and oil tanks), respectively. More specifically, our related works about these topics are also presented to verify the effectiveness of DL-based methods. Firstly, several DL methods (e.g., CNN, RNN, AEs, and GNN), commonly used to SAR-TDR, are briefly introduced. The systematic review of DL-based SAR-TDR (including our related works) is then presented. Finally, the current challenges and future possible research directions are deeply analyzed and discussed.

Index Terms—SAR images interpretation, deep learning, automatic targets recognition, targets detection, change detection

I. INTRODUCTION

OPERATING conditions of all weather, day-and-night and high-resolution imaging, synthetic aperture radar (SAR) is widely applied in remote sensing field in military and civil applications [1], [2], such as automatic targets recognition, land cover classification, and ships detection. As an active

This work was supported in part by the Natural Science Foundation of China (NSFC) under Grant 62401320, 62222102 and Grant 62171023. (*Corresponding author: Ping Lang*)

Ping Lang and Jian Yang are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: plang1220@sina.com; yangjian_ee@tsinghua.edu.cn).

Xiongjun Fu and Jian Dong are with the School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: fuxiongjun@bit.edu.cn; radarvincent@sina.com).

Huizhang Yang is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: hzyang@njust.edu.cn).

Junjun Yin is with School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: yinjj07@gmail.com)

Marco Martorella is with the Department of Information Engineering, University of Pisa, Pisa 56127, Italy (e-mail: marco.martorella@iet.unipi.it).

remote sensor, the SAR carries its own microwave illumination and does not depend on external sunlight like optical imaging. With the rapid development of military and electronic information technology, many types of SAR equipments have been practically applied, such as satellite-borne SAR (e.g., Sentinel-1/2 [3], Radarsat-1 [4], SIR-C/X [5], TerraSAR-X [6], Gaofen-3 [7]), airborne SAR [8] (e.g., AN/APD-10, ERIMX/SIR, E-SAR), and missile-based [9], [10]. Different SAR systems have its own specific configuration characteristics. The related characters include working frequency/wavelength, imaging mode (e.g., stripmap SAR, spotlight SAR, scan SAR), polarization, resolution in range and azimuth direction, synthetic apertures, etc. [1].

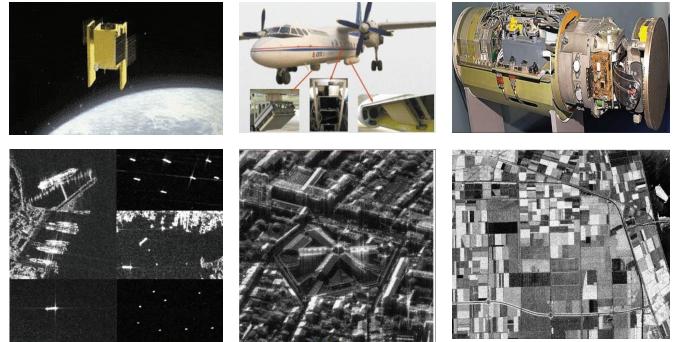


Fig. 1. SAR platform and SAR image samples, top: SAR platforms; down: SAR images.

The SAR image is a 2D (i.e., range and azimuth) high resolution image. The high resolution in range direction is obtained by the pulse compression technology of compressed transmitted signal, such as linear frequency modulation (LFM), phased coded signal. As for the high resolution in azimuth direction, a long synthetic aperture, formed along the trajectory of relative motion between target and radar platform, is used to store the magnitude and phase of successive radar echoes to guarantee the high resolution. Therefore, the relative motion between the target and radar platform is basically necessary for the formation of the SAR image. The downstream tasks of SAR image based include SAR targets recognition, detection, segmentation [11], [12], and so on. In this survey paper, we only focus on the SAR targets recognition and detection, since the article length and our research interests.

Here, we briefly give the goals of SAR targets recognition and detection. The SAR targets recognition methods are studied to accurately recognize the actual class of target on each SAR images with small size. It usually has only one target in each SAR image with high resolution, such as vehicles, ships, aircrafts, as shown in Fig.7. Some literatures regard the “classification” as the “recognition”. Therefore, we can not

specifically distinguish the “classification” and “recognition”, and unify the presentation as “recognition” in this paper. As for the SAR targets detection methods, which are used to usually locate and recognize the multiple targets on each SAR image with large size. There are more than one targets, such as ships, bridges, oil tanks, cars, as shown in Fig.8(c).

The effect and efficient SAR-TDR tasks are of great significance for the accurate remote information awareness. However, it poses severe challenges due to the complex SAR imaging mechanism and diverse characteristics of SAR images. As for the SAR imaging mechanism, there are some specific phenomena on SAR images, such as shadows, dead zones, top and bottom inversion. Compared to optical images, the SAR image has its own specific characteristics including i) noisy. SAR images are easily distributed by the multiplicative noise (known as speckle) due to the coherent and destructive interference scattered by many small reflectors within each resolution cell [13]; ii) sparsity. A SAR image usually consists of only a few discrete heavy scattering points of the target; iii) sensitivity. It is very sensitive to the changes of target’s postures and configurations, such as the shadows effect, the interaction of the target’s backscatter with the external environment (e.g., clutter, adjacent targets); and iv) time-variable. The SAR imaging information is determined by the amplitude and phase of the radar echoes, which are usually time-varying. Therefore, unlike optical images, SAR images are significantly different from the human visual cognitive system, which directly leads to the heavy difficulty on the effective SAR images interpretation for the human individuals.

The complex characteristics of SAR images have brought severe challenges to the accurate SAR-TDR via traditional physical models based or machine learning methods. In recent years, with the rapid development of deep learning (DL) [14], [15], DL has also been widely introduced into the SAR-TDR and effectively achieved good performances in accuracy, real-time, such as [16], [17], [69]. Compared to traditional methods, the advantages of DL can be roughly summarized as i) automatic features extraction; ii) end-to-end architecture for features extraction and substream tasks (e.g., classification, detection); iii) data-driven, instead of well-designed parametric physical model; and iv) good performance improvement. However, DL-based SAR-TDR methods still face some serious problems as following:

(1) The DL model is sensitive to the speckle noise. For example, DL models usually have many layers and parameters, which results in serious robustness with noisy inputs.

(2) The complex detection scenarios usually mislead the DL model to extract effective targets’ features, which may lead to performance degradation. For example, complex backgrounds may produce false alarms or false dismissal in ships detection scenarios, such as sea clutter, inshore zones, intensive docking of many ships.

(3) Small SAR data samples may lead to the overfitting of the DL model due to sensitivity to the SAR imaging parameters or detection background. This is mainly caused by three aspects: i) the number of data is inherently small, since the SAR imaging costs and conditions are seriously limited; ii) a large number of clear data is difficult to be quickly

obtained. SAR images are usually polluted by the unintentional narrowband / wideband interference or intentional hostile jammer signals, since SAR is a wideband imaging modal; and iii) public SAR image target datasets are severely absent, since the military or commercial secrets.

(4) The complexity of the DL model is usually large, and requires large computational costs, which is mainly represented from two aspects: i) large models, such as ResNet has more than 150 layers; and ii) requirements of large data training. Complex models are usually difficult to meet real-time requirements, making them difficult to deploy to portable devices, such as laptops.

Fortunately, there are many related works have been carried out in the SAR-TDR field to address these problems in the past few years. With the rapid development of DL and SAR-based remote sensing techniques in the past few years, it is necessary to make a comprehensive survey on the latest developments in SAR-TDR. More specifically, with the rapid development of multi-modal large language model and embodied artificial intelligence in 2022 and 2024, respectively, what are the directions of future intelligent SAR-TDR? How to take full use of these latest AI achievements into the SAR-TDR domain to generate new breakthroughs in the challenges of SAR-TDR? Therefore, this paper systematically reviews these works from targets recognition, such as [18]–[21], and targets detection, such as ships detection [22]–[24], aircrafts detection [227], [229], [231], change detection [248]–[250], and so on. Other applications of SAR images can be referred to our previous overview work in [25]. Deep learning algorithms include convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs), and graph neural network (GNN). Take the searching results of “web of science” database as an example, the published papers of “SAR target recognition” and “SAR target detection” in recent five years have been drawn in Fig.2, which has predicted that the achievements of this domain will be quickly increased along the tendency line.

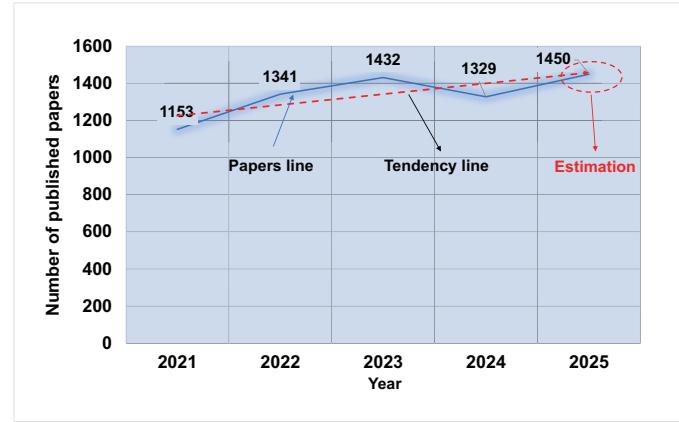


Fig. 2. The number of published papers related this topic.

Although there are some survey papers on the applications of DL in SAR-TDR [26], [27], [83]. We believe that the practical application scenarios of DL to SAR-TDR needs more attention. In other words, SAR-TDR are very different from the optical images, such as available labeled dataset, real-

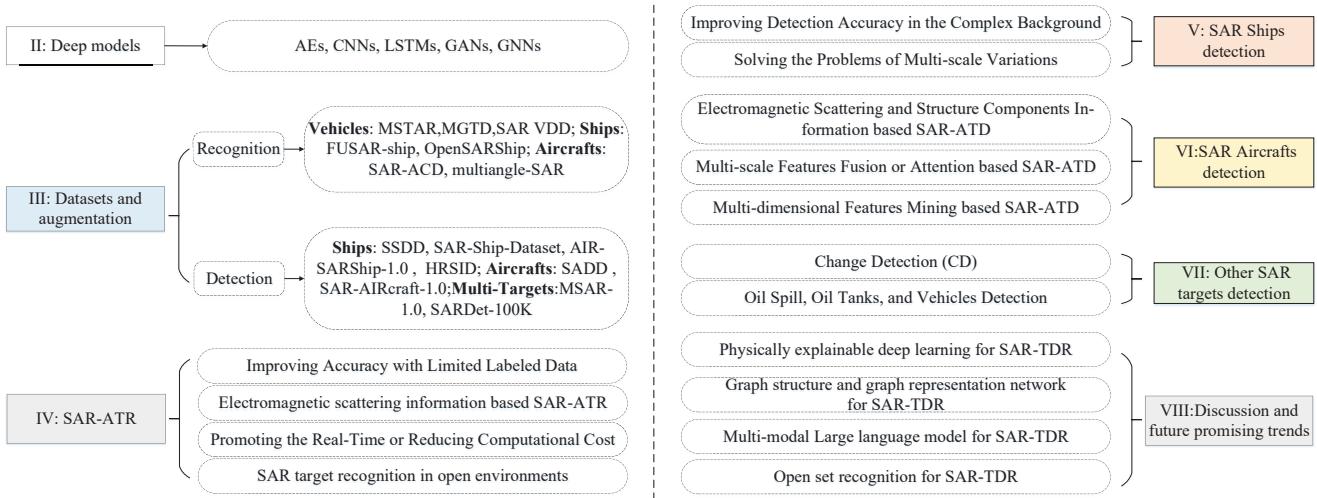


Fig. 3. The main contents of this paper.

time requirement. Therefore, we focus on the recent advances (such as graph CNN, complex value CNN) on the applications of small data sample and real-time, including SAR-ATR and targets detection. Specifically, we make a comprehensive survey from dataset augmentation, fine-grained DL model design, model train techniques, and evaluation metrics. Very different from existing review works aforementioned, we have proposed related novel methods to address these problems in SAR-CRD. Our latest works have also been introduced in details. The main contents of this paper is shown in Fig.3, except for the Introduction and Conclusion in section I and IX.

The motivation of this paper is mainly providing: i) a solid basis for new researchers and practitioners who are approaching this field for the first time; ii) an important reference for more experienced researchers who are working in this field; iii) existing terms for comparison for newly developed DL-based algorithms; iv) means to identify gaps; and v) a full understanding of strengths and limitations of DL-based approaches.

Hopefully, this paper can help relative researchers and practitioners to quickly and effectively determine potential facts of this topic by clearly knowing about key aspects and related body of research from the practical applications prospect. For this consideration, compared to existing related survey works, we make mainly three contributions:

(i) Aiming to the practical applications, we only make a systematical overview on three aspects, i.e., limited data, real-time, and complex interference scenarios;

(ii) The main concepts, motivations, and implications of enabling DL algorithms in SAR-TDR, are comprehensively proposed;

(iii) A profound discussion of future promising research opportunities and potential trends is proposed in this field.

Accordingly, the reminder of this review article is organized as follows. The basic principles of typical DL algorithms are briefly presented in Section II. Section III introduces some typical open SAR datasets and data augmentation methods, popularly used in SAR-TDR. State-of-the-art achievements about the SAR targets (including vehicles, ships and aircrafts)

recognition are reviewed in section IV. Section V presents a survey on the SAR ships detection. SAR aircrafts detection is reviewed in Section VI. In Section VII, the other targets, such as seafloor oil spill, oil tanks, change detection, are also surveyed. Open problems, challenges and possible promising research directions are profoundly discussed in section VIII. Finally, the conclusions of this article is presented in section IX.

II. DEEP MODELS

(1) Autoencoders (AEs)

As one of important unsupervised learning algorithms, AEs mainly perform the data compression or dimensionality reduction tasks. An AE model usually has three modules: encoder, activation function, and decoder [28].

Encoder f is used to extract the latent features of input x , weighted through weight matrix \mathbf{W} and bias vector \mathbf{b} , i.e., $f = \mathbf{W}x + \mathbf{b}$. Activation function σ can be regarded as a non-linear mapping, which transforms the f into the latent feature map h of input x , i.e., $h = \sigma(\mathbf{W}x + \mathbf{b})$. Contrary to encoder f , the decoder g is used to generate the reconstruction \tilde{x} of the input x , i.e., $\tilde{x} = g(\mathbf{W}^T h + \mathbf{b}')$.

The gap between the \tilde{x} and input x can be determined by the loss function $L(\tilde{x}, x)$. The training processing of AE is minimizing the $L(\tilde{x}, x)$.

(2) Convolutional Neural Networks (CNNs)

Inspired by human's brain visual information processing mechanism, CNNs are very popular in CV domain [29], such as target recognition/classification, detection. Many famous CNNs architectures have been proposed, which can be briefly shown in Fig.4. CNNs can effectively extract latent features of data through convolution layers (i.e., shared convolutional kernels) and pooling layers (i.e., downsampling). At the beginning, the increasing of model depth is a main fashion, e.g., from a 5-layer of LeNet in 1998 to hundreds of layers of ResNet in 2015 [30]. In order to improve the training efficiency, the lightweight DNN models, i.e., small volume of parameters, is increasingly emerging, e.g., ShuffleNet [31], MobileNet [32], EfficientNet [33].

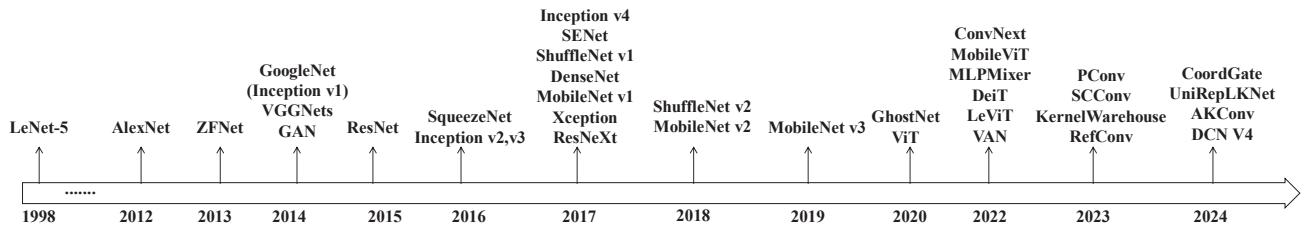


Fig. 4. Development pipeline of CNN models [29]

(3) Recurrent Neural Networks (RNNs)

Differently, inspired by the memory function of human brain information processing system, RNNs are used to predict the temporal memory series data, such as long short-term memory (LSTM) [34], gated recurrent unit (GRU) [35]. The LSTM is proposed to address the gradient vanishing issue of original RNN model during the training process.

Compared to original memory unit of RNN, LSTM module usually has two units: *long-term memory* unit (C) and *short-term memory* unit (h), both are m dimensions. C can selectively memorize valuable information previous data series, which can efficiently transmit the early information to current data series. LSTM has four gate control units: *forget*, *memory*, *information* and *output*, respectively. Each gate control unit contains a FCNN layer with m neurons, and the outputs of each gate are short time memory unit h and data features. Except for *information* gate is *tanh* activation function, the others are activation functions are *sigmoid*. The working procedure of LSTM model can be shown as follows:

Firstly, the *forget* gate unit discard some non-valuable information of inputs

$$f_t = \text{sigmoid}(\mathbf{W}_f[h_{t-1} \ x_t]) + b_f. \quad (1)$$

Then, *memory* and *information* gate units transmit valuable information to the *input* gate unit

$$i_t = \text{sigmoid}(\mathbf{W}_i[h_{t-1} \ x_t]) + b_i. \quad (2)$$

$$\tilde{C}_t = \text{Tanh}(\mathbf{W}_c[h_{t-1} \ x_t]) + b_c. \quad (3)$$

The long time memory (C) is then updated through

$$C_t = C_{t-1} * f_t + i_t * \tilde{C}_t. \quad (4)$$

The function of *output* gate unit can be expressed by

$$\tilde{o}_t = \text{sigmoid}(\mathbf{W}_o[h_{t-1} \ x_t]) + b_o. \quad (5)$$

Lastly, the short time memory (h) can be given by

$$h_t = o_t * \text{Tanh}(\tilde{C}_t). \quad (6)$$

Accordingly, the gradient vanishing issue can be alleviated via LSTM, since the *long time memory* unit (c) and *forget* gate unit. c can effectively preserve valuable information with large weights, while *forget* gate discards non-valuable information. In this way, the gradient can not become more

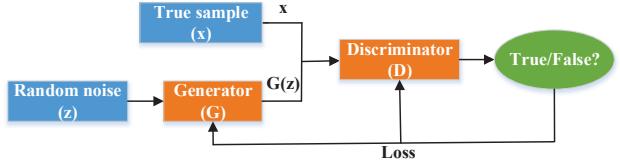


Fig. 5. The diagram of GAN.

smaller after layer-by-layer training, so as to avoid gradient vanishing. As an improved LSTM, GRU [35] is more simple, which combines *forget* gate with *input* gate (i.e., *memory* and *information* gates aforementioned) as a single *update* gate. GRU converges more faster than LSTM and does not cause overfitting.

(4) Generative Adversarial Networks (GANs)

GAN, composed of *Generator* G and *Discriminator* D , is also an unsupervised generative model [36], [37], as shown in Fig.5. The noisy sample with standard normal distribution is usually used as the input of G , to produce a new sample (i.e., $G(z)$) to deceive D via adversarial training. A two-class classifier D is used to discriminate whether the new sample is true or not, compared to original true sample. The G has only one loss function to minimize to gap between the generated new sample and true label. As for D , there are two loss functions, one is determined by true sample and true labels, denoted by $loss_1$, the other is determined by generated new sample and false label, denoted by $loss_2$. The total loss function of GAN can be shown as follows

$$\min_G \max_D V(G, D) = E_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x})] + E_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))], \quad (7)$$

where $p_{data}(\mathbf{x})$, \mathbf{z} , $p_z(\mathbf{z})$, and $G(\mathbf{z})$ denote true distribution of sample, noisy sample, distribution of noise sample, and generated new sample, respectively. The distribution of $G(\mathbf{z})$ can be represented by $p_G(\mathbf{x})$. $D(\mathbf{x})$ and $1 - D(G(\mathbf{z}))$ denote the loss of discriminator D and generator G , respectively.

The purpose of D is trying to correctly discriminate the original true sample and generated false sample, while generator G is trying to generate the true sample. Therefore, the training process of GAN is adversarial training style between discriminator D and generator G . GAN is firstly trained by discriminator D to maximize the expectation of discriminative results. The parameters of discriminator D are then fixed, the generator G is trained to minimize the Jensen-Shannon divergence [36] between the true samples and generated false sample. That is to say, generator G tries to make the distribu-

tion of the generated false sample close to the distribution of true sample as close as possible. The training process can not stop until the discriminative probability of discriminator D for the true sample and generated false sample is equal, i.e., 0.5. There are many improved GANs model, such as Wasserstein GAN (WGAN) [38], to improve the training effectiveness instead of suffering from the gradient vanishing problem. The objective function of WGAN can be expressed by

$$\min_G \max_{D \in \Omega} W(p_r, p_g) = E_{\mathbf{x} \sim p_r(\mathbf{x})}[D(\mathbf{x})] - E_{z \sim p_g(\mathbf{z})}[D(G(\mathbf{z}))], \quad (8)$$

where Ω is the 1-Lipschitz functions set, $p_g(\mathbf{z})$ is the generator G distribution, and $p_r(\mathbf{x})$ is the true sample distribution.

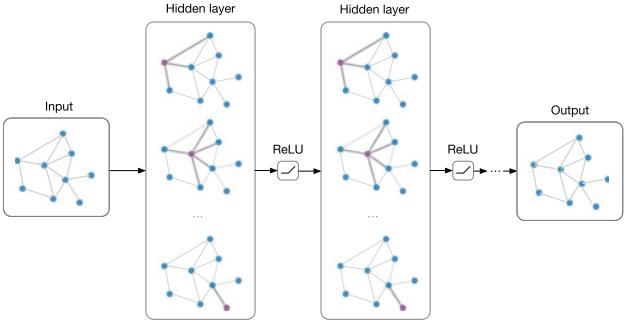


Fig. 6. An illustrative example of GCN [40].

(5) Graph convolutional networks (GCNs)

Graph structure is also used to improve the performance in SAR-TDR, such as [105], [106]. $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{A})$ can represent a graph network, where $\mathbf{V} = \{v_1, v_2, v_3, \dots, v_N\} \in \mathbb{R}^{N \times d}$ denotes the vertex set, N represents the vertex number, $\mathbf{E} = \{e_1, e_2, e_3, \dots, e_M\} \in \mathbb{R}^{M \times 1}$ is the edges set. $\mathbf{A} \in \mathbb{R}^{N \times N}$ represents the adjacency matrix, used to denote the relationships between vertices, which can be defined by [39], [40]

$$\mathbf{A}_{ij} = \begin{cases} 1, & \text{if } (v_i, v_j) \subseteq \mathbf{E} \\ 0, & \text{else} \end{cases}, \quad (9)$$

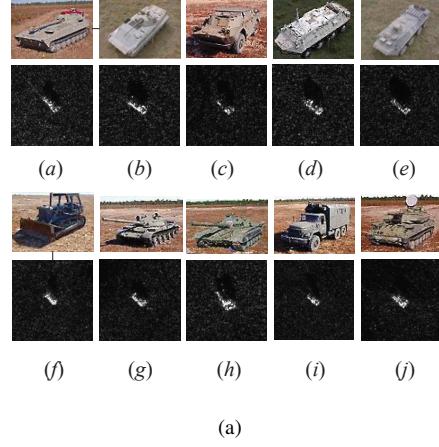
where $v_i, v_j \in \mathbf{V}$, $i \neq j, i, j = 1, 2, 3, \dots, N$ represents the index of different vertex.

There are three steps for the GCN learning process: aggregation, update, and prediction, as shown in Fig.6. The aggregation step is used to update information representation of each vertex via aggregating features of its all neighbor vertices in the latent years. The prediction results of each vertex in the output layer, can be performed by the crossentropy loss function. The GCN learning process can be mathematically expressed by [40]

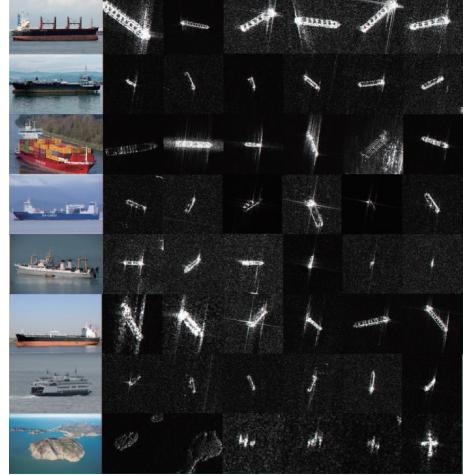
$$\mathbf{H}^{(l+1)} = f_\theta \left(\mathbf{H}^{(l)}, \mathbf{A} \right) = \sigma \left(\hat{\mathbf{A}}^{-\frac{1}{2}} \hat{\mathbf{A}} \hat{\mathbf{A}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right), \quad (10)$$

where $\mathbf{H}^{(l+1)} \in \mathbb{R}^{N \times C_{l+1}}$, $\mathbf{H}^{(l)} \in \mathbb{R}^{N \times l}$ is the extracted feature maps via convolutional network from the $(l+1)_{th}$ and l_{th} layer, respectively. $\hat{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, \mathbf{I} is identity matrix. N , C_{l+1} , and C_l are the number of vertices, dimensions of output features of the $(l+1)_{th}$ and l_{th} layer, respectively.

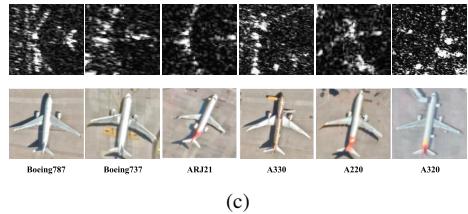
In the first layer, i.e., input layer, the $H^{(1)} = \mathbf{X}$. $\sigma(\bullet)$ is activation function. This transformation is used to change the diagonal elements of A from 0 to 1. That is to say, the information update process of each vertex is performed by both information of its neighbors and itself. The undirected graph can be normally represented by $\hat{\mathbf{D}}^{-\frac{1}{2}} \hat{\mathbf{A}} \hat{\mathbf{D}}^{-\frac{1}{2}}$, where $\hat{\mathbf{D}} \in \mathbb{R}^{N \times N}$ denotes the diagonal matrix of vertices' degree, $\hat{\mathbf{D}}_{ii} = \sum_{j=1}^N \hat{\mathbf{A}}_{ij}$. $\mathbf{W}^{(l)}$ represents the weight matrix of l_{th} layer.



(a)



(b)



(c)

Fig. 7. The samples of SAR target recognition datasets (SAR images and its optical images). (a) ground vehicles: MSTAR [42], (b) ships: FUSAR-ship [46], and (c) aircrafts: SAR-ACD [48]

III. DATASETS AND AUGMENTATION

A. Datasets

Dataset usually contains training datasets, validation datasets, and testing datasets. In order to better train DL model, a great number of available dataset is necessary.

However, the data collection and datasets establishment are challengeable, which usually spend huge human and financial resources. More specifically, the big SAR datasets are not easily collected, since military or commercial secretes, compared with general CV datasets. Luckily, there are still some public SAR datasets in target recognition and detection tasks [41], including SAR target recognition datasets, such as vehicles recognition (MSTAR [13], [42], [43], MGTD [44], SAR VDD [45]), ships recognition (FUSAR-ship [46], OpenSARShip [47]), aircrafts recognition (SAR-ACD [48], multiangle-SAR [49]); target detection datasets, such as ships detection datasets (SSDD [50], SAR-Ship-Dataset [51], AIR-SARShip-1.0 [52], HRSID [53]), aircrafts detection datasets (SADD [54], SAR-AIRcraft-1.0 [55]), multi-target detection datasets, (MSAR-1.0 [56], SARDet-100K [57], e.g., aircrafts, oil tanks, bridges, harbors, cars, and ships). Others multi-sources datasets, such as 3MOS [58], SOPatch [59], are also used to perform SAR-TDR tasks. The detail information of main public SAR datasets are shown as in Tab.I for the targets detection and recognition, and Tab.II for the change detection (CD) tasks. The dataset samples of SAR targets recognition, detection, and CD are shown as in Fig.7, Fig.8, and Fig.9, respectively. There, we take some datasets as the examples to briefly introduce.

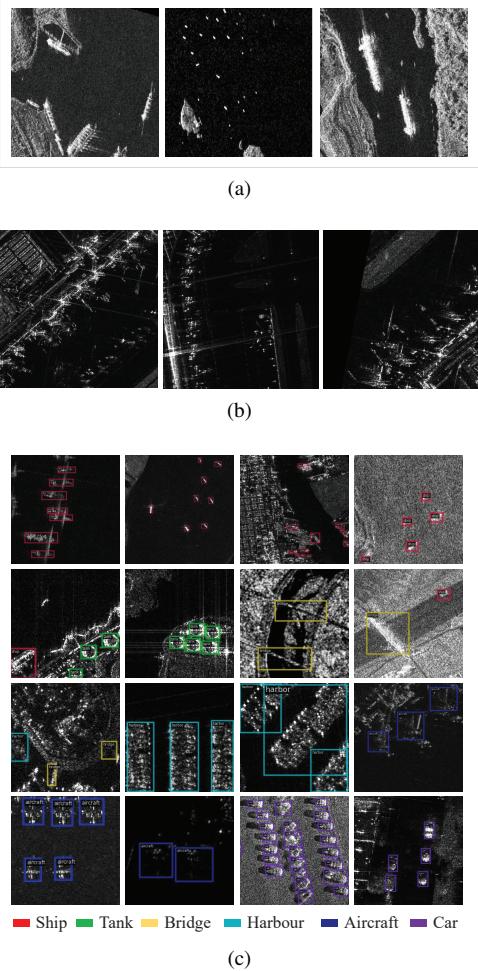


Fig. 8. The samples of SAR target detection datasets. (a) ships: SSDD [50], (b) aircrafts: SAR-AIRcraft-1.0 [55], and (c) multi-target: SARDet-100K [57]

The X-band SAR based MSTAR dataset contains 10 classes

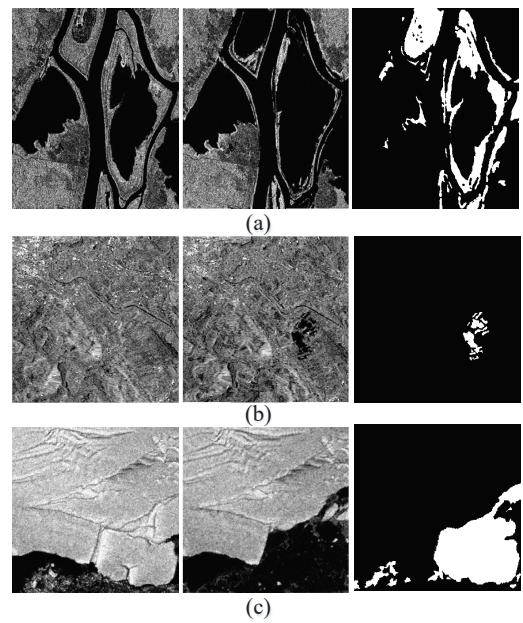


Fig. 9. Some dataset samples of CD tasks. (a) Ottawa [60], (b) Bern [62], and (c) Sulzberger [60]

of ground targets, e.g., T72 (main tank), BTR70 (armored personnel carrier), and BMP2 (infantry combat vehicle). The size of all images is $128 * 128$ with $0.3 \text{ m} * 0.3 \text{ m}$ resolution. Some example samples are as shown in Fig.7(a). MSTAR dataset is very popularly used to SAR-ATR tasks.

SSDD dataset [50] is used to ship detection, including 1,160 images and 2,456 ships in total. The construction process of SSDD is similar as PASCAL VOC [68]. In addition, SAR-Ship-Dataset [51] has 102 Chinese Gaofen-3 images and 108 Sentinel-1 images. There are 43,819 ship chips with the size of $256 * 256$. Some samples of SSDD can be shown in Fig.8(a).

AIR-SARShip-1.0 [52] has 31 images with the size of $3,000 * 3,000$, including spotlight mode and stripemap imaging mode. The backgrounds have port, island and different sea conditions. The ship targets have 10 classes, such as transport ship, oil ship, and fisher boat.

HRSID [53] is high-resolution SAR dataset with a number of 5,604 images and 16,951 ship instances, used for ship detection, instance segmentation tasks, and semantic segmentation. The construction process of HRSID is similar as COCO dataset. The scenarios of HRSID contain different sea conditions, sea areas, and coastal ports. The resolution is 0.5 m, 1 m, and 3 m.

B. Data Augmentation

Although some public available SAR datasets aforementioned can be used for training DL model, the number of labeled samples is still limited and maybe do not meet the requirements of complex DL model training. The data augmentation methods have been proposed, such as GANs [69], [70], or design novel small data learning model to learn with limited labeled data, such as transfer learning (TL) based methods [71], [72].

Wasserstein GAN (WGAN) was developed in [69] to produce new data samples based on MSTAR dataset. After data

TABLE I: The main public SAR image datasets for the targets detection and recognition, including vehicles, ships, and aircrafts, oil tanks, bridges, and so on.

Datasets	Classes	Number of instances	Number of images	Size	Tasks	Year
MSTAR [42]	10	5,950	5,950	128 × 128	Vehicles Recognition	1998
MGTD [44]	4	1,728	1,728	-	Vehicles Recognition	2017
SAR VDD [45]	-	2,958	104	-	Vehicles Recognition	2017
SSDD [50]	1	2456	1160	190-668	Ships detection	2017
SAR-Ship-Dataset [51]	1	59535	43819	256 × 256	Ships detection	2019
AIR-SARShip-1.0 [52]	1	461	31	3000 × 3000	Ships detection	2019
HRSID [53]	1	16,951	5604	800 × 800	Ships detection and segmentation	2020
FUSAR-Ship [46]	15	16,144	16,144	512 × 512	Ships detection and recognition	2020
OpenSARShip [47]	17	11,346	11,346	256 × 256	Ships detection and recognition	2017
SynthWakeSAR [64]	10	-	46,080	800-1500	Ships detection	2022
LS-SSDD-v1.0 [65]	1	-	9,000	800 × 800	Small ships detection	2020
SADD [54]	1	7,835	2,966	224 × 224	Aircrafts detection	2022
MSAR-1.0 [56]	4	60,396	28,449	256-248	Aircrafts, oil tanks, bridges, and ships detection	2022
SAR-ACD [48]	6	3,032	3,032	-	Aircrafts recognition	2022
multiangle-SAR [49]	2	144	144	128 × 128	Aircrafts recognition	2022
SPGAN-SAR [66]	10	5,040	5,040	158 × 158	Aircrafts, ships, and vehicles recognition (simulation)	2018
IRIS-SAR [67]	6	355	63,900	512 × 512	Aircrafts, ships, and vehicles recognition (simulation)	2020
SAR-AIRcraft-1.0 [55]	7	16,463	4,368	800-1500	Aircrafts detection	2023
SARDet-100K [57]	6	245,653	116,598	512 × 512	Aircrafts, Tanks, Bridges, Harbors, Cars, and Ships detection	2024

TABLE II: The main public SAR image datasets for the change detection tasks.

Datasets	Sensor	Size (pixels)	Date	Location	Event
Ottawa [60]	RADARSAT	290 × 350	May - August 1997	Ottawa	Flood
Sulzberger [60]	ESA/ENVISAT	256 × 256	March 2011	Pacific Ocean	Sea ice breakup
Farmland [61]	Radarsat-2	291 × 306	June 2008 - June 2009	Yellow River Estuary in China	Farmland changes
Bern [62]	ERS-2	301 × 301	April - May 1999	Bern	Flood
San Francisco [62]	ERS-2	256 × 256	August 2003 - May 2004	San Francisco	Flood
WenChuan [63]	ESA/ASAR	301 × 442	March 2008 - June 2008	China	Earthquake
Atlantico [63]	ALOS/PALSAR	729 × 1056	April 2010 - March 2011	Colombia	Flood
Katios NationalPark [63]	Sentinel 1A	879 × 1319	March 2019 - April 2019	Colombia	Fire

augmentation, the DL model can improve the recognition accuracy rate from 79% to 91.6% for three-class task and from 57.48% to 79.59% for ten-class task. A least squares GAN with TL technique was proposed for SAR data augmentation in [70]. Different from [69], [70], the authors in [71] augment the dataset via image processing methods, such as manual-extracting sub-images, filtering, adding noise, and flipping. In [73], the authors produced new noisy samples with different SNRs, partially occluded images, and multiresolution representations based on original images. In addition, in [19], [74], the authors proposed three types of data augmentation based on MSTAR dataset, i.e., translation of the targets, adding random speckle noise, and postures synthesis. Sparse representation was used to construct image samples or data augmentation based on attributed scattering centers in [75]–[77].

In [20], the authors proposed a domain-specific data augmentation method based on accuracy-translation map. This method can achieve SOTA classification accuracy rate of 99.6% on MSTAR. The authors in [78] produced adequate multi-view SAR data via a flexible mean based on limited data. As an alternative method, the authors in [79] developed

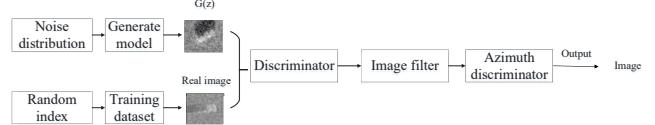


Fig. 10. The diagram of WGAN to generate new SAR images [69].

an electromagnetic simulation method to produce adequate bistatic SAR image samples. In [80], the amplitude and phase information of SAR image were both used as the inputs of CNN model to reduce the over-fitting.

In order to generate precise SAR target images, the authors in [81] developed an angle transformation GAN (ATGAN) based on image-to-image translation, to generate azimuth-controllable SAR target images while preserving the target details for the SAR-ATR task. The authors in [82] developed a SAR-to-optical image translation neural network for the SAR-ATR with generating a virtual optical image based on the SAR image target data using a 3-D model.

TABLE III: The overview of DL-based methods for the SAR-ATR.

Classes	Methods	Models
Improving accuracy with limited labeled data	Data / features augmentation	simulated SAR data [85], [86], code-image-code cyclic network [88], Wasserstein AE [89], cycleGAN [87], generative deconvolutional DNN [90], conv-biLSTM prototypical network [91], multi-view [79], [94], [95], multi-task [96], [97], multi-scale [16], [98], and multi-dimension [99], [100]
	Fine-grained DNN structures design	FCN [108], compressed CNN [109], regularization [110], parallel CNN [93]
	Transfer learning	optical and non-optical domains [117], [119], electro-optical domains [118], meta transfer learning [120], [121]
	Semi-supervised and unsupervised learning	semi-supervised via GANs or DCGANs [123]–[125], unsupervised multi-level adaptation learning [126]
Electromagnetic scattering information based SAR-ATR	-	scattering topology network [130], GCN-based spatial-structural association network [131], strong scattering point awareness network [132]
Promoting the real-time or reducing computational cost	-	ensemble learning [137], lightweight design [117], [140]–[143], graph representation [105], [106], parallel design [144], semi-random DNN [145], micro CNN [146], compression and acceleration [147], deep forest [148]
SAR target recognition in open environments	-	density coverage selection [149], random sampling strategy [150], unsupervised representation [151], GAN-based multitask [152], [153], contrastive learning [154], [155], sub-dictionary learning [156], spherical space adaptation [157], incremental learning [158], [159], [161]

IV. SAR AUTOMATIC TARGET RECOGNITION

In order to improve the resolution of SAR images, spotlight SAR usually observes the interest zone with long time. Therefore, the SAR images, produced by spotlight SAR, usually have small size (such as the size of image in MSTAR dataset is 128×128) with only one target on each SAR image. As for this type of SAR images, it is usually used for the SAR automatic target recognition (SAR-ATR) tasks. In addition, the large SAR images (usually contain multiple targets) can be cut into multiple small slices, each of which has only one target. In this way, these small slices can be constitute a new SAR-ATR dataset, such as FUSAR-Ship [46], OpenSARShip [47].

The goal of SAR-ATR tasks is to predict the classes of targets on the SAR images via mathematical algorithms. SAR-ATR is very popular in military and civilian domains, thus, it has become a research hotspot in SAR image processing community. However, seeing from the published literatures, the targets of SAR-ATR mainly focus on ground vehicles, airplanes and ships [83]. In this paper, we focus on four main aspects: i) accuracy with limited labeled data; ii) electromagnetic scattering information based SAR-ATR; iii) real-time or alleviating computational costs at practical application situations, which are study hotspots at present; and iv) SAR-ATR in open environments, also called as open set recognition (OSR). The overview of DL-based methods for the SAR-ATR and classification is briefly shown as in Tab.III. The suitable evaluation metrics for the SAR-ATR are very important to evaluate the proposed algorithms performances. There are some evaluation metrics particularly used in SAR-ATR, including accuracy (Acc), precision, recall, F1 score; and confusion matrix [103].

1) Acc. Acc can be defined by the ratio of number of correctly recognized samples to the total number of samples, which is mathematically presented by

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (11)$$

where TP , TN , FP , FN denotes true positives, true negatives, false positives, and false negatives, respectively.

2) Precision. Precision denotes the recognized positive samples that are actually positive, which can be defined by

$$Precision = \frac{TP}{TP + FP}. \quad (12)$$

3) Recall. Recall denotes the recognized positive samples of input samples, which can be defined by

$$Recall = \frac{TP}{TP + FN}. \quad (13)$$

4) F1 score. F1 is a combination of both precision and recall, which can be defined by

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}. \quad (14)$$

5) Confusion matrix (CM). CM is usually presented by an informative table. The diagonal entries of which are the number of correct recognition samples for each targets category and the non-diagonal entries are the confusions incorrectly recognizing actual targets samples in the rows into other classes.

A. Improving Accuracy with Limited Labeled Data

The limited data has two aspects: one is that the real number of data is small. Data augmentation method can be used to increase the data samples for the DL training. The other is that the number of labeled data is small, while there are many other unlabeled data. The semi-supervised or unsupervised learning can be used to improve the generalization ability of DL model. SAR-ATR with limited data also called few-shot learning

[84], the current solution methods can be roughly classified into four categories: i.e., data / features augmentation based, fine-grained DNN structure design-based, transfer learning based methods, and semi-supervised and unsupervised based methods respectively.

Data / features augmentation based methods Data augmentation mainly contains simulated SAR aircraft data via specific electromagnetic simulation software (such as FEKO, CST) [85], [86], GAN-based generated new SAR data [87]. In order to produce new samples from different azimuth angles, the authors in [88] proposed a code-image-code cyclic network to generate new data samples via adversarial training style. In order to generate data to be more similar to the real samples in image domain, the authors in [87] used the cycleGAN model to transfer simulated SAR aircraft data to be more similar with real samples. This method could increase about 10% in accuracy for the SAR aircraft classification task. Moreover, the authors in [89] used Wasserstein AE to generate SAR images. Except for directly generate new data samples, the generative models are also used to address the small data scenarios. In [90], a generative deconvolutional DNN framework was developed for the zero-shot learning in SAR-ATR task. The deconvolutional module learned the faithful hierarchical representation of known targets, and automatically constructed a continuous SAR target features space during training process.

Instead of directly SAR image data augmentation, the features space augmentation methods are also used to alleviate the limited samples. In [91], a conv-biLSTM prototypical network was trained to map SAR images into a new features space to enhance the few-shot SAR-ATR task. A multilayer AE model with a Euclidean distance constrain term was proposed in [92]. The authors in [78], [93] proposed a parallel deep CNN based multi-view DL framework with limited data. In this way, multi-view data could be used as the inputs of the DL model. In addition, multi-aspect data fusing methods, such as multi-view [79], [94], [95], multi-task [96], [97], multi-scale [16], [98], and multi-dimension [99]–[101] are also used in performance improvement of SAR-ATR tasks. The quantitative and qualitative training data enhancement strategies were developed in [102] to alleviate the lack of target data and phase defocusing of the target from its velocity in the vessel recognition task with the high-resolution SAR images.

A spherical space classifier based task-specific hierarchically designed network was developed in [103] to address the problem of intraclass diversity and interclass similarity in SAR ships recognition. Feature aggregation module and feature boost module were proposed to adaptively extract the ships' features, and spherical space classifier was proposed to expand the interclass margin and compress the intraclass feature distribution by fully taking advantage of the property of spherical space. Combined with traditional histogram of oriented gradient (HOG) features, the authors in [104] proposed DL network with HOG feature fusion for preferable SAR ship recognition.

In addition, to enrich the feature learning space, the authors in [105] transformed the raw SAR image into graph structure data through graph construction technique. The overfitting

problem can be alleviated via graph CNN model training. Moreover, a novel graph metalearning method was proposed in [107] to recognize SAR targets using few labeled SAR images.

Fine-grained DNN structure design based methods In order to alleviate overfitting, the authors in [108] developed a deep memory CNN model via replace full connection layer with convolutional layer, and achieved a recognition accuracy rate of more than 99% on MSTAR. A compressUnit-based deeper CNN model was developed in [109] for better generalization in SAR image recognition tasks. In order to enhance the generalization learning of DL model, the authors in [110] proposed an efficient transferred max-slice CNN in SAR-ATR tasks. More specifically, L2-regularization term was added in loss function to perform generalization learning with small data.

For the class imbalance problem, the authors in [111] proposed a CNN training method that combines deep metric learning (DML) with gradually balanced sampling. DML can obtain the center of each class in the feature space and perform clustering equally. Gradually balanced sampling adopted a smooth transition from instance-aware resampling to class-aware resampling to improve the recognition rate. Similarly, the authors in [112] developed a multibranch expert network and dual-environment sampling to address the long-tail problems in both interclass and intraclass scenarios. In order to achieve robustness and high accuracy ships recognition, MetaBoost ensemble learning was proposed in [113] to serve as heterogeneous DCNNs model ensemble. A multibranch embedding network with bi-classifier was designed in [114] for few-shot ship recognition.

In order to improve the multiscale features ability of ships, the authors in [115] proposed a SAR ship recognition method through multiscale feature attention and adaptive-weighted classifier to enhance features in each scale, and adaptively choose the effective feature scale for accurate recognition.

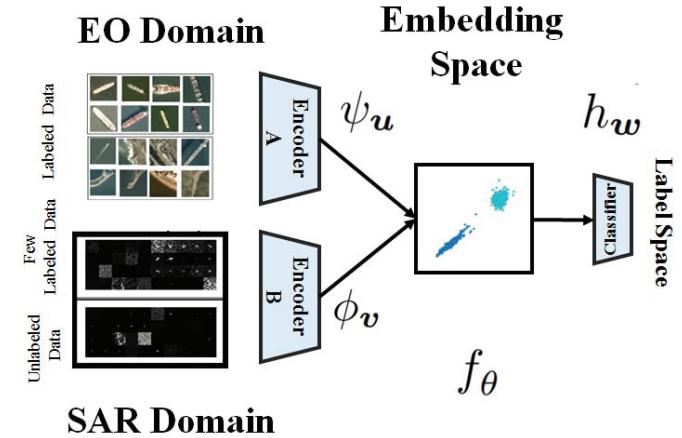


Fig. 11. The diagram of few shot learning from electro-optical domains to SAR domain [118].

Transfer learning based methods Transfer learning is a very important method to overcome the overfitting at small data scenarios. TL can transfer knowledge from one domain (i.e., source domain) to another domain (i.e., target domain). In this way, the knowledge of target domain can be enriched and can provide more learning space. In [116], the authors transferred

the knowledge from sufficient unlabeled SAR images to small labeled SAR target data. In [117], the authors transferred the prior knowledge from the optical, non-optical, hybrid optical and non-optical domains, to the SAR target recognition tasks. The authors in [118] transferred knowledge from electro-optical domains to SAR images domain, which can be shown as in Fig.11. A shared domain-invariant embedding features can be effectively learned via cross-domain knowledge transfer pattern.

To address the problems of sparse labeled samples and imbalanced categories, the authors in [119] proposed an attention-dense-CycleGAN (ADCG) method for the ship transfer learning task from optical to SAR domain, which includes a dense connection module and a lightweight convolutional block attention module.

Considering the SAR imaging variability of SAR aircraft targets, a target attitude angles (TAAs) based meta learning network was proposed in [120] to improve few-shot SAR aircraft classification performance. This network contains an angle self-adaption classifier to guide the network to focus on the positive sample pairs with different TAAs, and a pulse cosine transform based frequency embedded module to effectively extract valuable frequency-domain information. The test accuracy for 5-way 1-shot could be improved by 4.74%, compared to the baselines. In addition, the authors in [121] also proposed to combine the aircraft scattering information with the meta learning to improve the SAR-ATR at the few-shot sample scenarios.

Semi-supervised and unsupervised based methods The semi-supervised and unsupervised learning are also the effective methods to improve generalization learning capacity with small labeled data [122]. The authors in [123]–[125] proposed semi-supervised recognition methods via GANs or DCGANs model to improve the model generalization learning. Two discriminators share the same generators for joint training. As for unsupervised learning style, the authors in [126] proposed an unsupervised multi-level domains adaptation learning method through adversarial learning for the multi-band labeled SAR images recognition task.

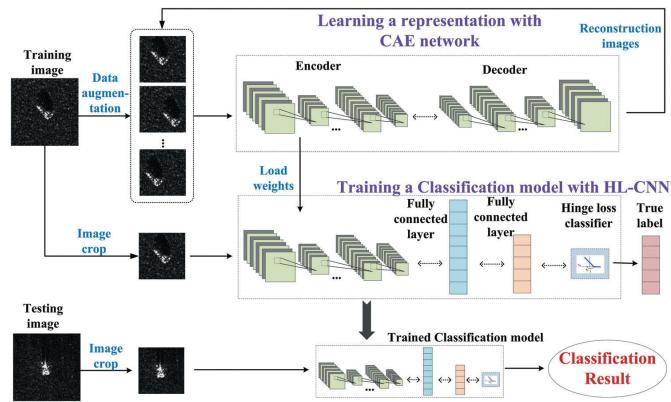


Fig. 12. Our proposed CAE-HL-CNN model [127].

We have studied the problems of few-shot samples in SAR-ATR mainly focus on the transfer learning on augmented SAR dataset [127], [128] and multi-attention model [129].

In [127], a semi-greedy neural network, namely CAE-

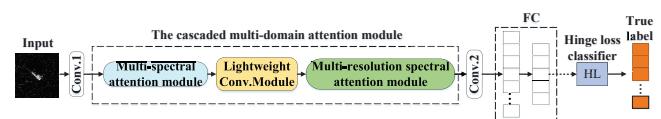


Fig. 13. Our proposed LW-CMDANet model [129]

TABLE IV: The test accuracy performances of different methods on different MSTAR data subsets [129]. Subset-100/50 represents the number of samples for each class is 100 and 50, respectively.

Models	Subset-100(%)	Subset-50(%)
Baseline CNN	89.69	83.57
A-ConvNet	90.93	82.24
ResNet	89.27	83.55
HL-CNN	90.82	83.91
MobileNet	90.02	83.91
CNN-SENet	91.73	86.11
CNN-FCANet	89.66	82.05
CNN-CBAMNet	90.84	82.25
CNN-CANet	73.87	70.41
YOLO-DMCCA	86.69	60.56
SAR-OVSM	76.35	50.57
SAR-VGG-KNN	88.54	65.66
SAR-HOG	74.58	48.97
SAR-BoVW	78.32	53.65
A-CNN-FCA-DWTNet	87.84	73.66
CNN-FCA-DWTNet	92.98	86.16
LW-CMDANet (proposed)	92.15	89.93

HL-CNN, was developed for SAR-ATR with limited training data. In this work, a semi-greedy convolutional auto-encoders (CAE) model based on transfer learning and hinge loss was constructed to balance the feature extraction and anti-overfitting capabilities of the DNN model. On the one hand, a non-greedy network with a hinge loss classifier could enhance the networks generalization performance on the small training data. On the other hand, the CAE-HL-CNN performed a complete features extraction to compensate for the degradation in features extraction via a greedy way aforementioned, through transfer learning with augmented SAR image data. The overview of CAE-HL-CNN can be shown as in Fig.12.

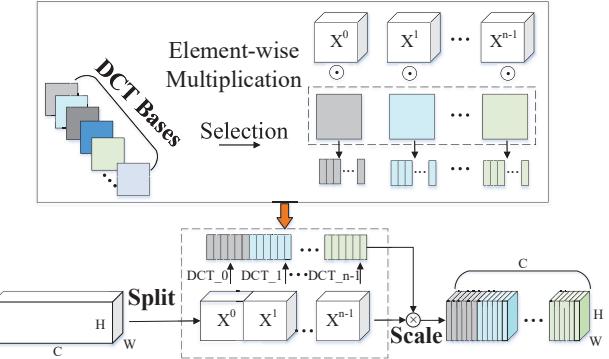


Fig. 14. FCA module [129]

Similarly, a few-shot recognition network based on inception and FCNN was proposed in [128]. Specifically, to improve the recognition accuracy rate and convergence speed rate, the amplitude domain multiplicative filtering was proposed to build the pretraining dataset, to optimize the model parameters.

In addition, the FCNN structure was effective to alleviate overfitting aforementioned.

Inspired by the attention mechanism, we have studied the effectiveness of attention in the few shot learning scenarios. Similarly to [127], we proposed a lightweight cascaded multidomain attention network, namely, LW-CMDANet in [129], as shown in Fig.13. This architecture firstly designed a four-layer CNN to perform hierarchical features extraction to alleviate the overfitting problem on the small dataset via the non-greedy learning based hinge loss function. The class-specific features extraction from both the frequency and wavelet transform domains was then performed via an attention module, embedded into the previous CNN model. Frequency channel attention module is shown as in Fig.14. In this way, the attention module could effectively compensate features extraction capacity of previous non-greedy learning style. Thus, it can improve the recognition performance. The experimental results on different number of SAR images samples shown in Tab.IV., Seeing from the experimental results has shown that our proposed LW-CMDANet has better performance in terms of accuracy on very small dateset.

B. Electromagnetic scattering information based SAR-ATR

Compared with optical images, SAR images usually lack the obvious texture and edge information, which can bring great difficulties in SAR-ATR tasks. Electromagnetic scattering information of SAR targets can be used as the prior knowledge to design effective classifier to alleviate the severe performance degradation, caused by the posture sensitivity and discrete scattering centers. In order to combine the SAR imaging mechanism with the topology geometric representation of aircraft, as shown in Fig.15, a scattering topology network was proposed in [130] to achieve desirable classification results in high-resolution SAR-ATR task, based on positional and semantic relationship between the scattering points of aircraft. More specifically, a scattering topology module was used to enhance the topology relations and scattering characteristics, through modeling the spatial relationships and semantic information interaction of discrete scattering points.

In order to take advantage of the spatial relationships between the discrete scattering centers, a GCN-based scattering features spatial-structural association network was proposed in [131] to obtain more valuable features for SAR-ATR. In this work, the strong scattering points of aircraft were extracted and converted into graph structure data. A GCN and a modified VGGNet were designed to extract the structural and high-level semantic electromagnetic scattering features and SAR image domain features of aircraft with extremely discrete appearances, respectively. Considering the SAR aircraft target imaging mechanism, a scattering characteristics analysis network was proposed in [120] to improve the aircraft targets recognition. A scattering extraction module was used to learn the number and distribution of the scattering points for the aircraft via network-based explicit supervision manner. In this way, scattering characteristics prior could assist the classification model to learn more valuable features. The test accuracy for 5-way 1-shot could be improved by 4.74%, compared to the baselines.

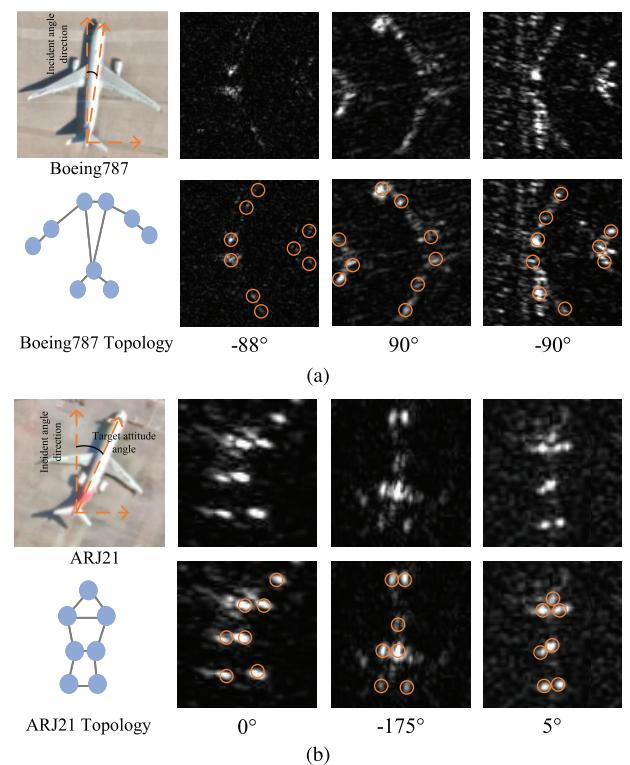


Fig. 15. The scattering topology of different aircraft [130]: (a) Boeing 787, (b) ARJ21.

As for the ships recognition, the authors in [132] developed a DL-based ship recognition network combined with scattering characteristics. A strong scattering point awareness network was specifically proposed to capture the strong scattering points that existed in the ship area. SPAN recognizes the ship category according to their distribution characteristics.

C. Promoting the Real-Time or Reducing Computational Cost

Computational cost is a significant metric for the SAR-ATR model performance. There are many works focus on this aspect. A squeezeNet was proposed in [133] for the real-time damaged regions classification based on SAR images. The authors in [134] proposed a single DCNN based direct SAR-ATR method on large-scene SAR images by encapsulating all of the computational cost. Experiments on MSTAR and large-scene SAR images showed taht this model outperformed CFAR+SVM and YOLOv2 [135]. In [136], the authors developed a novel stacked deep convolutional highway unit network.

The complex multi-view processing aforementioned, however, can cause huge computational costs in the SAR-ATR tasks. To address this problem, the authors in [137] proposed an optimal target viewpoints selection method in multi-view ATR algorithm to accelerate training and inferring learning process. An ensemble learning [138] based two-channel CNN was designed to perform multi-view learning. A heterogeneous CNN-based ensemble learning method was also employed in [139] to construct noncomplete connection scheme and multiple filters stacked.

In [140], the authors introduced depthwise separable convolution based channel and spatial attention to enhance the

feature extraction ability and reduce the computational cost. In addition, the authors in [141] alleviated the computational cost of DL model via BN layer, drop-out strategy, and concatenated ReLU activation function. In [105], the authors proposed graph representation network through target pixel grayscale decline to speed up the training time.

A shallow DL was proposed in [142] to simultaneously speed up the computation and improve classification accuracy for the full-polarization SAR images classification. A lightweight CNN, based on atrous convolution and inception module, was proposed in [117] to obtain rich global receptive fields and realize lightweight design. The authors in [143] proposed a high-speed apache spark clustering model for the classification of denoised SAR image patches. In addition, the authors in [144] developed an asymmetric parallel convolution module to alleviate the computational cost. A semi-random DNN was developed in [145] to exploit random fixed weights for real-time training and achieved the comparable accuracy, compared with general CNNs.

To effectively address the issues of low memory resources and large computational costs in SAR-ATR tasks, a two-layer micro CNN, compressed from a 18-layer DCNN via a knowledge gradual distillation, for real-time SAR recognition system was proposed in [146]. The memory footprint of this micro CNN was compressed by 177 times, and the computational cost was 12.8 times less, compared to original DCNN. Similarly, the authors in [147] developed three strategies of network compression and acceleration, to decrease the computational cost.

Instead of DNN model, the authors in [148] constructed a novel multi-grained cascade forest (gcForest) to perform SAR image targets recognition on MSTAR dataset. This is the first attempt to use non-neural network model in SAR-TDR. gcForest had better performances in computational cost and interpretability, compared with DNN-based methods.

D. SAR target recognition in open environments

Most of the existing SAR-ATR methods are performed at the closed set situation, in which the testing set has the same classes with training dataset. However, the robust DL model should not only identify the known target classes but also appropriately deal with the unknown ones. In the practical scenarios, it is common to encounter various kinds of new targets. Thus, the open set target recognition (OSTR) for SAR-ATR is very significant in open-world environments. This requires the DL model should have ability of detecting unknown categories while classifying known categories.

In order to alleviate the catastrophic forgetting of deep model, the authors in [149] developed a density coverage-based exemplar selection method to choose the key samples of the old class together with new class to jointly train the DL model in SAR-ATR task. Similarly, a random sampling combination strategy was also developed in [150]. As an unsupervised learning for the OSTR, the authors in [151] proposed an unsupervised based hierarchically reconstructive latent representation learning to complement the lost information in supervised representation and obtain a preliminary closed-set result with new samples.

GAN-based multitask (i.e., classification task and abnormal detection task) learning judgement methods were proposed in [152], [153] for the OSTR. Multitask loss was specifically devised to perform well on unknown categories detection and known categories recognition base on discriminator in GAN. In order to accurately judge the new class, the authors in [154] integrated an open-set recognition module and a new class discovery module to accurately distinguish unknown class data, and classify known class data based on a K-contrast loss. In order to enhance the capacity of identify unknown categories for the DL model, a contrastive learning based pseudo-unknown class (synthesized by known categories) guided OSTAR method was proposed in [155].

A joint training of class-specific sub-dictionary learning was developed in [156] for the OSTR. The class boundaries are determined by the matched and unmatched errors of each class between the targets and each sub-dictionary. The class of target can be determined by comparing the reconstruction errors on each sub-dictionary with these class boundaries. Aiming to the unknown classes crossing different satellites, the authors in [157] proposed a spherical space domain adaptation network under open set condition. In spherical space, features of the same class of SAR images are clustered together and features of different or unknown classes are separated on the hypersphere.

Except for detecting the new classes, quickly online recognizing the new classes is also important. Combined with incremental learning, the authors in [158] developed an incremental SAR-ATR method with new class targets, which not only classifies the targets from known classes and search targets from unknown classes but also incrementally updates the classification model with these unknown class targets. The strong separability and orthogonal distributed features based incremental learning method were also proposed in [159], [160] to address the forgetting problem on the new classes learning in OSTR. The intraclass and interclass scatters were employed to compute the feature separability loss, in order to enhance the linear separability of features during incremental learning. In order to reduce the dependence of old labeled classes and effectively mitigate the issue of catastrophic forgetting, the authors in [161] developed a class incremental learning to discover new classes with self-supervised learning and knowledge distillation based on the multiview and pseudolabeling strategies.

V. SAR SHIPS DETECTION

Ship targets detection (STD) is one of important tasks in modern maritime reconnaissance and surveillance. Except for optical imagery-based STD [162], SAR images based STD is also a wide application in SAR-TDR. Different from the SAR-ATR aforementioned, the SAR images of the STD, produced by scan SAR or stripmap SAR, usually have large scale, such as the scale of each image in AIR-SARShip-1.0 [52] is 3,000 * 3,000 pixels, which contains many different scale ship targets. Therefore, it firstly perform the location (i.e., located by rectangle with four coordinates) of targets before recognition, and there is usually only one class or multi-class in the dataset, i.e., ship, as shown in Fig.8.

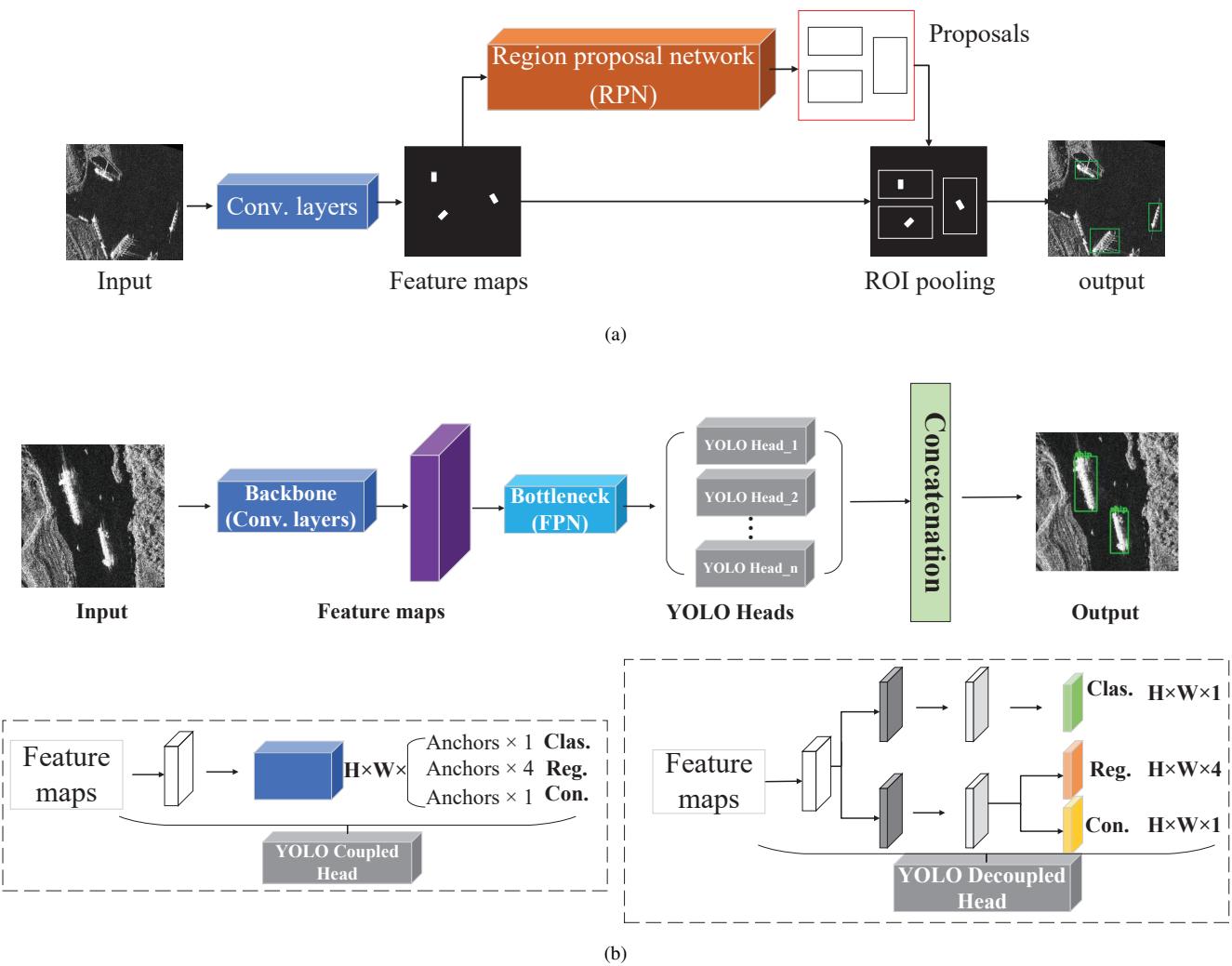


Fig. 16. The diagram of ships detection models. (a) two-stage; (b) single-stage.

Traditional STD approaches include constant false alarm rate (CFAR) based on the distribution of sea clutter [163], [164], manually features extraction based on ML algorithm [165]–[168], such as dictionary-based sparse representation, SVM, and KNN. Traditional methods, however, intensively depend on statistic modeling and experts' features extraction capacity. This may degrade the detection performances to some extend.

In recent years, DL-based methods have increasingly applied in objects detection domain. These DL methods can be roughly classified into two categories, i.e., two-stage methods and one-stage methods, respectively, as shown in Fig.16. The former contains two parts: generation of candidate regions and classifying (i.e., classification task), location of candidate regions (i.e., regression task), such as region convolutional neural networks (R-CNN) [169], fast R-CNN [170], faster R-CNN [171], mask R-CNN [172], cascade R-CNN [173], feature pyramid networks (FPNs) [174]. Compared to one-stage methods, the two-stage methods usually have higher accuracy, but involving high computational cost. Differently, the one-stage methods detect the objects by directly obtaining objects' position coordinates and the probability of objects' classes based on the extracted features of input image, which considers both accuracy and computational cost, such as You

Only Look Once (e.g., YOLOv1-v5 [135], [175]–[178], poly-v3 [179], YOLOX [180]), single shot multiBox detector (SSD) [181], RetinaNet [182]. These methods mainly include a backbone network (i.e., features extraction), bottleneck network (i.e., features fusion), and a detection head (i.e., prediction of target's category and position regression).

Furthermore, DL-based targets detection methods usually contain anchor-based and anchor-free methods, based on whether the anchors are used or not. In the anchor-based methods, such as faster R-CNN [171], RetinaNet [182], many anchors are built for each pixel of the image in advance. Dense candidate anchor regions usually need heavy computational cost. However, only a small number of positive samples are contained in candidate regions. Therefore, the problem of imbalanced positive and negative samples can severely degrade the detection efficiency. In addition, anchors-based methods usually have poor generalization ability since a large of parameters related to anchors. Differently, anchor-free methods, such as FCOS, CornerNet, can directly use the features, extracted through the backbone network, to predict the target's categories and positions. Thus, anchor-free methods have low computational cost and are more suitable for actual targets detection scenarios [183].

Take YOLOX model as an example of one-stage model, as

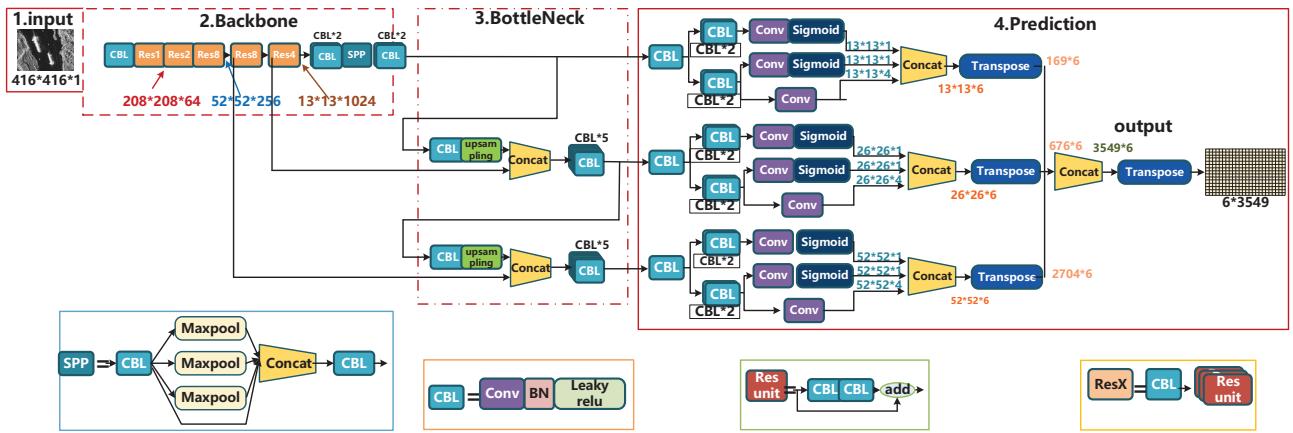


Fig. 17. The YOLOX model [180].

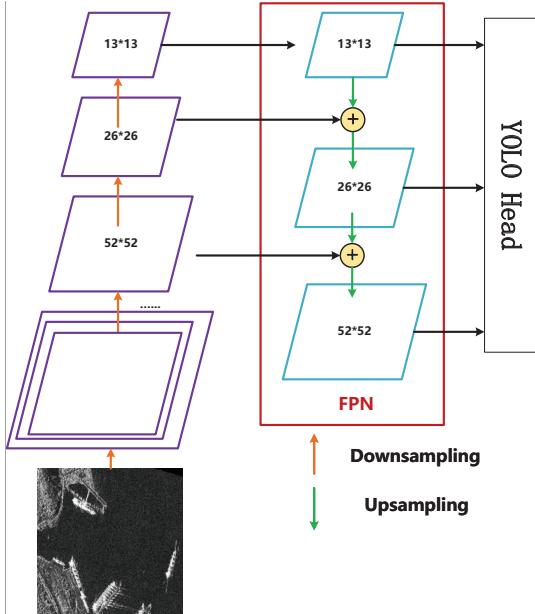


Fig. 18. The diagram of feature pyramid network.

shown in Fig.17. The network architecture of ships detection usually contains backbone, bottleneck, and YOLO heads, which is used to features extraction, features fusion (performed by FPN, as shown in Fig.18.), and the prediction of detection results. The prediction outputs usually contain six dimensions, i.e., target class, bounding box (four coordinate values), and confidence.

Nowadays, the SAR researchers have successfully applied DL-based objects detection in STD, which mainly focuses on the performances in terms of accuracy and real-time for the detection models: (i) ships on the SAR image often have a large aspect ratio and random directions. Traditional detection algorithms often unconsciously produce many false alarms. Thus, it is difficult to accurately locate the targets in complex detection backgrounds (such as background interference, clutter, inshore and outshore scenes, e.g., the GF-3 images has 86 backgrounds [184]); (ii) ships on the SAR images have many types of shape scales due to the multi-resolution SAR imaging modes (such as GF-3 has four resolutions: 3 m, 5 m, 8 m, and 10 m [184]), which poses a considerable challenge for accurate

TABLE V: The overview of DL-based methods for the SAR-STD.

Classes	Methods
General Model	faster R-CNN [22], [186], mask R-CNN [187], RetinaNet [184]
Improving Location Accuracy	FCN+clutter prior [188], multi-level sparse learning [24], U-Net based segmentation [23], attention mechanism [191], [192], loss function design [186]
Improving Detection Accuracy of Densely Arranged Ships	soft-NMS [192], [213], [219]
Solving the Problems of Multi-scale Variations	FPN [24], [184], densely connected multi-scale faster R-CNN [186], dense attention FPN [210], [211], loss function design [192], features fusion [191], [212], [213], transfer learning [135], [214], [215], [217], [218]

STD; iii) ships are often densely arranged at inshore scenes, thus, it is very complicated for the accurate detection of these ships; and (iv) the requirement of real-time detection.

In this section, we make a comprehensive survey on DL-based STD, which mainly focuses on solving accuracy and real-time challenges aforementioned. The overview of DL-based methods for the ships detection is shown as in Tab.V. The evaluation metrics of targets (including ships, vehicles, aircrafts, oil tanks, and so on) detection contains *Precision*, *Recall*, as shown in Eq.(12) and (13), respectively; *F1 score*, as shown in Eq.(14), *P-R curve* (as shown in Fig.19), *Average Precision*, and *mean Average Precision*, as shown in Eq.(15). In addition, computation cost, model size or parameters size are also used to evaluate the performance of ship detection algorithms.

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q}, \quad (15)$$

where *TP*, *FP*, *FN*, *AP*, and *Q* represent the samples of true positives, false positives, false negatives, average precision,

and number of targets on a SAR image, respectively.

The authors in [22], [185] investigated improved or lightweight faster R-CNN architectures [171] in STD. A new dataset and four strategies (such as features fusion, transfer learning, hard negative mining) were developed to achieve better performance (e.g., accuracy, computational cost) than the standard faster R-CNN. To solve the multi-scale and multi-scene STD problems, the authors in [186] proposed a densely connected multi-scale DNN model based on faster R-CNN framework. Connecting the ships detection with segmentation, an improved mask R-CNN model [172] was developed in [187] to accurately detect and segment ships. RetinaNet [182] model was also used for multi-scale ships detection in [184].

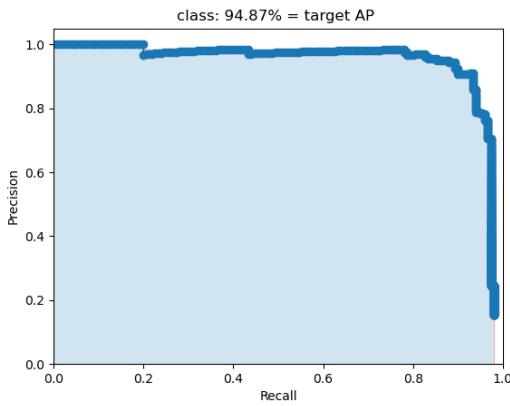


Fig. 19. The example of P-R curve, blue line is P-R curve, shadow area is average precision.

However, the objects detection models, originated from the optical image detection tasks, may not suitable for ships detection when directly applied them in STD tasks. Since the specific SAR imaging principles. Many STD specific algorithms are proposed to address the challenges in STD. These specific algorithms are still based on basic objects detection models, e.g., FPNs [174], R-CNN [170].

A. Improving Detection Accuracy in the Complex Background

Generally, it is difficult to accurately locate the targets under the heavy sea clutter or interference backgrounds. In order to reduce the false alarm rate, the authors in [184] applied RetinaNet in STD to alleviate the distortion of sea clutter. RetinaNet could achieve a more than 96% of mean average accuracy to detect multi-scale ships on GF-3 SAR images. The authors in [188] proposed a new land mask strategy to firstly separate the sea area from the land through FCN, and then design a probability distribution model of CFAR detector to accurately achieve STD on GF-3 SAR images. To fully fuse the polarization characteristics, a dualistic cascade CNN framework was proposed in [189] to alleviate the false alarm caused by clutter in PolSAR ships detection, including a backbone feature extraction network with parallel dualistic cascade architecture, a basic geometric feature extraction network, and a polarization feature enhancement network, which provided the comprehensive feature representation of targets via feature fusion. The authors in [190] designed a scattering characteristic-aware fully polarized SAR ship detection network to improve ships detection at different backgrounds,

which contains a four-component decomposition model and C-NR features extraction model. The four-component decomposition including surface scattering, double-bounce scattering, $\pm 45^\circ$ oriented dipole, and asymmetric scattering components.

The authors in [24] studied the multi-level sparse optimization method to extract discriminative features of SAR images, and effectively suppress the clutter and sidelobe interference. A U-Net based segmentation method was proposed in [23] to alleviate the false alarm issue caused by sea clutter. The attention mechanism was developed in [191], [192] to extract contextual features of SAR images and rule out the false alarm in complex scenarios. In order to alleviate the interference of background noise on the feature map, the authors in [208] designed a soft-threshold quantization module in the ship detection network based on the soft threshold function. The authors in [186] proposed a new training strategy to more focus on hard examples and reduce the weights of easy examples via loss function. In order to improve the ships positioning accuracy, the authors in [183] developed a category-position based FCN in STD. This model could enhance ships positioning performance via generating some guidance vectors from the features of classification branch.

In order to alleviate the specifical intraclass imbalance issue in a single-class based ships detection task, the authors in [193] proposed an explainable evidence learning method to correct biased learning under intraclass imbalance with uncertainty label of samples via contrastive learning. Considering the pixel-level information, the authors in [194] proposed a ship detection method based on a feature interaction network in SAR images combining object-level with pixel-level information to improve ships detection performance at the complex scenes. Combining the advantages of traditional non-DL methods and DL approaches, a fast progressive ships detection method in large SAR images was proposed in [195]. More specifically, a false alarm discrimination network was constructed to further remove false alarms.

To effectively address the issue of different distributions of training and test data, the authors in [196], [197] developed an automatic SAR image ship detection method based on feature decomposition and feature alignment crossing different satellites. More specifically, based on the adversarial domain adaptation learning strategy, the local features and global features were decomposed into domain specific features and domain-invariant features to improve object regression localization via vector decomposition method.

To reduce the dependence of network training on expensive target-level annotations, the authors in [198] developed a semisupervised SAR ship detection network via scene characteristic learning through utilizing the scene-level annotations of SAR images to improve the detection performance in the case of limited target-level annotations.

To improve the ships detection efficiency, an inception-text CNN model based non-image domain ships detection was studied in [199] in the SAR range-compressed domain. This method can filter out the invalid range-compressed data of the sea surface area, and can significantly reduce the amount of data for subsequent SAR imaging. Aiming to improve the open set ships detection, the authors in [200] proposed

a knowledge distillation based continual learning method to solve the problem of SAR ship targets incremental detection through predicted location probability representation and proposal selection.

B. Solving the Problems of Multi-scale Variations

The authors in [201] proposed a lightweight YOLOv4 based decoupled head and coordinate attention detection method for accurately detecting small targets in ship target detection. The authors in [202] proposed a YOLOX-based multiscale enhancement representation learning method to balance the accuracy and learning speed, which specifically developed an attention enhancement PAN module (including channel and spatial attention) and a novel detection head, namely ESPHead, to improve the detection capacity for targets with different scales. The mAP of proposed method can reach 0.977, while the Flops is only 9.86 G. The authors in [203] proposed a one-stage ship detector to capture multiscale contextual information through a receptive field increased module.

In order to learn more information about the small-scale ships in complex scenes, an attention-guided balanced pyramid was proposed in [204] to semantically balance the multiple features across different levels based on the anchor-free based feature balancing and refinement network, as shown in Fig.20. A feature refinement and reusable module was developed in [205] to refine small ship features based on an improved FCOS object detector. As a two-stage detector, an attention receptive pyramid network was proposed in [206] to improve the performance on detecting multiscale ships in SAR images, through representing the relationships among nonlocal features, and refining information at different feature maps.

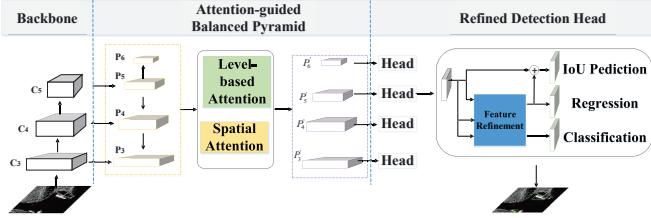


Fig. 20. The anchor-free based feature balancing and refinement network [204]

An anchor-free based balance attention network was developed in [207] to accurately detect the ships with diversity of scales and rotation angles, which balanced the local and nonlocal features of input SAR images through deformable convolution based local and non attention module, respectively. A local and global context fusion module was also designed in [208] to capture the contextual information of the target to enhance the detection of small targets. A multifeature transformation and fusion network was proposed in [209] to improve the small ship targets detection, which performed the fusion of local features and rich contextual features of ship targets.

More importantly, it is very significant to design the optimal STD algorithms to solve the issues of multi-scale variations. In [184], a focal loss function was used to address the classes imbalance. The authors in [186] proposed a densely

connected multi-scale model to densely connect feature maps from top to down of the model. Thus, each fused feature map could produce more positive proposals from multi-scale SAR images. Similarly, a dense attention FPN was proposed in [210], [211] to effectively fuse multi-layer feature maps.

As for the diverse scale issues, an improved generalized IOU loss was proposed in [192] to reduce the scale sensitivity during training learning process. The authors in [212] proposed a bi-directional features fusion based lightweight feature optimizing network, to improve the salient features extraction capacity. As for the detection of multiscale and arbitrary oriented ships in complex backgrounds, a multiscale adaptive recalibration network was developed in [213] to improve robustness of the network for the different angles of targets. More specifically, a pyramid anchor and a specific loss function were proposed to speed up the detection process.

In order to enhance the detection capacity of small ships, the authors in [191] proposed a contextual region hierarchical DNN model to fuse deep contextual semantic features and shallow high resolution features through multilayer fusion strategy. Similarly, the authors in [214], [215] proposed a scale features transfer technique to deeply mine the multi-scale features of ships on SAR images. The latent connection FPN module can densely connect feature maps from top to down via scale-transfer layer. In this way, the valuable information of small ships can be explicitly explored. The learning strategy of smaller sub-images division of SAR image was proposed in [24] to improve the features extraction ability of small targets. Considering the SAR special imaging mechanism, a sidelobe-aware small ship detection network for SAR imagery was proposed in [216]. Average pooling and max pooling was utilized to lowered the effects of strong scattering points outside of the ship body and enhanced the ship body information.

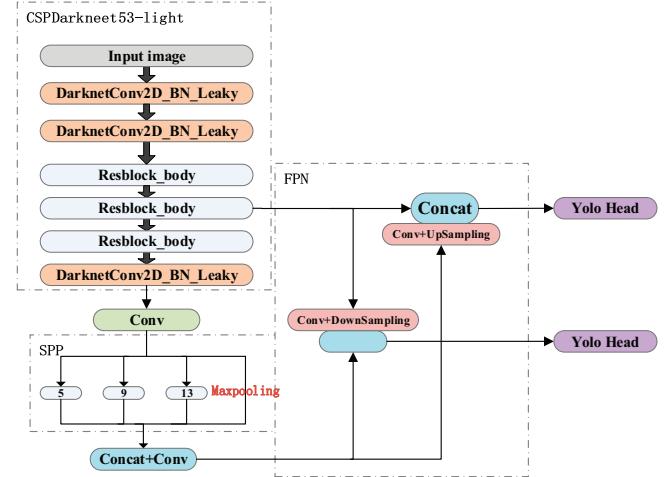


Fig. 21. The YOLO-V4-light [222].

A spatial attention block was also embedded in FPN module to enhance the extraction of spatial information. In order to study the effectiveness of TL technique in multi-scale ships detection, the authors in [217] applied a pre-trained YOLOv2 model [135] in STD task. In [218], TL strategy training based was also used to the scenario of limited number of datasets in

TABLE VI: The test accuracy performance [223].

Models	Average precision %	Model size MB
Yolov4-light	90.27	30.00
MSSDNet	95.60	25.8
Mor-FP Yolov4-tiny	96.36	-
ASAFE	95.19	-
OLF-light	95.50	10.30
ImYOLOX	97.45	4.35

TABLE VII: The accuracy performance [226].

Models	Average precision %	F1 score
SSD	68.35	80.15
FSSD	71.81	83.05
SSD+FPN	75.29	84.91
Faster-Rcnn	78.18	84.75
SSDRA	87.78	91.59

STD.

According to the previous references survey, most of existing SAR-STD methods mainly depend on low-resolution features representation. It is difficult to achieve accurate detection results in the region-level detection tasks. To address this issue, the authors in [219] proposed a high-resolution FPN based STD network to make full use of the high-resolution feature maps, through several parallel high-to-low resolution subnetworks.

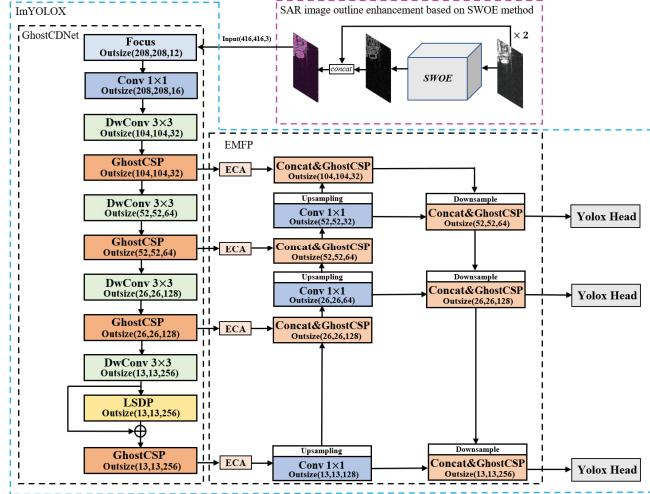
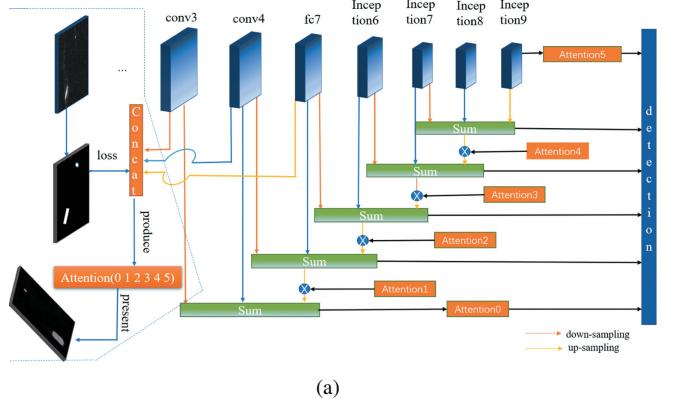


Fig. 22. The ultra-lightweight detection model based on YOLOX [223].

In addition, non-maximum suppression (NMS) method was proposed to accurately detect dense arranged ships. A soft-NMS method was introduced into the detection network model in [192], [219] to alleviate missed alarms at the dense ships overlap scenarios. The authors in [213] proposed a modified rotation NMS to address the issue of the large detection box overlap ratio.

Considering the rotation angle of each ships, the authors in [220] proposed a novel anchor-free and keypoint-based deep learning method for oriented ship detection in multiresolution SAR images. More specifically, the oriented nonnormalized Gaussian function was used to describe the center point of ship targets, while the nonuniform weighting of the different level



(a)

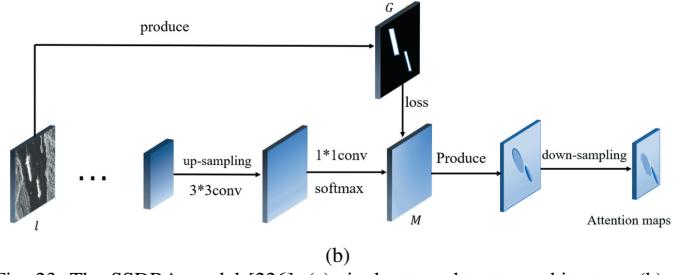


Fig. 23. The SSDRA model [226]. (a) single-stage detector architecture; (b) attention block.

loss functions was used to suppress the imbalanced sample distribution. Moreover, to alleviate the problem of multiscale, arbitrary orientation, and dense arrangement, the authors in [221] developed a YOLOV5-based SAR oriented ship detection method with contextual global attention mechanism and lightweight task-specific context decoupling, which can improve more than 4% mAP, compared to existing SOTA methods on multiple datasets.

About our works in ships detection domain, we have carried out some studies mainly focused on the lightweight model design and accuracy improvement. In [222], we proposed a high-speed lightweight ship detection model based on YOLO-V4 model, namely YOLO-V4-light, as shown in Fig.21. This model greatly reduced the number of convolutional layers in original CSPDarkNet53 (i.e., from 60 million parameters in YOLO-V4 to 6 million), which could significantly reduce the model size, detection time, number of computational parameters, and memory consumption. Moreover, the input of model was refined for the three-channel image to compensate for the loss of accuracy, due to features extraction degraded by the lightweight design of the model. Our proposed method could entirely run on a portable computer, instead of a largely powerful GPU, and achieved an average precision of 90.37% on SSDD dataset, which is better than current mainstream methods, such as YOLO-V4-tiny is 88.08%.

In order to deploy the model on the micro-SAR platforms, a YOLOX architecture based ultra-lightweight detection algorithm, namely ImYOLOX, based on ghost cross stage partial and lightweight spatial dilation convolution pyramid, was proposed in [223], as shown in Fig.22. Moreover, contextual semantic information and channel attention features were fused to perform the features enhancement, which could efficiently alleviate the performance degradation problems brought from the ship scales diversity and unbalanced distribution of channel

features information of SAR images. Experimental results on SSDD shown that our proposed method achieved an average accuracy of 97.45% with 3.31 MB parameters and 4.35 MB model size, which was better than most of current SAR ship detection algorithms, such as MSSDNet [224], LPENet [225], as shown in Tab.VI.

To improve the location accuracy of small objects under the complicated background, we proposed a single-stage detector based on the regional attention module in [226], namely SSDRA (as shown in Fig.23), which can extract detailed and semantic features information to effectively detect multi-scale and multidirectional targets. The background interference was firstly suppress through a box-level segmentation. The adaptive attention map is then produced via previous segmentation, to roughly locate the interest region via automatic learning. Finally, the multi-scale features were fused via a top-down FPN module, embedded in the multi-branch fusion structure. Experimental results on the SSDD dataset achieved a promising result, as shown in Tab.VII.

VI. SAR AIRCRAFTS DETECTION

As one of the high valuable targets, SAR-based aircraft targets detection (SAR-ATD) plays an important role in civil and military fields. Recent years, DL-based SAR-ATD has achieved great success in the performance improvement [227]. However, SAR-ATD has its specifical domain challenges and difficulties, compared to the ground vehicles and ships detection. i) Since the complex geometrical shape and special multi-component structure, the characteristics of of aircraft SAR image usually demonstrate that the scattering centers are extremely discrete without remarkable and detail edge contour features of aircrafts, as shown in Fig.24. This maybe seriously lead to false alarm detection; ii) the boarding bridges, ground vehicles, or other buildings around the aircraft can cause severe background interferences to some extent; and iii) small samples and multi-scale differences are usually difficult to achieve generalization learning with the deep neural network model. The SAR aircraft detection samples are shown as in Fig.8(b) and (c). This section presents a brief overview of the research findings in recent years in SAR-ATD.

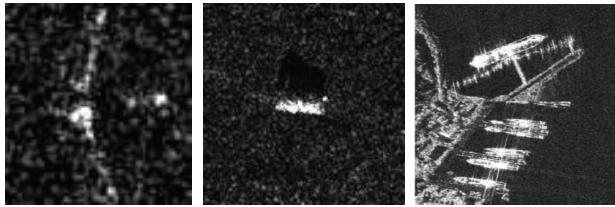


Fig. 24. The SAR images comparison between aircraft, vehicle, and ship. Left: Boeing737 aircraft, middle: vehicle, right: ship.

Similar to the previous ships detection, SAR-ATD task is also used to accurately detect the aircraft targets in large SAR images. These SAR-ATD methods can be roughly divided into three categories: i.e., electromagnetic scattering and structure components information based, multi-scale features fusion and attention based, and multi-dimensional features mining based, respectively, as shown in Fig.25.

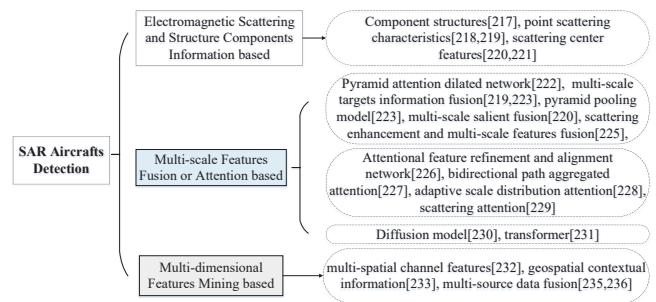


Fig. 25. The main contents of SAR-ATD.

A. Electromagnetic Scattering and Structure Components Information based SAR-ATD

In order to address the incompleteness of SAR aircraft image, caused by the highly discrete scattering centers, the prior knowledge of electromagnetic scattering features and aircraft target structure components information can be used to effectively assist the learning process of deep model, to improve the SAR-ATD performance. These methods usually fuse geometrical components structure information or electromagnetic scattering features of aircrafts with feature map features, extracted from the deep model during the learning process. One the one hand, these prior knowledge can enrich the input features of deep model. On the other hand, deep model can better learn the valuable features for the aircraft targets detection task with the guidance of these prior knowledge.

To address the sparsity and diversity brought by SAR scattering mechanism, depth features and component structures were both utilized in [228] for the SAR-ATD task. Aircraft components information could effectively assist the detection network to extract more valuable features. Take full use of scattering characteristics of aircraft SAR image can obviously improve the SAR-ATD task. The authors in [229], [230] proposed point scattering characteristics based incentive attention features fusion strategy, to improve features fusion efficiency and suppress the impact of redundant information on point features. A scattering feature relation network was proposed in [231] to obtain the desirable detection performance at the scenarios of complex interference and diversity of aircraft sizes. For the discrete strong scattering points of the aircraft, the scattering points relation module was used to fulfill the analysis and correlation of scattering points. Based on the spatial SAR imaging characteristics of the aircrafts, the authors in [232] proposed an aircraft scattering information extraction and detection model. The scattering center features extraction module was specifically designed to extract the potential scattering centers of the aircraft targets in the SAR images and acquire the overall structural features of the aircraft targets.

B. Multi-scale Features Fusion or Attention based SAR-ATD

Similar to the ships detection aforementioned, to address the multi-scale targets and complex background interferences in SAR-ATD, multi-scale features fusion or attention mechanism based methods are particularly proposed to alleviate these challenges. Multi-scale features fusion methods usually design high-efficient multi-scale features fusion module (such

as feature pyramid network, i.e., FPN), embedded into deep CNN model, to effectively fusion low-level textural and high-level semantic features. In this way, the different scale aircraft targets in SAR images can be efficiently detected with the multi-scale prediction module via trained model. The attention module is used to adaptively extract the local features corresponding to the detection tasks, so as to alleviate background interferences.

An effective pyramid attention dilated network was proposed in [233] to improve SAR-ATD accuracy. More specifically, dilated attention block was composed of two submodules, i.e., multibranch dilated convolution module (MBDCM) and convolution block attention module (CBAM). MBDCM and CBAM were used to enhance the relationship among discrete backscattering features of aircrafts, refine redundant information, and highlight significant features of aircrafts, respectively. Similarly, the authors in [230] proposed a scattering point intensity adaptive detection network to improve SAR-ATD performance with multi-scale targets information fusion and CBAM module. An improved pyramid pooling model and squeeze-and-excitation network were also proposed in [234] for the SAR-ATD.

As for the low SCNR SAR images, the authors in [235] proposed a coherent scattering enhancement and a fusion attention mechanism to suppress the clutter and speckle noise in SAR-ATD task. This network could achieve up to 91.7% of average precision in low SCNR, showing effectiveness and superiority over a number of benchmarks. In [231], a multi-scale salient fusion module was proposed to adaptively aggregate the rich semantic information and plentiful details from different scattering feature relation network. In order to reduce the false alarm and improve the localization accuracy, a contextual feature attention was also used to capture the global spatial and semantic information with a large receptive field. A scattering enhancement and multi-scale features fusion network was proposed in [236] to detect aircraft via combining traditional image processing and deep learning together. This network contained a scattering information extraction and feature enhancement module, a space-to-depth coordinate attention module, and a multi-scale features fusion pyramid module.

The authors in [237] proposed a single shot detector based attentional feature refinement and alignment network, to detect the aircrafts in SAR images with competitive accuracy and speed. This method contained attention feature fusion module, deformable lateral connection module, and anchor-guided detection module, to refine and align informative features of aircraft. A fast SAR-ATD method, namely, bidirectional path aggregated attention network, was proposed in [238]. Specifically, involution enhanced path aggregation (IEPA) module and effective residual shuffle attention (RSA) module were proposed to improve the detection accuracy of aircrafts. IEPA was used to better capture multi-scale scattering features through extracting semantic and spatial information of aircrafts, RSA could reduce false alarms via further enhancing the extracted features to overcome the interference of complex background and speckle noise.

Our team has proposed an adaptive scale distribution at-

tention (ASA) based lightweight network for multiclass SAR-ATD in [239]. The lightweight ASA was used to adaptively obtain the scattering information of multiscale aircraft targets to improve the detection accuracy, through aggregating the position and contour features of the targets. In [240], the authors designed a joint probability detector integrated with a scattering attention to alleviate the false alarms from the clutter scattering interference for SAR-ATD in large-scale SAR images.

In recent years, with the rapid development of generative artificial intelligence, the diffusion and transformer based SAR-ATD methods are also emerging. In order to explore the capacity of diffusion model in SAR-ATD, a diffusion-based SAR-ATD network, i.e., DiffDet4SAR was proposed in [241] to address the discrete scattering points and background clutter interferences. DiffDet4SAR mapped the SAR-ATD task to a denoising diffusion process of bounding boxes without heuristic anchor size selection, which can effectively enable large variations in aircraft sizes to be accommodated, and design a scattering feature enhancement module to further alleviate the clutter interference and enhance the target saliency. DiffDet4SAR can improve by 6% of mAP50, compared to the state-of-the-art methods. In addition, the authors in [242] proposed a transformer-based deformable scattering feature correlation network to address the discrete scattering points in SAR-ATD task. A scalable swin transformer backbone was used to extract hierarchical scattering features from multiscale aircraft. A deformable region correlation module was used to extract aircraft salient features in order to adapt for varying imaging appearances of SAR aircraft.

C. Multi-dimensional Features Mining based SAR-ATD

Different from designing specifical multi-scale or attention modules, multi-dimensional features mining can also improve the SAR-ATD task, including multi-spatial channel features, geospatial contextual information, and multi-instance formation.

An adaptive detection-oriented image enhancement algorithm was proposed in [243] to address the speckle effect and excessive dynamic range problems in SAR-ATD. This algorithm provided a pseudocolor SAR image through speckle suppression network and radiational features enhancement channels combination (derived from an adaptive quantization method based on the characteristics of amplitude distribution), including speckle noise suppression channel, strong and weak scattering features channels.

A novel geospatial contextual prior enabled knowledge reasoning framework was proposed in [244] for the fine-grained aircraft detection in panoramic SAR imagery with the assistance of geospatial information, including approximate airport location and the spatial relations between the airport and aircrafts. Considering the classification performance, G-ICNet was proposed in [245] for fine-grained SAR aircraft detection. A global instance-level contrast (GIC) module was used to improve interclass divergences and intraclass compactness through contrasting a large number of different aircraft targets while keeping a small batch size.

There are some other multi-source data fusion methods in SAR-ATD task. These multi-source data including data augmentation based simulation data, optical-SAR data [246], ships and vehicle SAR images as source domains to train deep model, and transferring the knowledge into SAR-ATD task [247].

VII. OTHER SAR TARGETS DETECTION

Except for the common valuable targets (i.e., vehicles, ships and aircrafts) detection aforementioned, there are some other SAR-based targets detection tasks, such as SAR images change detection, searsurface oil spill detection, oil tanks detection, and so on. This section also presents a brief survey for these studies, as shown in Fig.26.

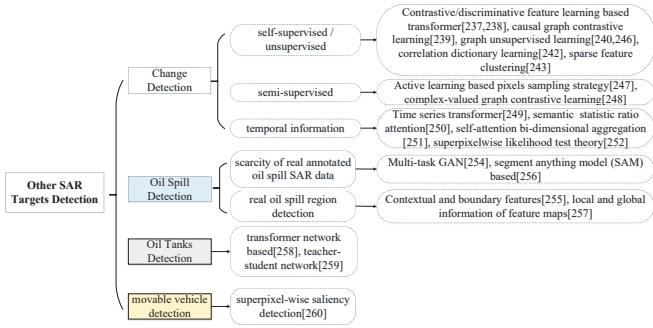


Fig. 26. The main contents of other SAR targets detection.

A. Change Detection (CD)

As one of the most active topics, change detection (CD) aims at recognizing changed regions along with the constant evolution of the earths surface. Therefore, CD can be applied in deforestation, urbanization, disaster assessment, and so on [248]. The common datasets used in CD tasks are as shown in Tab.II, including Ottawa, Sulzberger, Sulzberger, and so on.

In order to alleviate the requirement of large labeled dataset and scarce dataset to achieve robust detection, self-supervised or unsupervised CD frameworks are proposed. The authors in [248] proposed a self-supervised framework, i.e., weighted contrastive learning method with context-augmented transformer model, for the CD in multiresolution SAR image. This method can obtain rich local, global and multiscale context information, to achieve fine and robust feature expression via context-augmented swin transformer, global channel-wise aggregation module, and multiscale fusion structure. An wavelet transform-based multiscale self-supervised network was proposed in [249] for the SAR CD task, which maximizes the utilization of pseudo-labels and incorporates discriminative feature learning.

In order to exploit the capacity of causal reasoning in reducing the model's dependence on labeled samples, causal CD, i.e., causal graph contrastive learning based CD method, was proposed in [250] to obtain the optimal features representation, so as to improve performance of the self-supervised SAR CD model. Similarly, the authors in [251] proposed a graph-guided based unsupervised CD method with the with multi-attributes and multi-relationships, such as intensity and texture attributes of the SAR image.

In addition, an unsupervised feature fusion of information transfer network was proposed in [252] to obtain the discriminative feature representations, and effectively generate accurate change maps without the influence of the speckle noise. The authors in [253] also proposed an unsupervised CD method to reduce the instability of SAR image edges and details information extraction. This method contained a joint correlation dictionary learning algorithm, based on the k-singular value decomposition, with an iterative adaptive threshold optimization algorithm. In order to address the different levels of speckle noise, an unsupervised sparse feature clustering network was proposed in [254] for SAR image CD with a clustering regularization term in cross-entropy clustering loss to learn the discriminative representations of SAR images.

To use the advantage of Stockwell transform in speckle reduction, a Stockwell scattering network was proposed in [255] to perform noise-resilient feature representation in SAR image CD. The authors in [256] proposed a layer attention-based noise-tolerant network to adaptively weights the feature of different convolution layers and effectively suppresses the impact of noisy labels in SAR image CD. In order to reduce the speckle noise and extract structure features, an end-to-end dynamic graph-level neural network was embedded in [257] to exploit the local structure of each pixel neighborhood block at a graph level and learn a more discriminative graph for CD task.

In order to reduce the labeled samples for the supervised method and improve the performance of unsupervised methods, the authors in [258] proposed an active learning based pixels sampling strategy to choose the most informative samples for annotation with optical and SAR images pair. Similarly, a semi-supervised based structure-optimized complex-valued graph contrastive learning network was proposed in [259] to address the limited labeled data and inherent speckle noise in SAR image CD.

To mine more change temporal information and change pattern information from SAR image time series, an SAR time series transformer based framework was proposed in [260] for long time series SAR image CD and change pattern analysis. In semantic change detection, the authors in [261] proposed a statistic ratio attention based siamese U-Net architecture to locate land cover changes and identify their categories. In order to address the high-frequency information loss of existing attention mechanism, a wavelet-based self-attention bi-dimensional aggregation network was developed in [262] for SAR image CD. The authors in [263] introduced superpixels into the hypothesis test theory and proposed a superpixelwise PolSAR change detection method for built-up area extraction.

Except for the landmark changes detection, a global-context capturing network based coherent subtle change detection (such as vehicle tracks and footprints) method was proposed in [264] based on the amplitude and phase generates a difference image utilizing the repeat-pass repeat-geometry airborne SAR complex images.

B. Oil Spill, Oil Tanks, and Vehicles Detection

The oil spill incidents on marine surface can cause severe ecological pollution. The timely and effectively oil spill detection (OSD), therefore, is very significant. Recently, deep learning based OSD methods with SAR images have obtained great success, since the advantages of SAR sensor aforementioned. These methods mainly solve the following two challenges: i) scarcity of real annotated oil spill SAR data samples, and ii) effectively distinguishing the real oil spill region and look-alike oil spill regions (such as low wind speed areas, upwelling, leeward side, and etc.) in the SAR image [265]. In addition, the SAR-based OSD datasets are mainly based on the Sentinel-1, ERS-1/2, and GF-3 satellites. Other specific dataset such as deep-SAR Oil Spill (SOS) dataset was also proposed in [266].

As for the first issue, the authors in [265] proposed a multi-task GAN oil spill detection model to alleviate the limited training data issue via GAN-based more generated data. The authors in [267] proposed a composite oil spill detection framework to alleviate the issue of without finely annotated segmentation samples in the training stage. More specifically, a YOLOv8 detector integrated with a segment anything model (SAM) was developed to reduce the requirement of large annotated data samples, since SAM can offer impressive zero-shot segmentation performance by using detection bounding boxes and masks as input prompts.

For the second challenge, the authors in [266] developed an oil spill contextual and boundary-supervised detection network (CBD-Net) to extract refined oil spill regions by fusing multiscale features, such as spatial, channel, and boundary details features of oil spill SAR images. CBD-Net can achieve the highest mIoU of 83.42% and the highest F1 score of 87.87% on the SOS dataset, compared to the baselines. In order to take full use of the local and global information generated by the different feature maps, the authors in [268] proposed a dual-stream Unet for the OSD task. One module was an edge feature extraction module for extracting the local information and the other is an interscale alignment module for capturing the global information.

As for the oil tanks detection, the current existing deep learning based methods mainly solve the oil tanks detection problems under the condition of dense oil tanks, compactly arranged in SAR images due to the unique imaging mechanism of SAR. For example, Considering the incidence angle of the SAR image as prior, the authors in [269] proposed an end-to-end transformer network based 3-D detection method for the densely arranged oil tank targets from a single SAR image. This method can improve the more than 17% of AP, compared to the baselines. With the optical image enhancement, the authors in [270] proposed a multistage framework for oil tank detection in SAR images based on the teacher-student network.

As for the movable vehicle detection, a superpixel-wise saliency detection method was developed in [271]. The authors proposed to incorporate the superpixel generation and saliency calculation steps into an end-to-end trainable deep network. A differentiable superpixel generation method and a differentiable and computationally simple saliency model,

i.e., lacunarity cue were specifically developed to effectively highlight the vehicles' areas on real SAR images with different sizes and scenes.

VIII. DISCUSSION AND FUTURE PROMISING TRENDS

A. Discussion

The current DL-based SAR-TDR methods mainly benefits from four aspects: dataset, model design, training strategy, and computational power, respectively. Here, we only take a brief discussion for the first three aspects. Firstly, dataset is a very significant factor for the data-driven DL-based SAR-TDR methods. Different from the CV domain, there are few open SAR datasets can be used to train DL model, since the military or commercial secrets and it is certainly difficult to collect and label the clean SAR images. Although, there are some SAR datasets mentioned in section III, it is maybe far from sufficient to verify the complex DL-based SAR-TDR. More specifically, on the one hand, the current popular transformer-based DL model or large language model (LLM) has great requirement for the clean and large dataset. On the other hand, it is a challenge to collect a large SAR image dataset. Thus, the current popular transformer-based or LLM maybe not directly suitable for the SAR-TDR domain. It is necessary to design some special DL-based SAR-TDR model to match the downstream tasks.

Therefore, it is usually a small samples learning problem in DL-based SAR-TDR. There are mainly three solutions: data augmentation (such as generation of new SAR images via GAN or images processing, e.g., flipping, sub-image extraction), few-shot learning techniques (such as transfer learning, multi-view learning), and semi-supervised / unsupervised learning (such as teacher-student model, knowledge distillation), respectively. It is worth noting that, the data augmentation based domain knowledge bias problem between the generated data and practical data or different SAR platform data can be effectively alleviated via domain adversarial adaptive learning [196], [197]. Moreover, multi-source / multimodal data and microwave vision features extraction can also improve the performance at few samples scenarios.

In the practical scenarios, there is an open environment, thus, the open set target recognition and detection tasks are very significant. The traditional CNN model can combine with novel learning strategies, such as continual incremental learning [149], [200], causal reasoning learning [250], meta learning, inductive / transductive learning, to learn more robust and generalized features to improve open set target recognition and detection capacities in complex scenarios, such as zero-shot, real-time inference scenarios.

Secondly, model design is very important for the effective features extraction, which is directly related to generalization ability, accuracy, real-time, including attention structure, feature pyramid network, multi-parallel network, skip connection. Generally, more deeper model with better performance, which has been verified via a large number of CV tasks in the past ten years. However, differently, the targets of SAR image usually have only few heavy scatter points, such as SAR aircraft has extremely discrete scattering centers, which has not enough

information to be learned by the deep model. Therefore, it is useful to study lightweight and more semantical SAR image specific DL model, combined with the physical properties of targets' electromagnetic scattering. Moreover, the mobile equipment or other lightweight equipments will widely used in practical applications in the future, the computational costs and storage expenses are also significant factors to be considered during DL model design phase. In addition, transform-based generative AI model can be also designed to improve SAR-TDR tasks, such as speckle reduction, interference suppression, and data generation.

Finally, training strategy is another vital factor for the SAR-TDR performances' improvement, especially for the new class recognition in the open environment, such as adversarial training, contrastive training, transfer training, multi-phase training, multi-subnetwork training, task-specific loss function design, causal reasoning learning. Training time and learning effectiveness are mainly considered when designing the training strategy. In addition, most of the current DL-based SAR-TDR methods are mainly discriminant DL methods, which is carried out in the stable environment, i.e., the learning features space is constant. However, this discriminant learning paradigm may be difficult to produce effectiveness in practical application scenarios, since the scattering complexity and variability of SAR images. Therefore, it is necessary to develop the novel learning training strategies to match the practical requirement in SAR-TDR domain, such as award learning, exploration learning, or casual learning. For example, reinforcement learning framework based award or exploration learning can learn the more effective features in SAR-TDR via awarding or exploration behaviors, which can be more suitable for actual complex and variable environments.

B. Future Promising Trends

Despite the fact that current DL-based SAR-TDR methods have made great achievements, more attention should be paid to alleviate existing challenges in the following years when faced to the actual applications. Except for the challenges of small data samples and real-time discussed in this survey paper, current existing related research directions, such as GAN-based data augmentation, few/zero shot learning, deep model compression and accelerating methods, open set target recognition and detection, adversarial examples attack [272] are still main research contents. In addition, some other promising future research directions are listed below.

(1) Physically explainable deep learning for SAR-TDR

Black-box property of the DNN model [273] is still a severe challenge in practical DL-based SAR-TDR application domains. How does the DL model work? What features does the DL model extract on the SAR images to achieve SAR-ATR tasks? These questions are brought by the black box property of non-interpretable DL models. Therefore, in order to practically apply DL model in SAR-TDR domain, explainable DL model is a significant factor [274]. More specifically, a commander inherently wants to clearly understand the decision-making procedure of the DL models to trust the DL models, so as to effectively deploy next decision strategies. The current existing explainable artificial intelligence (XAI)

techniques, such as [275]–[282], maybe helpful to design the explainable DL-based SAR-TDR models, and speed up practical applications of these effective methods.

The pose sensitivity of targets is basically inherent characteristics for the SAR image. For example, the SAR aircraft image has obvious interclass similarity and intraclass variance since its special shape and structure. This specific characteristic has severe impact on data driven based DL model since it can cause features distribution confusion between different classes. Generally, physical imaging prior knowledge, such as SAR imaging condition, target's scattering characteristic, domain knowledges of geometry shape and structure, can be regarded as the explainable prior for the human individuals. It is very promising to adaptively guide the DL model for self-directed learning through embedding physical imaging prior knowledge during training and inferring phases. This can be also called as physical explainable machine learning [283]–[285]. Information acquisition ability of SAR image and information mining, cognition, and interpretation ability are core impact factors for the SAR-TDR tasks. It is possible to build more robust SAR target topology structure, based on the electromagnetic scattering parametrization model, as the inputs of DL model to guide the DL model performs explainable learning.

On the one hand, the traditional physical model can provide prior knowledge or domain knowledge and enrich the learning space (such as initial decomposition of original SAR images [222], [286], to effectively guide the representation learning of DL model. On the other hand, the physical model can adaptively obtain optimal parameters through the supervised learning of DL model, such as the determination of the filtering threshold value in wavelet denoising [287]. In addition, the combination of physical model and data driven may effectively improve the interpretation capacity of the DL based SAR-TDR. For example, determination of polarimetric decomposition parameters with deep learning approaches for PolSAR images may increase classification accuracy. The end-to-end mechanism-driven interpretable network for SAR-ATR under EOCs [288] is also a promising research direction.

(2) Graph structure and graph representation network for SAR-TDR

Most of the current SAR-TDR works mainly achieved in the Euclidean space, i.e., directly processing the input SAR images. However, this operation pattern may lead to limited performance, especially in semi-supervised and unsupervised learning scenarios, since only attributes information are used to SAR-TDR and few labeled data. Moreover, the SAR targets data usually has the characteristics of complex dynamic varying and diversity, which presents the non-Euclidean data characteristic. Compared to the current popular CNN, RNN and transformer models, the non-Euclidean space based GNN has obvious advantages of processing irregular data, such as graph topology. Graph structure learning can effectively extract the attributes information and relationship information of input graph vertex data simultaneously, such as graph convolutional networks (GCNs) [39], graph recurrent neural networks (Graph RNNs), graph autoencoders (GAEs), graph reinforcement learning (graph RL), and graph adversarial

methods [40], graph contrastive learning [250].

There are many latest work about graph neural networks applied in earth observation [289], which demonstrates the huge potentials of GNN in SAR-TDR domain. We have also explored the capacity of GNN in semi-supervised learning manner in literature [290]. The experimental results verify that GNN can reach more than 92.5% when the number of labeled samples is 1% for each class, i.e., less than 5 samples. In semi-supervised or unsupervised learning, the label propagation algorithm in GNN can alleviate the heavy dependence of labeled data. GNN can efficiently propagate the label information via edges-based connected vertex, to achieve SAR-TDR tasks in semi-supervised learning. Moreover, GNN can effectively extract features from huge number of unlabeled data to perform unsupervised learning. The optimized graph structure can be constructed based on the clustering algorithm, as shown in Fig.27.

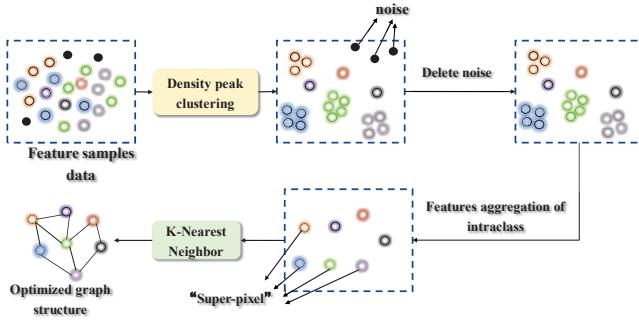


Fig. 27. The diagram of optimized graph construction.

Here, we take some examples to demonstrate the advantages of graph representation learning in SAR-TDR. For the SAR image speckle noise reduction, the superpixel-based vertex can be used in the GNN to alleviate the speckle noise, and improve the SAR-TDR performance. In the pixel-level PolSAR image land classification, every pixel of SAR image can be represented by the independent vertex of a graph and processed adaptively according to its relationship with others. Also, multiple pixel-centered subgraph can be designed and its attribute features and spatial structure feature extraction kernels are designed to complete graph convolution operation [291]. In this way, different pixels can effectively receive distinct information from the others during graph convolution process.

For the SAR aircraft target detection and recognition tasks, the scattering centers of aircraft can be regarded as the vertices of graph, edges can be the connections between the different scattering centers. Thus, a tight graph topology structure of SAR aircraft target can be constructed, which maybe alleviate the severe impact brought from extremely discrete scattering structure of aircraft SAR image. For the SAR narrowband interference (NBI) suppression, one vertex can represent a one-pulse signal, and the edges between vertices are the connections between different pulses. Thus, it is easy to imaging very tight connections between different pulses. The graph Laplacian embedding technique can be used to for time-varying NBI suppression [292]. Therefore, it is very promising to introduce the graph learning (such as graph classification)

into the SAR-TDR domain. Some works in [105], [106] have initially demonstrated the effectiveness of the graph learning in the SAR-TDR. It is significantly necessary to make a deep study for the applications of the graph learning in the SAR-TDR.

However, there are some still challenges when GNN applied in SAR-TDR: such as computation efficiency, multi-modal data merging and results explanation. Some reliable strategies, such as graph sampling (e.g., graphSAGE [293]), sparse matrix storage, heterogeneous graphs construction, multi-scale graph structure learning, dynamic vertex upgrade, physical domain knowledge embedding, maybe used to alleviate these challenges. For example, the physical attribute scattering center (ASC) can be embedded into the vertex of GNN, to increase the understanding for the GNN recognition results. The physical geometry shape graph structure design of SAR aircraft can improve the effective feature extraction and results explanation of GNN. In addition, graph regularization can also improve the robustness and explanation of GNN.

Combing with existing neural network, GNN can improve itself capacity. For example, GNN combines with CNN, GNN can improve the ability of multi-scale features extraction, since the local features extraction of CNN and the complex spatial relationship features extraction of GNN, e.g., the multi-resolution features extraction of multi-satellite SAR images. GNN combines with RNN, the time and spatial dynamic relationship of target can be conveniently modeled, since the time series features extraction of RNN and the time-space features extraction of GNN. Combining with transformer, GNN can improve the graph learning ability with the multi-head self-attention of transformer, such as graph transformer [294].

(3) Multi-modal Large language model for SAR-TDR

In recent years, with the excellent performances of ChatGPT and Sora, the multi-modal large language model (MLLM) has demonstrated great intelligence and obtained great achievements in many application domains [295], [296]. MLLM maybe a promising direction in SAR-TDR domain [297]–[299]. On the one hand, MLLM can collect the multi-modal images of targets, such as optical, SAR, infrared, etc., to fuse cross-modal knowledge representation learning [300], [301]. On the other hand, MLLM can effectively perform targets recognition in complex backgrounds through interaction between text languages and vision images, and improve the external explanation of the internal learning mechanism of DL models, since MLLM has great reasoning and cognition capacities.

The integration of multi-modal information in SAR image applications has become increasingly significant in recent years. SAR images offer unique structural and textural information, which is complementary to the spectral richness of optical images. These fusion methods include SAR and optical image fusion, pixel-level fusion, feature-level and decision-level fusion, and multi-resolution collaborative fusion. For example, SAR images, acquired by satellites like Sentinel-1, lack spectral information and are susceptible to speckle noise, which can degrade interpretability to some extent. In contrast, optical images from Sentinel-2 offer rich spectral data but are

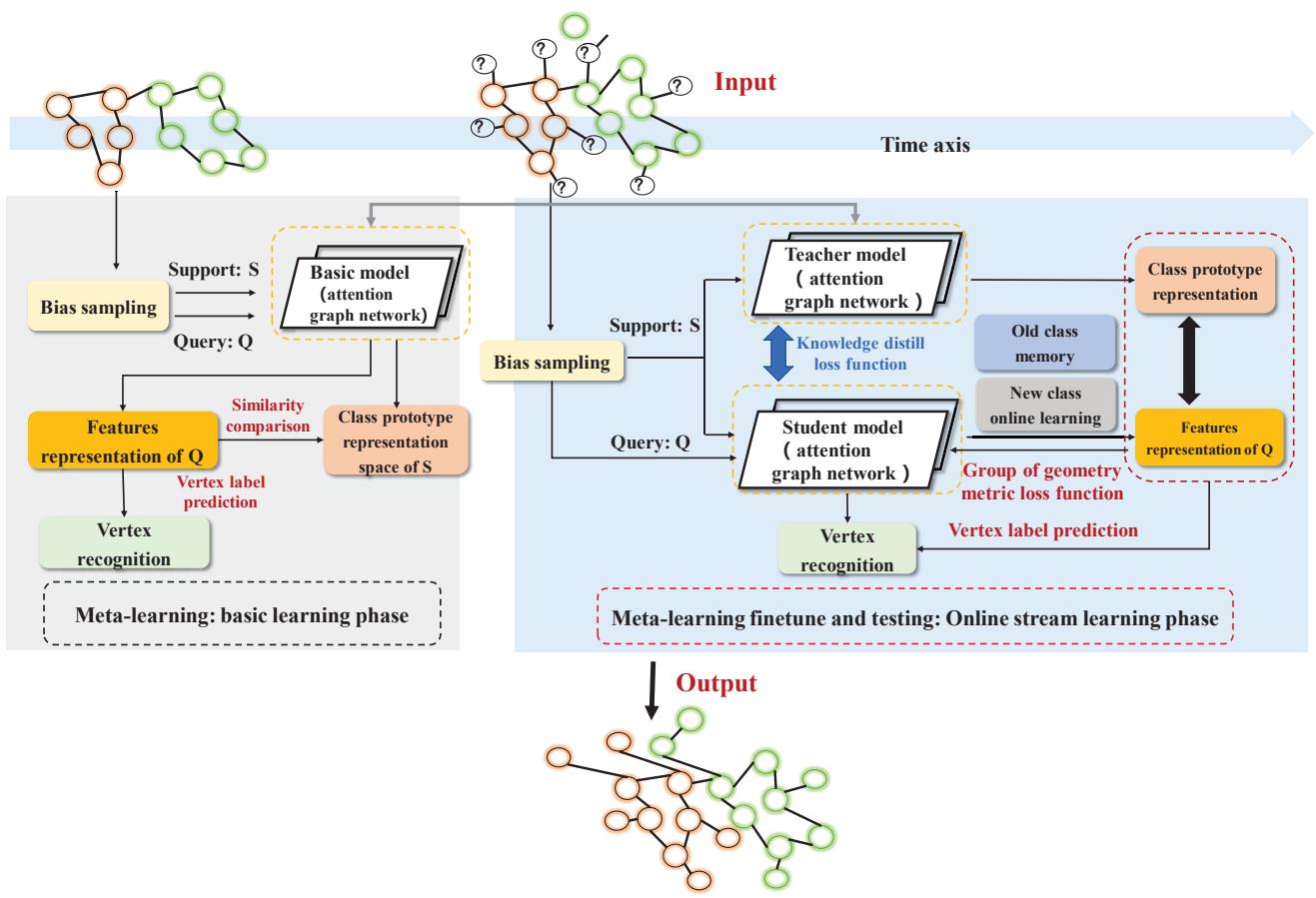


Fig. 28. The diagram of joint meta-transfer learning and graph online class incremental learning method.

limited by weather conditions and lighting. Fusion of these two types of imagery aims to leverage their complementary strengths. The multi-resolution collaborative fusion method based on SAR, multispectral, and hyperspectral images, can improve the performance of coastal wetlands mapping through cross-modal weights in both spatial and spectral dimensions to enhance robustness.

In order to improve the DL model performance, feature-level information (such as signal features, image semantic features) mining of input SAR images maybe an available method. Therefore, the novel DL model design or training methods, efficiently performed multi-domain or multi-source features extraction [302], [303], such as polarization features [304], [305], spatial features, Doppler features, is necessary. This can improve accuracy and generalization at limited data scenarios. For example, the spatial information mining from PolSAR data is very import in the land classification and ship detection. Tensor decomposition based spatial information mining, combined with DL methods, maybe a proposing direction for the PolSAR data processing.

In addition, the multi-modal data can be effectively represented by GNN, data types, such as optical, SAR, multispectral, can be denoted by vertex, edges can denote their complex dynamic relationship. For example, physical scattering or geometry shape structure prior knowledge can be added into the GNN as the learning constrain, to extract high-dimensional features with more generalization ability for

the downstream tasks, such as target detection, recognition, location, or behavior prediction.

(4) Open set recognition for SAR-TDR

In the actual dynamic open complex electromagnetic environment, unknown novel target categories are constantly emerging. Therefore, targets recognition is usually an open set recognition (OSR) problem in practical scenarios, which requires the already trained DL model to be able to quickly detect and recognize new target samples online. The incremental learning of few-shot classes maybe a promising direction to address this problem. On the one hand, the DL model can be quickly trained online with the new target classes. On the other hand, it can alleviate the catastrophic forgetting problem of class incremental learning for the old classes.

Therefore, a joint meta-transfer learning and graph online class incremental learning method maybe suitable for the OSR. Its schematic pipeline is shown as in Fig.28, which is a multi-head attention prototype (a vector representing a class in a metric space) representation learning architecture based on geometric relationship metric learning. The learning process can be roughly divided into two phase, i.e., meta basic learning and online stream learning, respectively, based on episode-based meta learning. In the meta basic learning phase, the DL model can be pre-trained through meta dataset, which enable the DL model has the ability of quick learning. While in the online stream learning phase, the pre-trained model can be used for the fine-tuning and testing for the new class

target based on geometry metric learning and teacher-student knowledge distillation.

IX. CONCLUSION

Recently, DL-based SAR-TDR tasks have drawn great attention in the remote sensing domain and achieved good performance. Contrary to the traditional physical model-based interpretation methods, deep learning can automatically extract complex semantic features representation of SAR images with multiple hierarchical convolutional layers. In this paper, we briefly introduce several DL methods (such as CNN, RNN, AEs, GNN) that are used to SAR automatic targets recognition and detection. A comprehensive survey of the existing state-of-the-art methods about DL-based SAR-TDR is then provided, including vehicles, ships, and aircrafts detection and recognition, change detection, oil spill and oil tanks detection. Our related works for SAR-TDR are also proposed and compared. These experimental results may provide some useful guidance for our future study in this domain. In addition, we have discussed some potentially promising directions that may inspire and help some researchers and practitioners to make progress in their works.

REFERENCES

- [1] S. A. Hovanessian, "Introduction to synthetic array and imaging radars," Dedham, Artech House, 1980.
- [2] M. D. Xing, H. Lin, J. L. Chen, and et al., "A review of imaging algorithms in multiplatform-borne synthetic aperture radar," *Journal of Radars*, vol. 8, no. 6, pp. 732-757, 2019.
- [3] B. Hosseiny, M. Mahdianpari, B. Brisco, and et al., "WetNet: A spatialCtemporal ensemble deep learning model for wetland classification using sentinel-1 and sentinel-2," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [4] H. Lee, H. Chae and S. J. Cho, "Radar backscattering of intertidal mudflats observed by radarsat-1 SAR images and ground-based scatterometer experiments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 5, pp. 1701-1711, May 2011.
- [5] E. R. Stofan, D.L. Evans, C. Schmullius, and et al., "Overview of results of spaceborne imaging radar-C, X-band synthetic aperture radar (SIR-C/X-SAR)," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 4, pp. 817-828, July 1995.
- [6] W. Pitz and D. Miller, "The TerraSAR-X satellite," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 615-622, Feb. 2010.
- [7] C. B.Ding, J. Y. Liu, B. Lei, and et al., "Preliminary exploration of systematic geolocation accuracy of GF-3 SAR satellite system," *Journal of Radars*, vol. 6, no. 4, pp. 11-16, 2017.
- [8] Y. Huang, G. Liao, J. Xu, and et al., "GMTI and parameter estimation via time-doppler chirp-varying approach for single-channel airborne SAR system," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4367-4383, Aug. 2017.
- [9] B. Deng, X. Li, H. Wang, and et al., "Fast raw-signal simulation of extended scenes for missile-borne SAR with constant acceleration," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 1, pp. 44-48, Jan. 2011.
- [10] S. Tang, L. Zhang, P. Guo, and et al., "An omega-K algorithm for highly squinted missile-borne SAR with constant acceleration," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 9, pp. 1569-1573, Sept. 2014.
- [11] F. Ma, F. Zhang, D. Xiang, and et al., "Fast task-specific region merging for SAR image segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.
- [12] F. Ma, F. Zhang, Q. Yin, and et al., "Fast SAR image segmentation with deep task-specific superpixel sampling and soft graph convolution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.
- [13] H. Wang, S. Chen, F. Xu, and Y. Jin, "Application of deep-learning algorithms to MSTAR data," in *Proc. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, 2015, pp. 3743-3745.
- [14] I.Goodfellow, Y. Bengio, and A. Courville, "Deep learning", *MIT Press*, ISBN: 9780262035613, 2016.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, 521 (7553), pp. 436-444, 2015.
- [16] P. F. Zhao, K. Liu, H. Zou, and et al., "Multi-stream convolutional neural network for SAR automatic target recognition," *Remote Sensing*, vol. 10, pp. 1473, 2018.
- [17] T. Zhang and X. Zhang, "High-speed ship detection in SAR images based on a grid convolutional neural network," *Remote Sensing*, vol. 11, no. 1206, 2019.
- [18] Z. Y. Cui, M. G. Zhang, and Z. J. Cao, "Image data augmentation for SAR sensor via generative adversarial nets," *IEEE Access*, vol. 7, pp. 42255-42268, 2019.
- [19] J. Ding, B. Chen, H. Liu, and et al., "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 364-368, March 2016.
- [20] H. Furukawa, "Deep learning for target Classification from SAR imagery data augmentation and translation invariance," *ArXiv*: 1708.07920, 2017.
- [21] Z. Zhao, L. Jiao, J. Zhao, and et al., "Discriminant deep belief network for high-resolution SAR image classification," *Pattern Recognition*, vol.61, pp. 686-701, 2017.
- [22] J. Li, C. Qu, and J. Shao, "Ship detection in SAR images based on an improved faster R-CNN," in *Proc.2017 SAR in Big Data Era: Models, Methods and Applications (BGSARDATA)*, Beijing, 2017, pp. 1-6.
- [23] Q. Fan, F. Chen, M. Cheng, and et al., "Ship detection using a fully convolutional network with compact polarimetric SAR images," *Remote Sensing*, vol. 11, no. 18, 2019.
- [24] F. Gao, W. Shi, J. Wang, and et al., "Enhanced feature extraction for ship detection from multi-resolution and multi-scene synthetic aperture radar (SAR) images," *Remote Sensing*, vol. 11, no. 22, 2019.
- [25] P. Lang, X. Fu, M. Martorella, and et al., "A comprehensive survey of machine learning applied to radar signal processing," *ArXiv*:2009.13702, 2020.
- [26] A. H. Oveis, E. Giusti, S. Ghio, and et al., "A survey on the applications of convolutional neural networks for synthetic aperture radar: recent advances," *IEEE Aerospace and Electronic Systems Magazine*, vol. 37, no. 5, pp. 18-42, 1 May 2022.
- [27] J. Li, J. Chen, P. Cheng, and et al., "A survey on deep-learning-based real-time SAR ship detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3218-3247, 2023.
- [28] G. Dong, G. Liao, H. Liu, and et al., "A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images," *IEEE Geoscience and Remote Sensing Magazine*, vol. 6, no. 3, pp. 44-68, 2018.
- [29] Z. Li, W. Yang, S. Peng, and et al., "A survey of convolutional neural networks: Analysis, applications, and prospects," *ArXiv*: 2004.02806v1, 2020.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *ArXiv*: 1512.03385, 2015.
- [31] N. Ma, X. Zhang, H. Zheng, and et al. "ShuffleNet V2: Practical guidelines for efficient CNN architecture design," in *Proc. European Conference on Computer Vision*, 2018, pp. 122-138.
- [32] M. S. A. Howard, G. Chu, L. Chen, and et al., "Searching for MobileNetV3," *ArXiv*: 1905.02244, 2019.
- [33] T. Mingxing and V. L. Quoc, "EfficientNet: Rethinking model scaling for convolutional neural networks," *ArXiv*: 1905.11946, 2019.
- [34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [35] M. Ravanelli, P. Brakel, M. Omologo, and et al., "Light gated recurrent units for speech recognition," *Audio and Speech Processing*, vol. 2, no. 2, pp. 92-102, 2018.
- [36] I. Goodfellow, J. Pougetabadi, M. Mirza, and et al., "Generative adversarial nets," in *Proc. Neural Information Processing Systems*, 2014, pp. 2672-2680.
- [37] Z. Wang, Q. She, T. E. Ward, and et al., "Generative adversarial networks: A survey and taxonomy," *ArXiv: Learning*, 2019.
- [38] I. Gulrajani, F. Ahmed, M. Arjovsky, and et al., "Improved training of wasserstein GANS," in *Proc. Advances Neural Information Processing Systems Conf.*, 2017.

- [39] P. Lang, X. Fu, J. Dong, and et al., "A novel radar signals sorting method via residual graph convolutional network," *IEEE Signal Processing Letters*, vol. 30, pp. 753-757, 2023.
- [40] Z. Zhang, P. Cui and W. Zhu, "Deep learning on graphs: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 249-270, 1 Jan. 2022.
- [41] C. Zhang, X. Zhang, G. Gao, and et al., "Development and application of ship detection and classification datasets: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 12, no. 4, pp. 12-45, Dec. 2024.
- [42] E. R. Keydel, S. W. Lee, and J. T. Moore, "MSTAR extended operating conditions: A tutorial," in *Proc. Aerospace/Defense Sensing and Controls*, 1996, pp. 228-242.
- [43] U. A. Force, "MSTAR overview," [online]. Available: <http://tinyurl.com/pc8nh3s>, 2013.
- [44] C. Belloni, A. Balleri, N. Aouf, and et al., "Sar image dataset of military ground targets with multiple poses for atr," in *Target and Background Signatures III*, vol. 10432. International Society for Optics and Photonics, 2017, pp. 104320N.
- [45] X. Lin, B. Zhang, F. Wu, and et al., "SIVED: A SAR image dataset for vehicle detection based on rotatable bounding box," *Remote Sensing*, vol. 15, no. 2825, 2023.
- [46] X. Hou, W. Ao, Q. Song, and et al., "FUSAR-Ship: Building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition," *Science China Information Sciences*, vol. 63, no. 4, pp. 140303, 2020.
- [47] L. Huang, B. Liu, B. Li, and et al., "OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 1, pp. 195-208, 2018.
- [48] X. Sun, Y. Lv, Z. Wang, and et al., "SCAN: Scattering characteristics analysis network for few-shot aircraft classification in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, no. 5226517, 2022.
- [49] R. Wang, H. Zhang, B. Han, and et al., "Multiangle SAR dataset construction of aircraft targets based on angle interpolation simulation," *Journal of Radars*, vol. 11, no. 4, pp. 637-651, 2022.
- [50] J. Li, C. Qu, and J. Shao, "Ship detection in SAR images based on an improved faster R-CNN," in *Proc. 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA)*, Beijing, 2017, pp. 1-6.
- [51] Y. Wang, C. Wang, H. Zhang, and et al., "A SAR dataset of ship detection for deep learning under complex backgrounds," *Remote Sensing*, vol. 11, no. 7, 2019.
- [52] S. Xian, Z. R. Wang, Y. R. Sun, and et al., "AIR-SARShip-1.0: High resolution SAR ship detection dataset," *Journal of Radars*, doi: 10.12000/JR19097. (in Chinese)
- [53] S. J. Wei, X. F. Zeng, Q. Z. Qu, and et al., "HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation," *IEEE Access*, vol. 8, pp. 120234-120254, 2020.
- [54] P. Zhang, H. Xu, T. Tian, and et al., "SEFEPNet: Scale expansion and feature enhancement pyramid network for SAR aircraft detection with small sample dataset," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 3365-3375, 2022.
- [55] Z. Wang, Y. Kang, X. Zeng, and et al., "SAR-AIRcraft-1.0: High-resolution SAR aircraft detection and recognition dataset," *Journal of Radars*, vol. 12, no. 4, pp. 906-922, 2023.
- [56] J. Chen, Z. Huang, R. Xia, and et al., "Large-scale multi-class SAR image target detection dataset-1.0[OL]", *Journal of Radars*, 2022. <https://radars.ac.cn/web/data/getData?dataType=MSAR>.
- [57] Y. Li, X. Li, W. Li, and et al., "SARDet-100K: Towards Open-Source Benchmark and ToolKit for Large-Scale SAR Object Detection", *arXiv preprint arXiv: 2403.06534*, 2024.
- [58] Y. Ye, X. Teng, S. Chen, and et al., "3MOS: Multi-sources, Multi-resolutions, and Multi-scenes dataset for Optical-SAR image matching", *arXiv preprint arXiv: 2404.00838*, 2024.
- [59] W. Xu, X. Yuan, Q. Hu, and et al., "SAR-optical feature matching: A large-scale patch dataset and a deep local descriptor," *International Journal of Applied Earth Observation and Geoinformation*, vol. 122, no. 103433, 2023.
- [60] X. Qu, F. Gao, J. Dong, and et al., "Change detection in synthetic aperture radar images using a dual-domain network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022, Art no. 4013405.
- [61] Y. Gao, F. Gao, J. Dong, and et al., "Change detection from synthetic aperture radar images based on channel weighting-based deep cascade network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 11, pp. 4517-4529, Nov. 2019.
- [62] F. Gao, J. Dong, B. Li, and et al., "Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine" *Journal of Applied Remote Sensing*, vol. 10, no. 4, Dec. 2016.
- [63] D. A. Jimenez-Sierra, H. D. Benitez-Restrepo, H. D. Vargas-Cardona, and et al., "Graph-based data fusion applied to: Change detection and biomass estimation in rice crops," *Remote Sensing* vol. 12, no. 17, 2020.
- [64] I. G. Rizaev and A. Achim. "SynthWakeSAR: A synthetic SAR dataset for deep learning classification of ships at sea", *Remote Sensing*, vol. 14, no. 16, pp. 3999, 2022.
- [65] T. Zhang, X. Zhang, X. Ke, and et al., "LS-SSDD-v1.0: A deep learning dataset dedicated to small ship detection from large-scale sentinel-1 SAR images", *Remote Sensing*, vol. 12, no. 18, pp. 2997, 2020.
- [66] W. Liu, Y. Zhao, M. Liu, and et al., "Generating simulated SAR images using Generative Adversarial Network," *SPIE 10752, Applications of Digital Image Processing XL*, San Diego, USA, 2018, pp. 32-42.
- [67] A. Ahmadibeni, L. Borooshak, B. Jones, and et al., "Aerial and ground vehicles synthetic SAR dataset generation for automatic target recognition," *SPIE 11393, Algorithms for Synthetic Aperture Radar Imagery XXVII*, California, United States, 2020, pp. 96-107.
- [68] M. Everingham, S. M. Ali Eslami, and L. V. Gool, "The PASCAL visual object classes challenge: A retrospective," *International Journal of Computer Vision*, vol. 111, pp. 98-136, 2015.
- [69] Z. Y. Cui, M. R. Zhang, Z. J. Cao, and et al., "Image data augmentation for SAR sensor via generative adversarial nets," *IEEE Access*, vol. 7, pp. 42255-42268, 2019.
- [70] Y. Ma, L. Yan, W. Y. Zhang, and S. Yan, "SAR target recognition based on transfer learning and data augmentation with LSGANs," in *Proc. Chinese Automation Congress (CAC) 2019*, 2019, pp. 2334-2337.
- [71] Z. Wang, L. Du, J. Mao, B. Liu, and et al., "SAR target detection based on SSD with data augmentation and transfer learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, pp. 150-154, Jan. 2019.
- [72] Z. L. Huang, Z. X. Pan, and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sensing*, vol. 9, pp. 907, 2017.
- [73] Y. Yan, "Convolutional neural networks based on augmented training samples for synthetic aperture radar target recognition," *Journal of Electronic Imaging*, vol. 27, pp. 1, 2018.
- [74] J. Ding, B. Chen, H. Liu, and et al., "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 364-368, March 2016.
- [75] J. Y. Lv and Y. Liu, "Data augmentation based on attributed scattering centers to train robust CNN for SAR ATR," *IEEE Access*, vol. 7, pp. 25459-25473, 2019.
- [76] J. Y. Lv and Y. Liu, "Data augmentation based on attributed scattering centers to train robust CNN for SAR ATR", *IEEE Access*, vol. 7, pp. 25459-25473, 2019.
- [77] B. Y. Ding, G. J. Wen, X. H. Huang, and et al., "Data augmentation by multilevel reconstruction using attributed scattering center for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 6, pp. 979-983, 2017.
- [78] J. F. Pei, Y. L. Huang, and W. B. Huo, "SAR automatic target recognition based on multiview deep learning framework," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 4, pp. 2196-2210, 2018.
- [79] J. F. Pei, W. B. Huo, Q. H. Zhang, and et al., "Multi-view bistatic synthetic aperture radar target recognition based on multi-input deep convolutional neural network," in *Proc. 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) , 2018*, pp. 2314-2317.
- [80] C. Coman and R. Thaens, "A deep learning SAR target classification experiment on MSTAR dataset," in *Proc. 2018 19th International Radar Symposium (IRS)*, 2018, pp. 1-6.
- [81] Z. Zeng, X. Tan, X. Zhang, and et al., "ATGAN: A SAR target image generation method for automatic target recognition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 6290-6307, 2024.
- [82] I. H. Lee and C. G. Park, "SAR-to-virtual optical image translation for improving SAR automatic target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [83] K. El-Darymli, E. W. Gill, P. McGuire, and et al., "Automatic target recognition in synthetic aperture radar imagery: A state-of-the-art review," *IEEE Access*, vol. 4, pp. 6014-6058, 2016.

- [84] J. Yin, C. Duan, H. Wang, and et al., "A review on the few-shot SAR target recognition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 16411-16425, 2024.
- [85] R. Wang, H. Zhang, B. Han, and et al., "Multiangle SAR dataset construction of aircraft targets based on angle interpolation simulation," *Journal of Radars*, vol. 11, no. 4, pp. 637-651, 2022. (in Chinese)
- [86] A. Ahmadibeni, B. Jones, L. Borooshak, and et al., "Automatic target recognition of aerial vehicles based on synthetic SAR imagery using hybrid stacked denoising auto-encoders," *Defense + Commercial Sensing*, 2020, pp. 71-82.
- [87] L. Liu, Z. Pan, X. Qiu, and et al., "SAR target classification with cycleGAN transferred simulated samples," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, Spain, 2018, pp. 4411-4414.
- [88] Q. Song, F. Xu, and Y. Jin, "SAR image representation learning with adversarial autoencoder networks," in *Proc. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 9498-9501.
- [89] K. Wang, G. Zhang, Y. Leng, and et al., "Synthetic aperture radar image generation with deep generative models," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 6, pp. 912-916, June 2019.
- [90] Q. Song and F. Xu, "Zero-shot learning of SAR target feature space with deep generative neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2245-2249, Dec. 2017.
- [91] L. Wang, X. Bai, and F. Zhou, "Few-shot SAR ATR based on Conv-BiLSTM prototypical networks," in *Proc. 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Xiamen, China, 2019, pp. 1-5.
- [92] S. Deng, L. Du, C. Li, and et al., "SAR automatic target recognition based on euclidean distance restricted autoencoder," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3323-3333, July 2017.
- [93] Y. Hou, Y. Bai, T. Xu, and et al., "Deep convolutional neural network structural design for synthetic aperture radar image target recognition based on incomplete training data and displacement insensitivity," *Journal of Electronic Imaging*, vol. 28, no. 5, 2019.
- [94] F. Zhang, C. Hu, and Q. Yin, "Multi-aspect-aware bidirectional LSTM networks for synthetic aperture radar target recognition," *IEEE Access*, vol. 5, pp. 26880-26891, 2017.
- [95] J. Oh, G. Youm, and M. Kim, "SPAM-Net: A CNN-based SAR target recognition network with pose angle marginalization learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 2, pp. 701-714, Feb. 2021.
- [96] C. Ning, W. B. Liu, G. Zhang, and et al., "Synthetic aperture radar target recognition using weighted multi-task kernel sparse representation," *IEEE Access*, vol. 7, pp. 181202-181212, 2019.
- [97] X. Z. Zhang, Y. J. Wang, Z. Y. Tan, and et al., "Two-stage multi-task representation learning for synthetic aperture radar (SAR) target images classification," *Sensors*, vol. 17, pp. 2506, 2017.
- [98] Q. Z. Yu, H. B. Hu, X. P. Geng, and et al., "High-performance SAR automatic target recognition under limited data condition based on a deep feature fusion network," *IEEE Access*, vol. 7, pp. 165646-165658, 2019.
- [99] M. Touafria and Q. Yang, "A concurrent and hierarchy target learning architecture for classification in SAR application," *Sensors*, vol. 18, pp. 3218, 2018.
- [100] O. Kechagias-Stamatis and N. Aouf, "Fusing deep learning and sparse coding for SAR ATR," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 2, pp. 785-797, 2019.
- [101] L. Li, L. Ma, L. Jiao, and et al., "Complex contourlet-CNN for polarimetric SAR image classification," *Pattern Recognition*, vol. 100, pp. 107110, 2019.
- [102] J. Song, D. Kim, J. Hwang, and et al., "Effective vessel recognition in high resolution SAR images using quantitative and qualitative training data enhancement from target velocity phase refocusing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-14, 2024.
- [103] Y. Shang, W. Pu, C. Wu, and et al., "HDSS-Net: A novel hierarchically designed network with spherical space classifier for ship recognition in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-20, 2023.
- [104] T. Zhang, X. Zhang, X. Ke, and et al., "HOG-ShipCLSN: A novel deep learning network with HOG feature fusion for SAR ship classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-22, 2022.
- [105] H. Zhu, N. Lin, H. Leung, and et al., "Target classification from SAR imagery based on the pixel grayscale decline by graph convolutional neural network," *IEEE Sensors Letters*, vol. 4, no. 6, pp. 1-4, June 2020.
- [106] H. Liu, S. Yang, S. Gou, and et al., "Fast classification for large polarimetric SAR data based on refined spatial-anchor graph," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1589-1593, Sept. 2017.
- [107] L. Li, J. Liu, L. Su, and et al., "A novel graph metalearning method for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [108] R. H. Shang, J. M. Wang, L. C. Jiao, and et al., "SAR targets classification based on deep memory convolution neural networks and transfer parameters," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 8, pp. 2834-2846, 2018.
- [109] Y. Zhang, X. Sun, H. Sun, and et al., "High resolution SAR image classification with deeper convolutional neural network," in *Proc. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 2374-2377.
- [110] Y. K. Zhai, W. B. Deng, Y. Xu, and et al., "Robust SAR automatic target recognition based on transferred MS-CNN with L2-Regularization," in *Computational Intelligence and Neuroscience*, vol. 2019, no. 3, pp. 1-13, November 2019.
- [111] Y. Zhang, Z. Lei, H. Yu, and et al., "Imbalanced high-resolution SAR ship recognition method based on a lightweight CNN," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [112] Y. Liu, F. Zhang, L. Ma, and et al., "Long-tailed SAR target recognition based on expert network and intraclass resampling," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [113] H. Zheng, Z. Hu, J. Liu, and et al., "MetaBoost: A novel heterogeneous DCNNs ensemble network with two-stage filtration for SAR ship classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [114] G. Gao, M. Wang, P. Zhou, and et al., "A multibranch embedding network with bi-classifier for few-shot ship classification of SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1-15, 2025.
- [115] C. Wang, J. Pei, Siyi Luo, and et al., "SAR ship target recognition via multiscale feature attention and adaptive-weighted classifier," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [116] Z. L. Huang, Z. X. Pan, and B. Lei, "Transfer Learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sensing*, vol. 9, pp. 907, 2017.
- [117] Z. Ying, C. Xuan, Y. Zhai, and et al., "TAI-SARNET: Deep transferred atrous-inception CNN for small samples SAR ATR," *Sensors*, vol. 20, no. 6, 2020.
- [118] M. Rostami, S. Kolouri, E. Eaton, and et al., "SAR image classification using few-shot cross-domain transfer learning," in *Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, 2019, pp. 907-915.
- [119] G. Gao, Y. Dai, X. Zhang, and et al., "ADCG: A cross-modality domain transfer learning method for synthetic aperture radar in ship automatic target recognition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-14, 2023.
- [120] X. Sun, Y. Lv, Z. Wang, and et al., "SCAN: Scattering characteristics analysis network for few-shot aircraft classification in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-17, 2022.
- [121] Y. Lv, Z. Wang, P. Wang, and et al., "Scattering information and meta-learning based SAR images interpretation for aircraft target recognition," *Journal of Radars*, vol. 11, no. 4, pp. 652-665, 2022. (in Chinese)
- [122] Z. Yue, F. Gao, Q. Xiong, and et al., "A novel semi-supervised convolutional neural network method for synthetic aperture radar image recognition," *Cognition Computation*, vol. 13, pp. 795-806, 2021.
- [123] W. Zhang, Y. Zhu, and Q. Fu, "Semi-supervised deep transfer learning-Based on adversarial feature learning for label limited SAR target recognition," *IEEE Access*, vol. 7, pp. 152412-152420, 2019.
- [124] C. Zheng, X. Jiang, and X. Liu, "Semi-supervised SAR ATR via multi-discriminator generative adversarial network," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7525-7533, 1 Sept., 2019.
- [125] Gao F, Yang Y, Wang J, and et al., "A deep convolutional generative adversarial networks (DCGANs)-based semi-supervised method for object recognition in synthetic aperture radar (SAR) images," *Remote Sensing*, vol. 10, no. 6, 2018.
- [126] W. Zhang, Y. Zhu and Q. Fu, "Adversarial deep domain adaptation for multi-band SAR images classification," *IEEE Access*, vol. 7, pp. 78571-78583, 2019.

- [127] R. Qin, X. Fu, J. Dong, and et al., "A semi-greedy neural network CAE-HL-CNN for SAR target recognition with limited training data," *International Journal of Remote Sensing*, vol. 41, no. 20, pp. 7889-7911, 2020.
- [128] H. Chen, X. Fu, and J. Dong, "SAR target recognition based on inception and fully convolutional neural network combining amplitude domain multiplicative filtering method," *Remote Sensing*, 2022, 14, 5718.
- [129] P. Lang, X. Fu, C. Feng, and et al., "LW-CMDANet: A novel attention network for SAR automatic target recognition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 6615-6630, 2022.
- [130] Y. Kang, Z. Wang, H. Zuo, and et al., "ST-Net: Scattering topology network for aircraft classification in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-17, 2023.
- [131] C. Zhao, S. Zhang, R. Luo, and et al., "Scattering features spatial-structural association network for aircraft recognition in SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [132] Y. Sun, Z. Wang, X. Sun, and et al., "SPAN: strong scattering point aware network for ship detection and classification in large-scale SAR imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1188-1204, 2022.
- [133] Y. B. Bai, C. Gao, S. Singh, and et al., "A framework of rapid regional tsunami damage recognition from post-event TerraSAR-X imagery using deep neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 43-47, 2018.
- [134] G. Q. Huang, X. G. Liu, J. P. Hui, and et al., "A novel group squeeze excitation sparsely connected convolutional networks for SAR target classification," *International Journal of Remote Sensing*, pp. 1, 2019.
- [135] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21-26 July 2017, pp. 7263-7271.
- [136] Z. Lin, K. F. Ji, M. Kang, and et al., "Deep convolutional highway unit network for SAR target classification with limited labeled training data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1091-1095, 2017.
- [137] J. F. Pei, Y. L. Huang, and W. B. Huo, "Multi-view SAR ATR based on networks ensemble and graph search," in *Proc. 2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0355-0360.
- [138] T. G. Dietterich, "Ensemble methods in machine learning," *Multiple Classifier Systems*, pp. 1-15, 2000.
- [139] Y. Xue, J. F. Pei, Y. L. Huang, and et al., "Target recognition for SAR images based on heterogeneous CNN ensemble," *2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0507-0512.
- [140] J. Q. Shao, C. W. Qu, J. W. Li, and et al., "A lightweight convolutional neural network based on visual attention for SAR image target classification," *Sensors*, vol. 18, pp. 3039, 2018.
- [141] Z. Wang and X. Xu, "Efficient deep convolutional neural networks using CReLU for ATR with limited SAR images," *The Journal of Engineering*, vol. 2019, no. 21, pp. 7615-7618, 2019.
- [142] D. Li, Y. Gu, S. Gou, and et al., "Full polarization SAR image classification using deep learning with shallow feature," in *Proc. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Fort Worth, TX, 2017, pp. 4566-4569.
- [143] C. Ozcan, O. K. Ersoy, I. U. Ogul, and et al., "Fast texture classification of denoised SAR image patches using GLCM on spark," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 28, no. 1, pp. 182-195, 2020.
- [144] Y. Hou, Y. Bai, T. Xu, and et al., "Deep convolutional neural network structural design for synthetic aperture radar image target recognition based on incomplete training data and displacement insensitivity," *Journal of Electronic Imaging*, vol. 28, no. 5, 2019.
- [145] H. Syed, R. Bryla, U. K. Majumder, and et al., "Semi-random deep neural networks for near real-time target classification," *Algorithms for Synthetic Aperture Radar Imagery XXVI*, 2019.
- [146] R. Min, H. Lan, Z. Cao, and et al., "A gradually distilled CNN for SAR target recognition," *IEEE Access*, pp. 42190-42200, 2019.
- [147] H. Chen, F. Zhang, B. Tang, and et al., "Slim and efficient neural network design for resource-constrained SAR target recognition," *Remote Sensing*, vol. 10, no. 10, 2018.
- [148] J. H. Zhang, H. J. Song, and B. B. Zhou, "SAR target classification based on deep forest model," *Remote Sensing*, vol. 12, pp. 128, 2020.
- [149] B. Li, Z. Cui, Y. Sun, and et al., "Density coverage-based exemplar selection for incremental SAR automatic target recognition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, 2023.
- [150] X. Geng, G. Dong, Z. Xia, and et al., "SAR target recognition via random sampling combination in open-world environments," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 331-343, 2023.
- [151] Y. Lu, J. Pei, Y. Zhang, and et al., "A SAR open-set recognition method aided by hierarchically reconstructive latent representation learning," *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, Athens, Greece, 2024, pp. 8966-8970.
- [152] X. Ma, K. Ji, L. Zhang, and et al., "An open set recognition method for SAR targets based on multitask learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [153] Y. Li, H. Ren, X. Yu, and et al., "Threshold-free open-set learning network for SAR automatic target recognition," *IEEE Sensors Journal*, vol. 24, no. 5, pp. 6700-6708, 2024.
- [154] M. Chen, J. -Y. Xia, T. Liu, and et al., "Open set recognition and category discovery framework for SAR target classification based on K-contrast loss and deep clustering," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 3489-3501, 2024.
- [155] Y. Li, H. Ren, L. Miao, and et al., "Pseudo-unknown class guided-based open-set learning network for SAR automatic target recognition," *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, Athens, Greece, 2024, pp. 9548-9551.
- [156] X. Ma, K. Ji, L. Zhang, and et al., "SAR target open-set recognition based on joint training of class-specific sub-dictionary learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [157] S. Zhao, Z. Zhang, T. Zhang, and et al., "Transferable SAR image classification crossing different satellites under open set condition," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [158] X. Ma, K. Ji, S. Feng, and et al., "Open set recognition with incremental learning for SAR target classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-14, 2023.
- [159] F. Gao, L. Kong, R. Lang, and et al., "SAR target incremental recognition based on features with strong separability," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-13, 2024.
- [160] L. Kong, F. Gao, X. He, and et al., "Few-shot class-incremental SAR target recognition via orthogonal distributed features," *IEEE Transactions on Aerospace and Electronic Systems*, early access, doi: 10.1109/TAES.2024.3443014.
- [161] H. Huang, F. Gao, J. Sun, and et al., "Novel category discovery without forgetting for automatic target recognition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 4408-4420, 2024.
- [162] U. Kanjir, H. Greidanus, and K. Ostir, "Vessel detection and classification from spaceborne optical images: A literature survey," *Remote Sensing of Environment*, vol. 207, pp. 1-26, 2018.
- [163] G. Gao, L. Liu, L. Zhao, and et al., "An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1685-1697, 2008.
- [164] A. Farrouki and M. Barkat, "Automatic censoring CFAR detector based on ordered data variability for nonhomogeneous environments," in *Proc. IEE Proc.-Radar Sonar Navigation*, 2005, pp. 43-51.
- [165] X. Huang, W. Yang, H. Zhang, and et al., "Automatic ship detection in SAR images using multi-scale heterogeneities and an acontrario decision," *Remote Sensing*, vol. 7, pp. 7695-7711, 2015.
- [166] X. Xing, K. Ji, and H. Zou, "Ship classification in TerraSAR-X images with feature space based sparse representation," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, pp. 1562-1566, Nov. 2013.
- [167] F. Ma, F. Gao, J. Wang, and et al., "A novel biologically inspired target detection method based on saliency analysis in SAR imagery," *Neurocomputing*, 2019.
- [168] T. Xie, W. Zhang, and L. Yang, "Inshore ship detection based on level set method and visual saliency for SAR images," *Sensors (Basel)*, vol. 18, no. 11, pp. 3877, 2018.
- [169] R. Girshick, J. Donahue, and T. Darrell, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 23-28 June 2014, pp. 580-587.
- [170] R. Girshick, "Fast r-cnn," in *Proc. of the IEEE International Conference on Computer Vision*, Santiago, Chile, 7-13 Dec. 2015, pp. 1440-1448.
- [171] S. Ren, K. He, and R. Girshick, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2016.

- [172] K. He, G. Gkioxari, and P. Dollr, "Mask r-cnn. in *Proc. of the IEEE International Conference on Computer Vision*, Venice, Italy, 22-29 Oct. 2017, pp. 2961-2969.
- [173] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18-23 June 2018, pp. 6154-6162.
- [174] T. Y. Lin, P. Dollr, and R. Girshick, "Feature pyramid networks for object detection," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21-26 July 2017, pp. 2117-2125.
- [175] J. Redmon, S. Divvala, and R. Girshick, " You only look once: Unified, real-time object detection," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27-30 June 2016, pp. 779-788.
- [176] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *ArXiv*: 1804.02767, 2018.
- [177] B. Alexey, W. Chien-Yao, and L. H. Y. Mark, "YOLOv4: Optimal speed and accuracy of object detection," *ArXiv*: 2004.10934, 2020.
- [178] G. jocher, "yolov5", <https://github.com/ultralytics/yolov5>, 2021.
- [179] P. Hurtik, V. Molek, J. Hula, and et al., "Poly-YOLO: higher speed, more precise detection and instance segmentation for YOLOv3," *arXiv*: 2005.13243, 2020.
- [180] Ge Z, Liu S, Wang F, and et al., "YOLOX: Exceeding YOLO series in 2021," *ArXiv*: 2107.08430, 2021.
- [181] W. Liu, D. Anguelov, D. Erhan, and et al., "Ssd: Single shot multibox detector," in *Proc. of the 14th European Conference on Computer Vision*, Amsterdam, The Netherlands, 8-16 October 2016, pp. 21-37.
- [182] T. Y. Lin, P. Goyal, R. Girshick, and et al., "Focal loss for dense object detection," in *Proc. of the IEEE International Conference on Computer Vision*, Venice, Italy, 22-29 October 2017, pp. 2980-2988.
- [183] Z. Sun, M. Dai, X. Leng, and et al., "An anchor-free detection method for ship targets in high-resolution SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing*, vol. 14, pp. 7799-7816, 2021.
- [184] Y. Y. Wang, C. Wang, H. Zhang, and et al., "Automatic ship detection based on RetinaNet using multi-resolution Gaofen-3 imagery," *Remote Sensing*, vol. 11, no. 5, 2019.
- [185] Y. Li, S. Zhang, and W. Q. Wang, "A lightweight faster R-CNN for ship detection in SAR images," *IEEE Geoscience Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [186] J. Jiao, Y. Zhang, and H. Sun, "A densely connected end-to-end neural network for multiscale and multiscene SAR ship detection," *IEEE Access*, vol. 6, pp. 20881-20892, 2018.
- [187] X. Nie, M. Y. Duan, and H. X. Ding, "Attention mask R-CNN for ship detection and segmentation from remote sensing images," *IEEE Access*, vol. 8, pp. 9325-9334, 2020.
- [188] Q. An, Z. Pan, and H. You, "Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network," *Sensors (Basel)*, vol. 18, no. 2, pp. 334, 2018.
- [189] G. Gao, Q. Bai, C. Zhang, and et al., "Dualistic cascade convolutional neural network dedicated to fully PolSAR image ship detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 202, pp. 663-681, 2023.
- [190] G. Gao, C. Zhang, L. Zhang, and et al., "Scattering characteristic-aware fully polarized SAR ship detection network based on a four-component decomposition Model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-22, 2023.
- [191] M. Kang, K. Ji, X. Leng, and et al., "Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection," *Remote Sensing*, vol. 9, no. 8, 2017.
- [192] C. Chen, C. He, and C. H. Hu, "A deep neural network based on an attention mechanism for SAR ship detection in multiscale and complex scenarios," *IEEE Access*, vol. 7, pp. 104848-104863, 2019.
- [193] Y. Liu, G. Yan, F. Ma, and et al., "SAR ship detection based on explainable evidence learning under intraclass imbalance," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024.
- [194] Q. Hu, S. Hu, S. Liu, and et al., "FINet: A feature interaction network for SAR ship object-level and pixel-level detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022.
- [195] H. Jia, X. Pu, Q. Liu, and et al., "A fast progressive ship detection method for very large full-scene SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024.
- [196] S. Zhao, Y. Luo, T. Zhang, and et al., "A feature decomposition-based method for automatic ship detection crossing different satellite SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022.
- [197] S. Zhao, Z. Zhang, W. Guo, and et al., "An automatic ship detection method adapting to different satellites SAR images with feature alignment and compensation loss," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-17, 2022.
- [198] Y. Du, L. Du, Y. Guo, and et al., "Semisupervised SAR ship detection network via scene characteristic learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-17, 2023.
- [199] H. Zeng, Y. Song, W. Yang, and et al., "An Incept-TextCNN model for ship target detection in SAR range-compressed domain," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [200] Y. Tian, Z. Cui, J. Ma, and et al., "Continual learning for SAR target incremental detection via predicted location probability representation and proposal selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024.
- [201] Q. Li, D. Xiao, and F. Shi, "A decoupled head and coordinate attention detection method for ship targets in SAR images," *IEEE Access*, vol. 10, pp. 128562-128578, 2022.
- [202] H. Wan, J. Chen, Z. Huang, and et al., "AFSar: An anchor-free SAR target detection algorithm based on multiscale enhancement representation learning," *IEEE Transactions Geoscience Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [203] X. Yang, X. Zhang, N. Wang, and et al., "A robust one-stage detector for multiscale ship detection with complex background in massive SAR images," *IEEE Transactions Geoscience Remote Sensing*, vol. 60, pp. 1-12, 2022.
- [204] J. Fu, X. Sun, Z. Wang, and et al., "An anchor-free method based on feature balancing and refinement network for multiscale ship detection in SAR images," *IEEE Transactions Geoscience and Remote Sensing*, vol. 59, no. 2, pp. 1331-1344, Feb. 2021.
- [205] S. Yang, W. An, S. Li, and et al., "An improved FCOS method for ship detection in SAR images," *IEEE Journal of Selected Topics Applied Earth Observations Remote Sensing*, vol. 15, pp. 8910-8927, 2022.
- [206] Y. Zhao, L. Zhao, B. Xiong, and et al., "Attention receptive pyramid network for ship detection in SAR images," *IEEE Journal of Selected Topics Applied Earth Observations Remote Sensing*, vol. 13, pp. 2738-2756, 2020.
- [207] Q. Hu, S. Hu, and S. Liu, "BANet: A balance attention network for anchor-free ship detection in SAR images," *IEEE Transactions Geoscience and Remote Sensing*, vol. 60, pp. 1-12, 2022.
- [208] C. Zhang, G. Gao, J. Liu, and et al., "Oriented ship detection based on soft thresholding and context information in SAR images of complex scenes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024.
- [209] M. Zha, W. Qian, W. Yang, and et al., "Multifeature transformation and fusion-Based ship detection with small targets and complex backgrounds," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [210] Z. Y. Cui, Q. Li, and Z. J. Cao, "Dense attention pyramid networks for multi-scale ship detection in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8983-8997, 2019.
- [211] Q. Li, R. Min, and Z. Y. Cui, "Multiscale ship detection based on dense attention pyramid network in Sar images," in *Proc. 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 5-8.
- [212] X. H. Zhang, H. P. Wang, and C. G. Xu, "A lightweight feature optimizing network for ship detection in SAR image," *IEEE Access*, vol. 7, pp. 141662-141678, 2019.
- [213] C. Chen, C. He, C. H. Hu, and et al., "MSARN: A deep neural network based on an adaptive recalibration mechanism for multiscale and arbitrary-oriented SAR ship detection," *IEEE Access*, vol. 7, pp. 159262-159283, 2019.
- [214] N. Y. Liu, Z. Y. Cui, and Z. J. Cao, "Scale-transferrable pyramid network for multi-scale ship detection in Sar images," in *Proc. 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 1-4.
- [215] Y. C. Gui, X. H. Li, and L. Xue, "A scale transfer convolution network for small ship detection in SAR images," in *Proc. 2019 IEEE 8th Joint International on Information Technology and Artificial Intelligence Conference (ITAIC)*, 2019, pp. 1845-1849.
- [216] Y. Zhou, H. Liu, F. Ma, and et al., "A sidelobe-aware small ship detection network for synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-16, 2023.
- [217] H. M. Khan and Y. Z. Cai, "Ship detection in SAR Image using YOLOv2," in *Proc. 2018 37th Chinese Control Conference (CCC)*, 2018, pp. 9495-9499.

- [218] Y. Y. Wang, C. Wang, and H. Zhang, "Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images," *Remote Sensing Letters*, vol. 9, no. 8, pp. 780-788, 2018.
- [219] S. J. Wei, H. Su, J. Ming, and et al., "Precise and robust ship detection for high-resolution SAR imagery based on HR-SDNet," *Remote Sensing*, vol. 12, pp. 167, 2020.
- [220] J. Zhang, M. Xing, G. C. Sun, and et al., "Oriented Gaussian function-based box boundary-aware vectors for oriented ship detection in multiresolution SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022.
- [221] G. Gao, Y. Wang, Y. Chen, and et al., "An oriented ship detection method of remote sensing image with contextual global attention mechanism and lightweight task-specific context decoupling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1-18, 2025.
- [222] J. Jiang, X. Fu, R. Qin, and et al., "High-speed lightweight ship detection algorithm based on YOLO-V4 for three-channels RGB SAR image," *Remote Sensing*, vol. 13, no. 1909, 2021.
- [223] S. Li, X. Fu, and J. Dong, "Improved ship detection algorithm based on YOLOX for SAR outline enhancement image," *Remote Sensing*, vol. 14, no. 4070, 2022.
- [224] K. Zhou, M. Zhang, H. Wang, and et al., "Ship detection in SAR images based on multi-scale feature extraction and adaptive feature fusion," *Remote Sensing*, vol. 14, no. 55, 2022.
- [225] Y. Feng, J. Chen, Z. Huang, and et al., "A lightweight position-enhanced anchor-free algorithm for SAR ship detection," *Remote Sensing*, vol. 14, no. 1908, 2022.
- [226] X. Qi, P. Lang, X. Fu, and et al., "A regional attention-based detector for SAR ship detection," *Remote Sensing Letters*, vol. 13, no. 1, pp. 55-64, 2021.
- [227] R. Luo, L. Zhao, Q. He, and et al., "Intelligent technology for aircraft detection and recognition through SAR imagery: Advancements and prospects," *Journal of Radars*, vol. 13, no. 2, pp. 307-330, 2024.
- [228] C. He, M. Tu, D. Xiong, and et al., "A component based multi-layer parallel network for airplane detection in SAR imagery," *Remote Sensing*, vol. 10, no. 7, 2018.
- [229] J. Chen, H. Wang, and H. Lu, "Aircraft detection in SAR images via point features," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [230] X. Ye and C. Du, "Integrated multi-Scale aircraft detection and recognition with scattering point intensity adaptiveness in complex background clutter SAR images," *Remote Sensing*, vol. 16, no. 13, 2024.
- [231] Y. Kang, Z. Wang, J. Fu, and et al., "SFR-Net: Scattering feature relation network for aircraft detection in complex SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-17, 2022.
- [232] S. Lin, T. Chen, X. Huang, and et al., "Synthetic aperture radar image aircraft detection based on target spatial imaging characteristics," *Journal of Electronic Imaging*, vol. 32, no. 2, 2022.
- [233] Y. Zhao, L. Zhao, C. Li, and et al., "Pyramid attention dilated network for aircraft detection in SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 4, pp. 662-666, April 2021.
- [234] W. Zhu, L. Zhang, C. Lu, and et al., "FEMSFNet: Feature enhancement and multi-scales fusion network for SAR aircraft detection," *Remote Sensing*, vol. 16, no. 9, 2024.
- [235] X. Zhang, D. Hu, S. Li, and et al., "Aircraft detection from low SCNR SAR imagery using coherent scattering enhancement and fused attention pyramid," *Remote Sensing*, vol. 15, no. 18, 2023.
- [236] B. Huang, T. Zhang, S. Quan, and et al., "Scattering enhancement and feature fusion network for aircraft detection in SAR images," *IEEE Transactions on Circuits and Systems for Video Technology*, early access, doi: 10.1109/TCSVT.2024.3470790.
- [237] Y. Zhao, L. Zhao, Z. Liu, and et al., "Attentional feature refinement and alignment network for aircraft detection in SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.
- [238] R. Luo, L. Chen, J. Xing, and et al., "A fast aircraft detection method for SAR images based on efficient bidirectional path aggregated attention network," *Remote Sensing*, vol. 13, no. 15, pp. 2940, 2021.
- [239] H. Chang, X. Fu, J. Dong, and et al., "MLSDNet: Multiclass lightweight SAR detection network based on adaptive scale distribution attention," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [240] X. Xiao, H. Jia, Q. Wang, and et al., "Aircraft detection and classification based on joint probability detector integrated with scattering attention," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 2, pp. 1722-1739, April 2024.
- [241] J. Zhou, C. Xiao, B. Peng, and et al., "DiffDet4SAR: Diffusion-based aircraft target detection network for SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [242] Y. Chen, Y. Cong, and L. Zhang, "Deformable scattering feature correlation network for aircraft detection in SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [243] Y. Suo, Y. Wu, T. Miao, and et al., "Adaptive SAR image enhancement for aircraft detection via speckle suppression and channel combination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024.
- [244] R. Luo, Q. He, L. Zhao, and et al., "Geospatial contextual prior-enabled knowledge reasoning framework for fine-grained aircraft detection in panoramic SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, early access, doi: 10.1109/TGRS.2024.3487780.
- [245] D. Zhao, Z. Chen, Y. Gao, and et al., "Classification matters more: global instance contrast for fine-grained SAR aircraft detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-15, 2023.
- [246] Q. Gao, Z. Feng, S. Yang, and et al., "Multi-Path interactive network for aircraft identification with optical and SAR images," *Remote Sensing*, vol. 14, no. 16, 2022.
- [247] P. Zhang, H. Xu, T. Tian, and et al., "SEFEPNet: Scale expansion and feature enhancement pyramid network for SAR aircraft detection with small sample dataset," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 3365-3375, 2022.
- [248] H. Dong, W. Ma, L. Jiao, and et al., "Contrastive learning with context-augmented transformer for change detection in SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 17710-17724, 2024.
- [249] H. Zong, E. Zhang, X. Li, and et al., "Multiscale self-supervised SAR image change detection based on wavelet transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [250] H. Li, B. Zou, L. Zhang, and et al., "CausalCD: A causal graph contrastive learning framework for self-supervised SAR image change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-16, 2024.
- [251] J. Wang and A. Zhang, "SAR image change detection based on heterogeneous graph with multiatributes and multirelationships," *IEEE Access*, vol. 10, pp. 44347-44361, 2022.
- [252] J. Ma, D. Li, X. Tang, and et al., "Unsupervised SAR image change detection based on feature fusion of information transfer," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [253] Q. Yu, M. Zhang, L. Yu, and et al., "SAR image change detection based on joint dictionary learning with iterative adaptive threshold optimization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 5234-5249, 2022.
- [254] W. Zhang, L. Jiao, F. Liu, and et al., "Sparse feature clustering network for unsupervised SAR image change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022.
- [255] G. Chen, Y. Zhao, Y. Wang, and et al., "SSN: Stockwell scattering network for SAR image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [256] D. Meng, F. Gao, J. Dong, and et al., "Synthetic aperture radar image change detection via layer attention-based noise-tolerant network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [257] R. Wang, L. Wang, X. Wei, and et al., "Dynamic graph-level neural network for SAR image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [258] Y. Xia and X. Xu, "Active learning-enhanced pyramidal convolution Unet for change detection in optical and SAR remote sensing image pairs," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 17503-17521, 2024.
- [259] H. Li, B. Zou, L. Zhang, and et al., "Semi-supervised SAR image change detection via structure-optimized complex-valued graph contrastive learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [260] W. Li, P. Ma, H. Wang, and et al., "SAR-TSCC: A novel approach for long time series SAR image change detection and pattern analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-16, 2023.
- [261] S. Chen, X. Su, L. Zheng, and et al., "Statistic ratio attention-guided siamese U-Net for SAR image semantic change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.

- [262] J. Xie, F. Gao, X. Zhou, and et al., "Wavelet-based bi-dimensional aggregation network for SAR image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [263] F. Zhang, X. Sun, F. Ma, and et al., "Superpixelwise likelihood ratio test statistic for PolSAR data and its application to built-up area extraction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 209, pp. 33-248, March 2024.
- [264] J. Zhang, M. Xing, W. Liu, and et al., "Joint exploitation of coherent change detection and global-context capturing network for subtle changed track detection with airborne SAR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 8324-8338, 2024.
- [265] J. Fan and C. Liu, "Multitask GANs for oil spill classification and semantic segmentation based on SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2532-2546, 2023.
- [266] Q. Zhu, Y. Zhang, Z. Li, and et al., "Oil spill contextual and boundary-supervised detection network based on marine SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-10, 2022.
- [267] W. Wu, M. S. Wong, X. Yu, and et al., "Compositional oil spill detection based on object detector and adapted segment anything model from SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
- [268] C. Li, M. Wang, X. Yang, and et al., "DS-UNet: Dual-stream U-Net for oil spill detection of SAR image," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [269] C. Ma, Y. Zhang, J. Guo, and et al., "End-to-end method with transformer for 3-D detection of oil tank from single SAR image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-19, 2022.
- [270] R. Zhang, H. Guo, F. Xu, and et al., "Optical-enhanced oil tank detection in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-12, 2022.
- [271] F. Ma, X. Sun, F. Zhang, and et al., "What catch your attention in SAR images: saliency detection based on soft-superpixel lacunarity Cue," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-17, 2023.
- [272] W. Xia, Z. Liu, and Y. Li, "SAR-PeGA: A generation method of adversarial examples for SAR image target recognition network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 2, pp. 1910-1920, April 2023.
- [273] D. Castelvecchi, "Can we open the black box of AI," *Nature*, vol. 538, no. 7623, pp. 20-23, 2016.
- [274] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, pp. 52138-52160, 2018.
- [275] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. European Conference on Computer Vision*, 2014, pp. 818-833.
- [276] B. Letham, C. Rudin, T. H. McCormick, and et al., "Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model," *The Annals of Applied Statistics*, vol. 9, no. 3, pp. 1350-1371, 2015.
- [277] A. S. Maruan, D. Avinava, and P. X. Eric, "Contextual explanation networks," *ArXiv*: 1705.10301, 2018.
- [278] N. Tishby and N. Zaslavsky, "Deep learning and the information bottleneck principle," in *Proc. Information Theory Workshop*, 2015, pp. 1-5.
- [279] Q. Li, B. Wang, M. Melucci, and et al., "CNM: An interpretable complex-valued network for matching," *ArXiv: Computation and Language*, 2019.
- [280] M.Z. Guo, Q.P. Zhang, X.W Liao, and et al., "An interpretable machine learning framework for modelling human decision behavior," *ArXiv*: 1906.01233, 2019.
- [281] S. Murray, N. Kyriacos, C. Antonia, and et al., "An explicitly relational neural network architecture," *ArXiv*: 1905.10307v1, 2019.
- [282] J. Y. Mao, C. Gan, K. Pushmeet, and et al., "The neuro-symbolic concept learner: Interpreting scenes, words, and sentences from natural supervision," in *Proc. International Conference on Learning Representation (ICLR)*, 2019.
- [283] Z. Huang, X. Yao, and J Han, "Progress and perspective on physically explainable deep learning for synthetic aperture radar image interpretation," *Journal of Radars*, vol. 11, no. 1, pp. 107-125, 2022.
- [284] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, and et al., "Physics-informed machine learning", *Nature Reviews Physics*, vol. 3, no., 6, pp. 422-440, 2021.
- [285] N. Thuerey, P. Holl, M. Mueller, and et al., "Physics-based deep learning", *arXiv*: 2109.05237, 2021.
- [286] R. Qin, X. Fu and P. Lang, "PolSAR image classification based on low-frequency and contour subbands-driven polarimetric SENet," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4760-4773, 2020.
- [287] R. Qin, X. Fu, J. Chang, and et al., "Multilevel wavelet-SRNet for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [288] L. Liao, L. Du, J. Chen, and et al., "EMI-Net: An end-to-end mechanism-driven interpretable network for SAR target recognition under EOCs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-18, 2024.
- [289] S. Zhao, Z. Chen, Z. Xiong, and et al., "Beyond grid data: exploring graph neural networks for earth observation," *IEEE Geoscience and Remote Sensing Magazine*, early access, 2024, doi: 10.1109/MGRS.2024.3493972.
- [290] P. Lang, X. Fu, J. Dong, and et al., "A novel radar signals sorting method via residual graph convolutional network," *IEEE Signal Processing Letters*, vol. 30, pp. 753-757, 2023.
- [291] F. Liu, J. Wang, X. Tang, and et al., "Adaptive graph convolutional network for polSAR image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [292] Y. Huang, L. Zhang, X. Yang, and et al., "An efficient graph-based algorithm for time-varying narrowband interference suppression on SAR system," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8418-8432, Oct. 2021.
- [293] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *arXiv:1706.02216v4*, <https://doi.org/10.48550/arXiv.1706.02216>.
- [294] J. Yang, Z. Liu, S. Xiao, and et al., "GraphFormers: GNN-nested transformers for representation learning on textual graph," *arXiv:2105.02605v3*, <https://doi.org/10.48550/arXiv.2105.02605>.
- [295] J. Li, D. Li, S. Savarese, and et al., "Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models," *arXiv preprint arXiv:2301.12597*, 2023.
- [296] D. Zhu, J. Chen, X. Shen, and et al., "Minigpt-4: Enhancing vision-language understanding with advanced large language models," *arXiv preprint arXiv:2304.10592*, 2023.
- [297] W. Zhang, M. Cai, T. Zhang, and et al., "EarthGPT: A universal multimodal large language model for multi-sensor image comprehension in remote sensing domain," *arXiv preprint arXiv:2401.16822*, 2023.
- [298] Y. Hu, J. Yuan, C. Wen, and et al., "Rsgpt: A remote sensing vision language model and benchmark," *arXiv preprint arXiv:2307.15266*, 2023.
- [299] K. Kuckreja, M. S. Danish, M. Naseer, and et al., "Geochat: Grounded large vision-language model for remote sensing," *arXiv preprint arXiv:2311.15826*, 2023.
- [300] G. Gao, M. Wang, X. Zhang, and et al., "DEN: A new method for SAR and optical image fusion and intelligent classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1-18, 2025.
- [301] G. Gao, P. Zhou, L. Yao, and et al., "A bi-prototype BDC metric network with lightweight adaptive task attention for few-shot fine-grained ship classification in remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-16, 2023.
- [302] Y. Guan, X. Zhang, G. Gao, and et al., "A new indicator for assessing fishing ecological pressure using multi-source data: A case study of the South China Sea," *Ecological Indicators*, vol. 170, Jan. 2025.
- [303] G. Gao, L. Yao, W. Li, and et al., "Onboard information fusion for multisatellite collaborative observation: Summary, challenges, and perspectives," *IEEE Geoscience and Remote Sensing Magazine*, vol. 11, no. 2, pp. 40-59, June 2023.
- [304] B. Shen, T. Liu, G. Gao, and et al., "A low-cost polarimetric radar system based on mechanical rotation and its signal processing," *IEEE Transactions on Aerospace and Electronic Systems*, doi: 10.1109/TAES.2024.3507776.
- [305] X. Zhang, G. Gao and S. -W. Chen, "Polarimetric autocorrelation matrix: A new tool for joint characterizing of target polarization and doppler scattering mechanism," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-22, 2024.