

Improving Deep Subdomain Adaptation by Dual-Branch Network Embedding Attention Module for SAR Ship Classification

Shuangmei Zhao and Haitao Lang , *Member, IEEE*

Abstract—This study aims at improving fine-grained ship classification performance under the condition that there is no labeled samples available in SAR domain (target domain) by transferring the knowledge from optical remote sensing (ORS) domain (source domain), which has rich labeled samples. The proposed method improves the original deep subdomain adaptation network (DSAN) by designing a dual-branch network (DBN) embedding attention module to extract more discriminative deep transferable features, thereby improving the performance of the subdomain adaptation. Specifically, we utilized a deep base network (ResNet-50) and a shallow base network (ResNet-18) to build the DBN, and embedded the convolutional block attention module after the first and the last convolutional layer of each branch. Extensive experiments demonstrate that the proposed method, which is termed as DSAN++, is feasible and achieves remarkable improvement than the state-of-the-art methods on the task of fine-grained ship classification.

Index Terms—Deep subdomain adaptation network (DSAN), domain adaptation (DA), dual-branch network (DBN), ship classification, synthetic aperture radar (SAR), transfer learning (TL).

I. INTRODUCTION

MARINE security has a great impact on economic development and the environment. As the most important carrier of human activities at sea, ships have always been the focus of maritime surveillance by coastal countries. Spaceborne synthetic aperture radar (SAR) has been widely used in most marine affairs because it can provide high-resolution, day-and-night, and weather-independent images [1]. Accurately identifying the category of ship in SAR images is very significant for many tasks, such as combating irregular immigration, conducting maritime rescue, and traffic monitoring.

With the increasing resolution of SAR, the classification of ship images becomes possible. Most of existing researches on ship classification in SAR images are based on supervised learning methods and have achieved great success [2], [3],

[4], [5], [6], [7], [8], [9], [10], but these approaches rely on a large amount of labeled SAR ship data. As we know, collecting sufficient training data is often expensive, time-consuming, or even unrealistic, which may hinder the further development of supervised learning methods in practical SAR ship classification scenarios. Recent years, transfer learning (TL), which aims to improve the performance of learner on the target domain by transferring the knowledge from the different but related source domain [11], [12], [13], has been introduced to solve the problem of ship classification under the condition that there are very few or even no samples that are not enough to train a good classifier in SAR domain. Lang et al. [14] used the automatic identification system (AIS) data as the source domain to extract naive geometric features (NGFs) of the ships, and designed a multiclass adaptive support vector machine as classifier to realize knowledge transfer between two domains. Xu et al. [15] proposed the method of discriminative adaptation regularization framework-based transfer learning (D-ARTL), which is an improvement to the original ARTL by adding a novel source discriminative information preservation regularization term to achieve a transfer from AIS domain to SAR domain. Xu et al. [16] proposed the method of geometric transfer metric learning (GTML), which improves the ship classification performance in SAR domain through joint application of TL and metric learning. Rostami et al. [17] proposed to transfer the knowledge of the high-resolution optical remote sensing (ORS) domain to the SAR domain to realize the classification of ships and nonships. In the feature extraction part, they designed two deep encoders, which are coupled to map data points into a common feature space, and then utilized the sliced wasserstein distance (SWD) [18] to measure and minimize the distribution discrepancy between the source and target domain. Song et al. [19] proposed to use CycleGAN [20] to transfer the labeled ORS images into SAR-style intermediate images with attribute labels first, and then use a domain adaptation (DA) network combining adversarial learning and metric learning to classify military-civilian ships. Yang et al. [21] proposed a dynamic joint correlation alignment network to achieve semisupervised heterogeneous TL from AIS domain to SAR domain. Lang et al. [22] proposed a multisource heterogeneous TL method for SAR ship classification. Analyzing the above methods, [14], [17], [19], [21], and [22] require the support of a small number of labeled samples in SAR domain, while [15] can work without

Manuscript received 24 May 2022; revised 21 July 2022 and 2 September 2022; accepted 11 September 2022. Date of publication 15 September 2022; date of current version 22 September 2022. This work was supported by the National Natural Science Foundation of China under Grant 62071030. (*Corresponding author: Haitao Lang.*)

The authors are with the College of Mathematics and Physics, Beijing University of Chemical Technology, Beijing 100029, China (e-mail: smzhao@buct.edu.cn; haitaolang@hotmail.com).

Digital Object Identifier 10.1109/JSTARS.2022.3206753

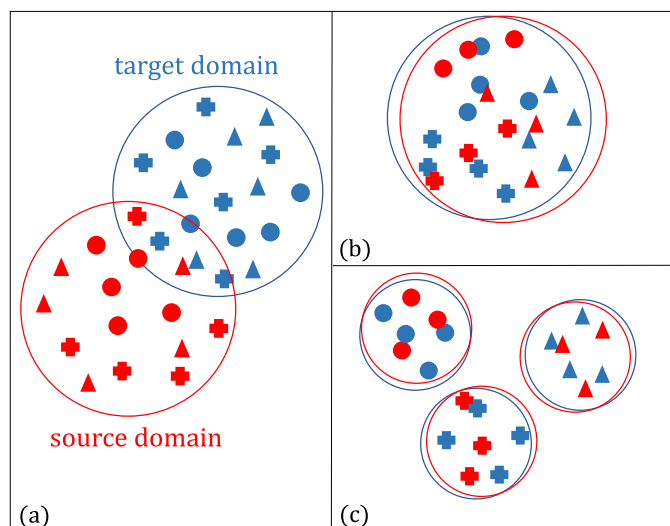


Fig. 1. Schematic diagram of DA. The triangle, circle, and cross symbols represent three subclasses (subdomains), whose colors in red and blue indicate that they belong to the source domain and target domain, respectively. (a) Before DA. (b) DA only with marginal distribution adaptation (MDA) and (c) DA with joint distribution adaptation, that is conducting both MDA and class-conditional distribution adaptation (CDA).

labeled samples in SAR domain, and Xu and Lang [16] can handle both labeled samples available and unavailable situations. In terms of source domain usage, AIS data is used as the source domain by [14], [15], [16] and [21]. The authors in [17] and [19] utilized ORS images. While Lang et al. [22] used both AIS and ORS data as source domains. In terms of TL solution, the work in [14] is a parameter transfer based method and the others are the methods of DA, which aims at learning a well-performing model from a source data distribution to a different target data distribution. To alleviate the domain shift of data distributions across the source and target domains, the authors in [17] and [19] aligned the class-conditional distribution between the two domains, the authors in [15], [16], and [21] aligned both the marginal distribution and class-conditional distribution.

Many existing DA methods demonstrate that if only marginal distributions are aligned, although the distributions of the two domains appear to be roughly aligned overall, different subdomains (subclasses) may also be mixed together, making it difficult to classify them correctly [23], [24], [25], [26]. This situation is illustrated in Fig. 1(a) and (b). While if taking the relationship between subdomains (subclasses) into account and aligning both marginal and class-conditional distribution (or say joint distribution), as shown in Fig. 1(a) and (c), the methods will achieve a better classification performance [15], [16], [21], [27], [28], [29], [30], [31], [32], [33].

Finding a correct distance metric and using the learned metric to fit a good classifier is very important in fine-grained classification tasks. Transfer metric learning methods, which combine TL and metric learning techniques, have been widely used in many applications [34]. For example, Deng et al. [35] proposed a deep metric learning feature embedding model suitable for unsupervised TL, and it can learn the similarity between sample pairs. Dong et al. [36] proposed joint distance transfer metric

learning to increase the interclass distance while reducing the intraclass distance on the basis of maximum mean discrepancy (MMD). The authors in [37] and [38] reduced the distance between source and target domains by minimizing empirical risk while maximizing the consistency of the manifold structure of the data with the classifier.

This study aims at improving fine-grained ship classification performance under the condition that there is no labeled samples available in SAR domain (target domain) by transferring the knowledge of ORS domain (source domain) with rich labeled samples under the framework of DA based on deep neural networks. The motivation of this study mainly stems from the following three aspects.

1) No labeled samples available in SAR domain is a common application scenario, which is more valuable and challenging for research, but there is very few reports at present [15], [16].

2) Compared with AIS data, ORS images can provide richer semantic feature, which is not limited to geometric information. Both ORS and SAR belong to image data, although their imaging mechanisms are different, it is assumed that there are some internal connections between them. Whether a ship is captured by a SAR sensor or an optical camera, the image data should share some common macroscopic features (such as geometric features) and some microscopic features (such as texture features), which can be extracted by the deep neural network and be utilized as transferable features. Based on this consensus, several existing works [17], [19], and [22] have also conducted studies on knowledge transfer from ORS to SAR, and the results prove that such transfer is not only feasible but also effective. Therefore, there is reason to believe that ORS can become a better source domain to assist fine-grained ship classification in SAR domain.

3) Recent studies have shown that deep neural networks can learn more transferable features for DA, which is achieved by embedding DA modules in the pipeline of deep feature learning to extract domain-invariant representations.

Specifically, this article proposed to improve the deep subdomain adaptation network (DSAN) [33] by utilizing a dual-branch network (DBN) embedding attention mechanism to enhance deep transferable feature extraction capability and realize subdomain adaptation from ORS domain (source) to SAR domain (target). DSAN is a newly proposed deep DA method, which can align both global (marginal) and local (class-conditional) distributions between two domains as well as learn transferable representations simultaneously, by integrating deep feature learning network (DFLN) and subdomain adaptation network (SDAN) into an end-to-end deep learning mode [see Fig. 2(a)]. In this study, we improve the DFLN of the original DSAN by using a dual-branch architecture, and embed the attention mechanism into each branch [see Fig. 2(b)–(c)]. By fully mining the image information of two domains, it can extract more discriminative deep transferable features, which further boosts the performance of subsequent subdomain adaptation process. In this sense, we refer to the proposed method as DSAN++. In-depth analysis and extensive experiments show that in the common subspace mapped via the deep feature extracted by the DSAN++, the marginal distribution shift between the target and source domain

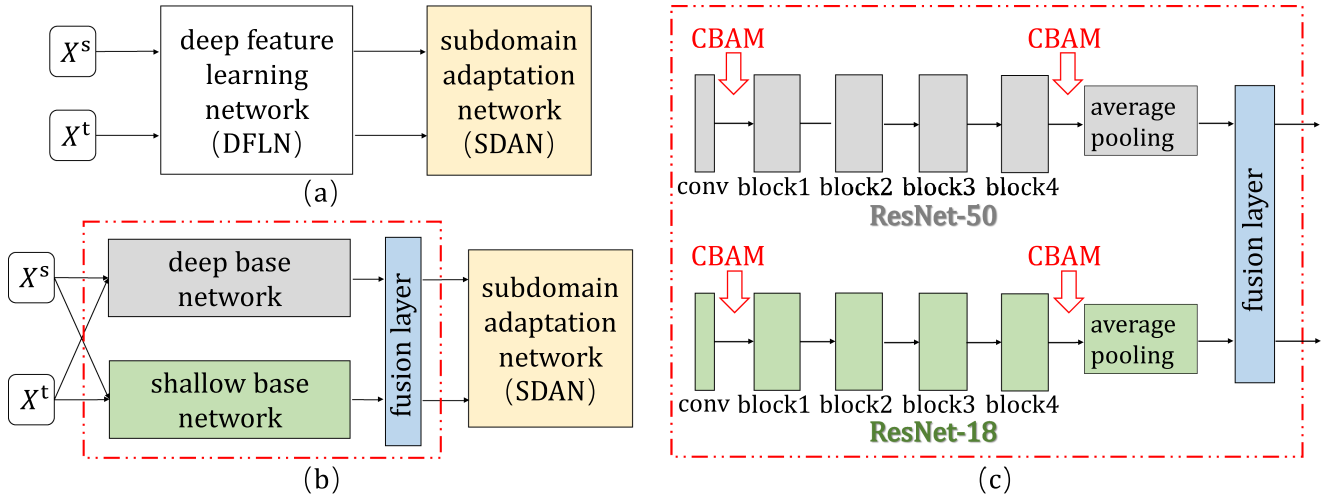


Fig. 2. Architecture of the proposed DSAN++. (a) Original DSAN is composed of a DFLN and a SDAN. (b) The proposed DSAN++ uses a DBN embedding attention module to excute deep feature extraction. (c) ResNet-50 and ResNet-18 are used as the deep and shallow network, CBAM are embedded into each branch of the DBN.

and the class-conditional distribution discrepancy cross each subdomain are minimized (compared with original DSAN and other comparison methods).

The main contribution of this article is three-fold.

1) This study proposes DSAN++ which improves the original DSAN by designing a DBN embedding attention mechanism to extract the deep feature. This change allows the network to extract more discriminative transferable features, further improving the performance of the subdomain adaptation.

2) This is the first work to solve the fine-grained ship classification problem in SAR domain by transferring the knowledge from ORS domain as source domain, focusing on the application scenario that there is no labeled samples available. In contrast, although the authors in [17] and [19] studied the TL from ORS domain to SAR domain, Rostami et al. [17] only realized the classification between ships and nonships, and Song et al. [19] studied the classification of military and civilian ships, neither of which is a more complex fine-grained ship classification. The application scenario of [22] is supervised, which is different from us.

3) Extensive experiments demonstrate the effectiveness and reliability of the proposed method.

The rest of this article is as follows. Section II introduces the proposed method in detail, including the overall framework, DBN, and the method of CBAM embedding. Section III describes the datasets and experimental protocol. Next, we report and analyze the experimental results in Section IV. Finally, Section V concludes this article.

II. PROPOSED METHOD

A. Overall Framework

DSAN can learn a transfer network by aligning the relevant subdomain distributions (i.e., class-conditional distribution) of multiple domain-specific layers across source and target domains on the basis of local maximum mean discrepancy (LMMMD) [33]. As shown in Fig. 2(a), it is composed of a DFLN

followed by an SDAN. Experiments have shown that by embedding DA modules in the pipeline of deep feature learning helps us to extract domain-invariant feature representations. More importantly, subdomain adaptation has the ability to capture the fine-grained information for each category, thereby improving fine-grained classification performance.

As reported by previous studies [7], [16], fine-grained ship classification in SAR images is not a trivial task since ships only have subtle visual appearance variation between different categories. Therefore, it is particularly important to extract features that are sufficiently discriminative. As shown in Fig. 2(b), for the purpose of fully exploring the feature representation potential of ORS and SAR images, this article improves the DFLN of the original DSAN with a DBN, which is composed of a deep base network, a shallow base network, and a fusion layer. This improvement is inspired by [39], which demonstrates that the deeply fused network is able to learn multiscale feature representations due to the complementary contribution of multiple base networks with different depths. Another improvement is shown in Fig. 2(c), this study embeds the convolutional block attention module (CBAM) [40] after the first and the last convolutional layer of each branch to further improve feature representation power.

B. Dual-Branch Network

Following the suggestion of previous work [39], [41], the proposed DBN is formed by one deep (ResNet-50) and one shallow (ResNet-18) base networks as shown in Fig. 2(b) and (c). The features extracted from ResNet-50 and ResNet-18 are different. Comparatively speaking, the features extracted by ResNet-18 have a higher resolution and contain more location and details information, but lower semantics information and higher noise. The high-level features extracted by ResNet-50 have stronger semantic information, but the resolution is very low, and the perception of details is poor. In this article, we try to efficiently integrate the two and take advantage of their strengths

to improve the feature representation capability. Another reason we choose ResNet-50 and ResNet-18 as the backbone network is that the residual structure is implemented in the form of shortcut connection, which solves the degradation problem of deep neural networks. It avoids the loss of information to a certain extent when transmitting information, and protects the integrity of the information. The features of the input ship images are extracted by two branch networks, respectively, then are fused in the fusion layer, where the concatenation operation is applied to connect the two features.

C. CBAM Embedding

We assume that the transferability of different semantic contents in an image is different and DA methods need to focus on the meaningful knowledge that are highly relevant to the task while ignoring the irrelevant information. Based on this assumption, we propose to embed attention mechanism into the feature learning process to exploit the effective semantic features. The essence of the attention mechanism is to reweight the feature maps for adaptive feature optimization, so that the important parts of the image will be given higher weights, and the unimportant parts will be given lower weights, which will enhance the feature representations. There are many well-performing attention modules that can be used to realize our concept [42], such as squeeze-and-excitation (SE) networks, efficient channel attention network (ECA-Net), and CBAM [40]. Considering that CBAM can both center on “what” is meaningful for a input image with its channel attention module and concentrate on “where” is the informative part in the input image with its spatial attention module, which is more suitable to realize our conception, this study utilizes CBAM. As shown in Fig. 2(c), we embed CBAM after the first convolutional layer and the last convolutional layer for the purpose of not changing the network structure of ResNet so that we can use pretraining parameters instead of training from scratch.

D. Loss

Through the end-to-end training of deep neural networks, we aim to minimize the distribution discrepancy between the related subdomains activated in the domain-specific layers $L = \{1, 2, 3, \dots\}$. The total loss is

$$\min_f \frac{1}{n_s} \sum_{i=1}^{n_s} J(f(x_i^s), y_i^s) + \lambda \sum_{l \in L} \hat{d}_l(p, q) \quad (1)$$

where the first item is the cross-entropy loss defined as

$$J(f(x_i^s), y_i^s) = -[y_i^s \log f(x_i^s) + (1 - y_i^s) \log(1 - f(x_i^s))]. \quad (2)$$

The cross-entropy loss is used to calculate the classification error, which needs to be minimized on source domain, where n_s is the number of samples in the source domain, y_i^s and $f(x_i^s)$ represent the ground-truth and predicted label of the sample x^s in the source domain, respectively.

The second item in (1) is the DA loss defined as

$$\begin{aligned} \hat{d}_l(p, q) = \frac{1}{C} \sum_{c=1}^C \left[\sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \omega_i^{sc} \omega_j^{sc} \kappa(z_i^{sl}, z_j^{sl}) \right. \\ \left. + \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \omega_i^{tc} \omega_j^{tc} \kappa(z_i^{tl}, z_j^{tl}) \right. \\ \left. - 2 \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} \omega_i^{sc} \omega_j^{tc} \kappa(z_i^{sl}, z_j^{tl}) \right] \quad (3) \end{aligned}$$

where p and q are the distributions of source domain and target domain. $\hat{d}_l(p, q)$ is the unbiased estimator of LMMD following the definition in [33] and $l \in L$. C is the number of classes of ships. n_t is the number of samples in the target domain, ω_i^{sc} and ω_j^{tc} represent the weight of each sample in the source and target domain. And z_i^{sl} and z_j^{tl} denote the activation features generated by the network at layer l , respectively. $\lambda > 0$ is a tradeoff parameter between the classification loss and DA loss.

III. EXPERIMENT

A. Datasets

1) *Target domain*: There are two SAR ship datasets for this research. The first is GF-SAR dataset, which consists of 150 Gaofen-3 images (3 ship classes, namely bulk carrier, container ship, and oil tanker, 50 images per class), of which 88 are from the FUSAR-Ship dataset [43] with about 1.0 m azimuth resolution and about 1.7 m slant range resolution, and the other 62 are collected by the authors with about 0.5 m azimuth resolution and 0.3 m range resolution. All class labels have been matched by AIS information. The second is HR-SAR dataset, which was collected by Xing et al. [2] from six TerraSAR-X stripmap-mode SAR imagery with 2.0 m azimuth resolution and 1.5 m range resolution. This dataset also contains three classes of ships, including cargo, container ship, and oil tanker, 50 images per class. Some ship samples of above two SAR datasets can be seen in Fig. 3(b) and (c), respectively.

2) *Source domain*: For the research purpose, we specially collected ship images from Google Earth with submeter resolution to build the ORS dataset. This dataset includes four classes of cargo, bulk carrier, container ship and oil tanker, with 1000 images per class, whose information are matched with the official website.¹ Unlike some existing ORS ship datasets [17], whose single image slice may contain multiple ships or even incomplete ship. As shown in Fig. 3(a), the vast majority of image slices in our dataset are segmented more carefully, and each of them only contains a single ship that is more suitable for ship classification research. The ORS dataset is available at <https://github.com/BUCT-RS-ML/MS-HeTL-via-MS-HFA>.

In the experiments of this study, for two target domain datasets GF-SAR and HR-SAR, we select the corresponding ship classes from source domain ORS dataset to conduct knowledge transfer task, respectively.

¹[Online] Available: <http://www.marinetraffic.com/>.

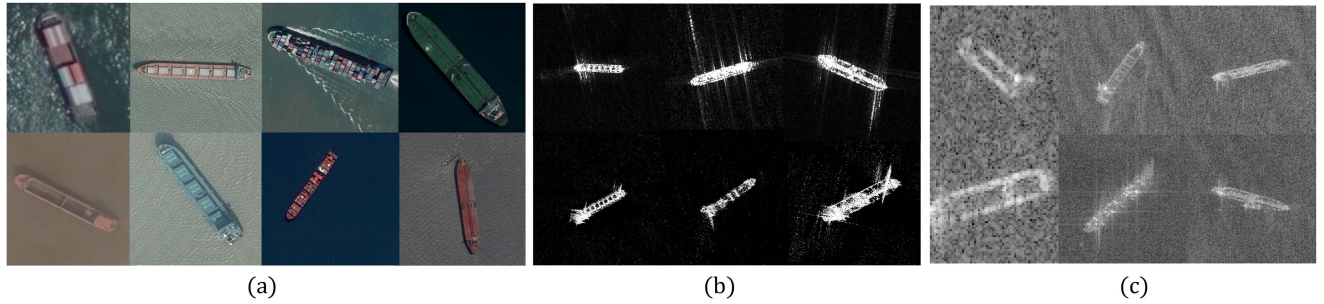


Fig. 3. Ship samples in the source and target domains. (a) ORS dataset, from left to right are cargos, bulk carriers, container ships, and oil tankers. (b) GF-SAR dataset, including three categories of bulk carriers, container ships and oil tankers. (c) HR-SAR dataset, including three categories of cargos, container ships and oil tankers.

TABLE I
METHODS OF COMPARISON

Category	Methods	Description
Traditional	D-ARTL [15]	D-ARTL aligns the marginal distribution and class-conditional distribution of source and target domains on the SAR ship classification task. It adds a source domain discriminative information preservation regularization term on the basis of ARTL [45], and transfers the knowledge learned in the source domain to the target domain [15].
	GTML [16]	GTML aligns the marginal distribution and class-conditional distribution of source and target domains on the SAR ship fine-grained classification task by combining transfer learning and metric learning [16].
Deep learning	DANN [25]	DANN aligns the marginal distribution and realizes the extraction of domain invariant features, which consists of three components: feature extractor, class predictor, and domain discriminator. This is a pioneering work combining domain adaptation with adversarial learning [25].
	DAAN [44]	DAAN aligns the marginal distribution and class-conditional distribution between source and target domains by exploiting dynamic adversarial factor to dynamically learn domain-invariant feature representations, and using the factor to evaluate the relative importance of the two distributions [44].
	DAN [23]	DAN aligns the marginal distribution between source and target domains, which is proposed based on DDC [46]. DAN adapts the last three fully connected layers of the network, and expands the MMD into a multi-kernel MMD [23].
	Deep-CORAL [24]	Deep-CORAL aligns the marginal distribution between source and target domains, which applies the CORAL [47] to the convolutional neural network. It uses a nonlinear transformation to minimize CORAL loss between the two domains. [24]
	JAN [27]	JAN aligns the joint distribution (the marginal and class-conditional distributions) between source and target domains. It extends MMD to JMMD [27].
	MRAN [31]	MRAN aligns the class-conditional distribution between source and target domains through conditional-MMD (CMMD). Unlike the previous convolutional neural networks that only extract features from the source and target domain data by a single network structure, MRAN extracts multi-representation features based on inception attention module (IAM) [31].
DSAN [33]	DSAN aligns the marginal distribution and class-conditional distribution between source and target domains with LMMD while considering the weights of different samples [33].	

B. Experimental Protocol

1) *Evaluate criteria*: Following the setting of [13] and [16], we adopted accuracy as the criteria to evaluate the SAR ship classification performance, which is defined as

$$\text{accuracy} = \frac{|\{x|x_j^t \in D_t \wedge f(x_j^t) = y_j\}|}{|D_t|} \quad (4)$$

where D_t denotes the test data, $f(x_j^t)$ and y_j denote the predicted label and the ground-truth label of sample x_j^t , respectively. And the operator $|\cdot|$ represents the number of elements in the set.

2) *Comparison methods*: In this study, we compared the proposed method with nine state-of-the-art TL methods, including DAN [23], Deep-CORAL [24], DANN [25], JAN [27], DAAN [44], MRAN [31], D-ARTL [15], GTML [16], and DSAN [33], which conducted same TL task on the same datasets. The specifics of these comparison methods can be seen in Table I.

3) *Parameter setting*: For DSAN and the proposed DSAN++, we used minibatch stochastic gradient descent optimizer with a momentum of 0.9, and the batch size is set to 16. The learning rate follows the formula: $\eta_\theta = \eta_0 / (1 + \alpha\theta)^\beta$, where $\eta_0 = 0.01$, $\alpha = 10$, $\beta = 0.75$, and θ is changing from 0 to 1, which is optimized to promote convergence and reduce error on source domain [27], and the settings of these parameters are the same as those in [33]. In order to suppress noisy activations in the early stages of training, λ is not fixed. Instead, λ follows a progressive schedule: $\lambda_\theta = 2 / \exp(-\gamma\theta) - 1$, where $\gamma = 10$ is defined by the experiments [25]. The progressive schedule can not only stabilize parameter sensitivity but also ease model selection for DSAN++. The epoch is set to 200, and the dropout rate is 0.5 to avoid feature redundancy. For the rest methods, the parameters are set to the default values or the recommended values mentioned in their original articles. For fair comparison, ResNet-50 is selected as the backbone network of all deep learning-based methods. The pretrained models trained on the

TABLE II
RESULTS OF THE ABLATION STUDY ON GF-SAR DATASET

Methods	Accuracy(%)	Distribution discrepancy(LMMD)	Distribution discrepancy(MMD)
DSAN	82.67	5.58	18.17
DSAN + DBN	84.67	4.51	16.03
DSAN + CBAM	86.00	3.92	13.25
DSAN++	89.11	2.29	9.06

TABLE III
RESULTS OF THE ABLATION STUDY ON HR-SAR DATASET

Methods	Accuracy(%)	Distribution discrepancy(LMMD)	Distribution discrepancy(MMD)
DSAN	84.00	4.90	16.63
DSAN + DBN	85.11	4.21	15.03
DSAN + CBAM	86.67	3.77	12.92
DSAN++	88.89	2.38	9.24

ImageNet were used as the initialization model. For traditional methods, we follow [15] and [16] to extract NGFs from source and target domains for the transfer task.

The experiments were performed with Intel i9-9980XE CPU 3.00 GHz and GeForce RTX 2080 Ti. Each method was implemented three times, and the average accuracy were adopted as the experimental results.

IV. RESULTS AND ANALYSIS

A. Effectiveness of DBN and CBAM Embedding

In order to validate the effectiveness of the proposed two improvements, i.e., DBN and CBAM embedding, we conducted the following three comparison experiments on two target domain datasets, respectively:

- 1) Only replacing the original DFLN with the DBN without attention module embedding to obtain DSAN + DBN;
- 2) Only embedding CBAM into the original DFLN to obtain DSAN + CBAM;
- 3) Replacing the original DFLN with DBN embedding CBAM to obtain the proposed DSAN++.

From the classification performance listed in Tables II and III, it is found that both DBN and CBAM embedding can effectively improve the performance of DSAN, which implies that: 1) Compared with the original DFLN, which is a single branch network, the proposed DBN has a stronger deep discriminative feature extraction ability and 2) CBAM embedding can further improve the deep feature extraction capability of the network. The proposed method combines the two strategies, fully exploiting their specialties in deep feature extraction, and greatly improves the performance of the original DSAN by 6.44% (from 82.67% to 89.11%) on GF-SAR dataset, and 4.89% (from 84.00% to 88.89%) on HR-SAR dataset. Additionally, we use LMMD to measure the discrepancy in local subdomain feature distributions, and MMD to measure the global distribution discrepancy between the source and target domain. The smaller the two values, the smaller the discrepancy in the distribution of features within the local subdomain and between

the two domains. From Tables II and III, we can find that the proposed DSAN++ achieves the lowest LMMD and MMD on both SAR datasets, that is, DSAN++ achieves the closest feature distribution of the corresponding ship classes in the ORS images and the SAR images, thereby obtaining the highest the ship classification accuracy. These results illustrate that with the increase of the discriminative power of the deep feature, the distribution discrepancies measured by LMMD and MMD are gradually reduced.

Taking the classification results on the GF-SAR dataset as an example, as shown in Fig. 4, we randomly select a part of samples from the source domain and target domain, and use t-SNE technique [48] to visualize the feature distribution of the four methods of DSAN, DSAN+DBN, DSAN+CBAM, and DSAN++ in the common feature space. In Fig. 4, different colors are used to distinguish source and target domains, and different markers indicate different ship classes. Since the deep features extracted by the deep neural network have extremely high dimensions and are difficult to be visualized, t-SNE maps those high-dimensional features into two-dimensional space through dimensionality reduction so as to visually display their distribution. t-SNE pays more attention to preserving the distribution relationship of the original data: in this two-dimensional space, the data that was originally close in distance would also be close after dimensionality reduction; similarly, the distance that was originally far away would be far after dimensionality reduction. Based on the abovementioned cognition, we can find that the original DSAN [see Fig. 4(a)] uses ResNet-50 to extract the deep features of ORS ship images and SAR ship images, and then uses SDAN to align the class-conditional distribution and marginal distribution between the two domains. Compared to the other three methods, the distances between subclass samples are tighter. By learning and extracting more discriminative deep features, DSAN+DBN and DSAN+CBAM [see Fig. 4(b) and (c)] can further separate samples between subclasses. While the proposed method DSAN++ [see Fig. 4(d)] obtains transferable features with more significant class discrimination ability, and makes the intraclass samples more closely clustered and widens the distance between the subclass samples, which is

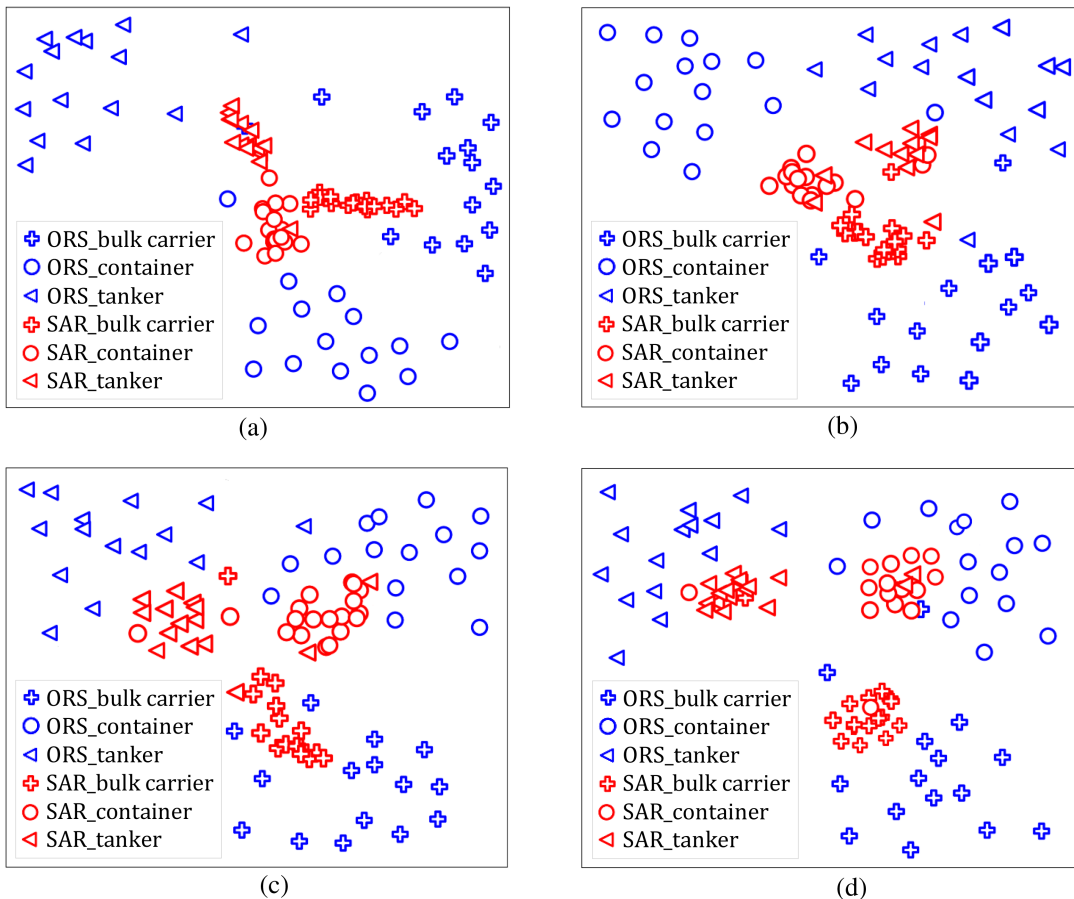


Fig. 4. Visualization of feature distribution in the common feature space on GF-SAR dataset. (a) DSAN. (b) DSAN + DBN. (c) DSAN + CBAM. (d) DSAN++.

more conducive to the transfer task, thereby improving the fine-grained classification performance.

B. Various Attention Mechanism Modules Embedding

In our study, we proposed to embed CBAM attention mechanism module into the feature learning process to exploit the effective semantic features. Considering that there are also various attention mechanism modules possessing similar effects, for the purpose of helping researchers to have a more comprehensive and in-depth understanding of the proposed architecture as shown in Fig. 2, in this section, we conducted an experiments with the other two well-performing attention mechanism modules: SE [49] and efficient channel attention (ECA) [50], which were embedded into the proposed architecture using the same strategy, and reported and compared the performance with the CBAM adopted in this article. Both SE and ECA module are channel attention mechanisms. The difference is that SE reweights each channel of the feature map through the fully connection operation, while ECA pays attention to the relationship between neighbor channels of the feature map, and it turns the two fully connected operations in the channel attention mechanism into a one-dimensional convolution to reassign the weight. SE module and ECA module are embedded in DBN, respectively (added to each residual block), and the classification

TABLE IV
PERFORMANCE WITH DIFFERENT ATTENTION MECHANISMS EMBEDDING
EVALUATED ON ACCURACY (%)

Different attention module	GF-SAR	HR-SAR
DBN + SE	86.67	85.33
DBN + ECA	88.00	87.55
DBN + CBAM	89.11	88.89

results of different attention mechanism embeddings are listed in the Table IV.

It can be found that embedding these attention mechanisms all can increase the classification accuracy. But relatively, the CBAM module performs better than the SE module and the ECA module embedding on the two target datasets. The reason for this phenomenon may be that CBAM considers the channel relationship and spatial relationship of feature maps, while SE and ECA modules only consider the relationship between channels of feature maps. This experiment demonstrates the effectiveness of CBAM module embedding in our proposed method. Observing the classification confusion matrices on GF-SAR dataset in Fig. 5, we can find that embedding CBAM module can handle ship subclass classification better. Compared with embedding SE module, embedding CBAM has higher classification accuracy in each subclass. Although the performance of embedding

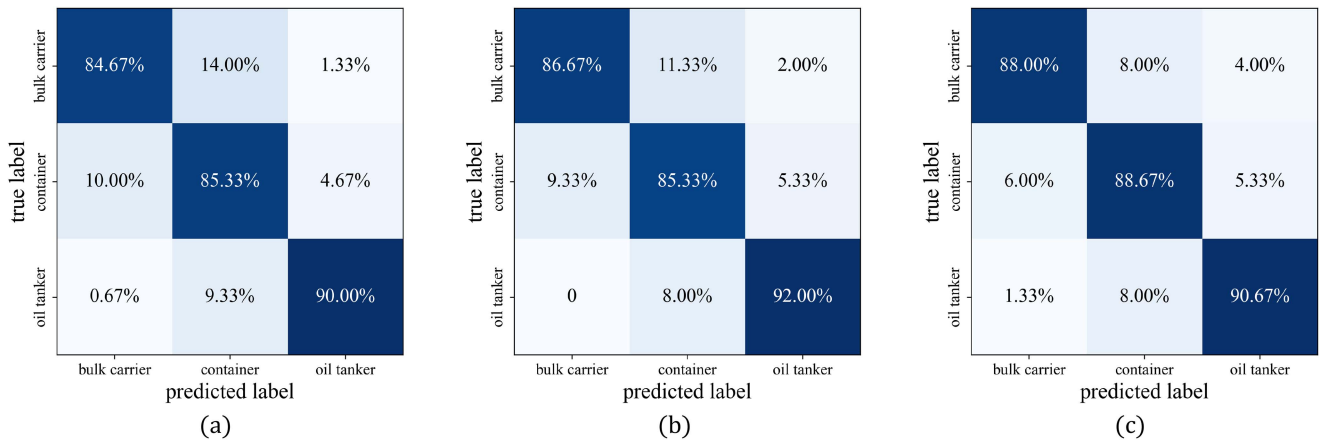


Fig. 5. Classification confusion matrix after three attention mechanisms are embedded into the DBN on GF-SAR dataset. (a) DSAN + SE. (b) DSAN + ECA. (c) DSAN + CBAM.

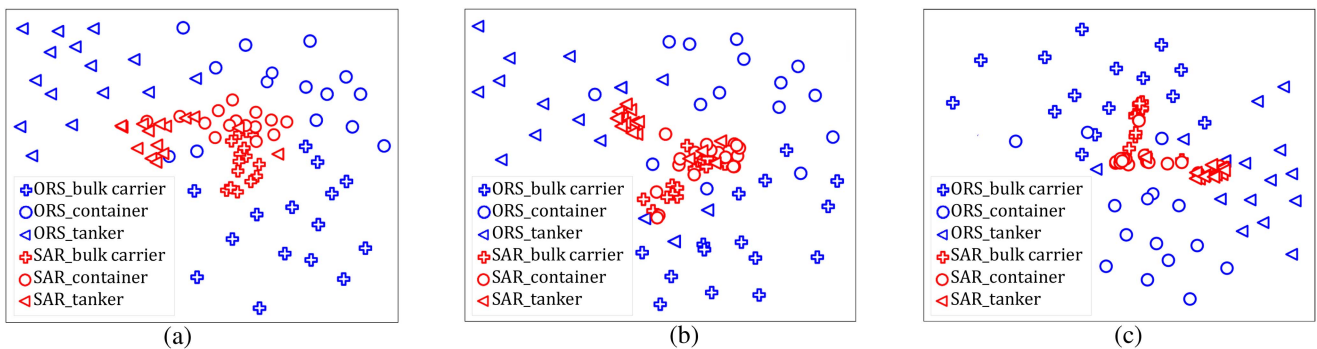


Fig. 6. Visualization of feature distribution in the common feature space of comparison methods to DA on GF-SAR dataset. (a) DAN. (b) DANN. (c) MRAN.

CBAM is slightly lower than embedding ECA on the “oil tanker” subclass (90.67% versus 92.00%), embedding CBAM achieves higher classification accuracy on the other two subclasses.

C. Comparison With State-of-the-Art

In the last experiment, we compared the proposed DSAN++ with the state-of-the-art methods as shown in Table I, which can be roughly divided into the following three categories, where D-ARTL [15] and GTML [16] are traditional methods (non-deep learning based methods), DANN [25] and DAAN [44] are adversarial-based methods, and the other methods together with the proposed method belong to statistic moment matching-based methods. Specifically, for GF-SAR dataset, the experimental results in Table V show that most of deep neural network based DA methods (the only exception is Deep-CORAL [24]) outperform two traditional methods D-ARTL (59.67%) and GTML (66.00%), which illustrates the importance of more discriminative deep feature. The poor performance of Deep-CORAL (65.33%) is mainly due to the fact that it only aligns the global domain shift. The similar situation happens to DAN [23] (67.33%), which just aligns the global distribution. Two adversarial-based methods DAAN (67.33%) and DANN (71.33%) are slightly better than the previous methods but not as good as the other statistic moment matching-based methods.

TABLE V
SAR SHIP CLASSIFICATION PERFORMANCE OF DIFFERENT ALGORITHMS

Methods	GF-SAR		HR-SAR	
	accuracy(%)	run time(s)	accuracy(%)	run time(s)
D-ARTL	59.67	3.945	71.33	3.811
GTML	66.00	0.496	68.67	0.454
DANN	71.33	0.096	76.67	0.095
DAAN	67.33	0.084	78.67	0.081
DAN	67.33	0.045	72.00	0.045
Deep-CORAL	65.33	0.079	74.67	0.079
JAN	72.00	0.062	79.33	0.062
MRAN	81.00	0.080	84.00	0.079
DSAN	82.67	0.067	84.00	0.065
DSAN++	89.11	0.069	88.89	0.069

JAN [27] adapts the joint distribution difference of multiple layers with joint-MMD (JMMD) and achieves 72.00% accuracy. MRAN [31] improves the accuracy to 81.00% by extracting multiple feature representations from a single perspective using an inception attention module (IAM) and minimizing CMMD. Thanks to utilizing LMMD, which can measure the distribution of related subdomains with considering the weight of each sample in both source and target domains, DSAN [33] captured more fine-grained information and improved classification performance to 82.67%. As mentioned before, with the exact

same SDAN as DSAN, the proposed DSAN++ explodes the performance to 89.11% (which outperforms DSAN by 6.44%) through improving the DFLN. Similar results appeared on another HR-SAR dataset, which are listed in Table V. The traditional methods D-ARTL and GTML obtain 71.33% and 68.67% classification accuracies, respectively, which are slightly lower than those based on deep learning. It also can be observed that the ship classification performance of those DA methods that only align marginal distributions, DAN (72.00%), DANN (76.67%), and Deep-CORAL (74.67%) are slightly lower. While those methods that align class-conditional distributions, JAN (79.33%), DAAN (78.67%), MRAN (84.00%), and DSAN (84.00%) achieve relatively good classification accuracy. On the basis of DSAN, the proposed method DSAN++ gets the best classification accuracy of 88.89%. In terms of running time, the proposed DSAN++ is comparable to JAN, DSAN, only slightly slower than DAN, and faster than other methods, especially traditional methods.

Different unsupervised DA methods explore the transferable features in different ways. In order to provide a more intuitive comparison and reveal the reasons behind the classification performance, we also utilize t-SNE to visualize the feature distribution in the common feature space of GF-SAR dataset and corresponding ORS ship dataset for several typical comparison methods, DAN, DANN, and MRAN, as shown in Fig. 6, where the meanings of different colors and markers are the same as those defined in Fig. 4, and to compare with DSAN and the proposed DSAN++, as shown in Fig. 4(a) and (d). It can be seen that the data points in both domains are not distinguishable so well in these three subplots. The performance of the target domain will be affected by the source domain. DAN and DANN [see Fig. 6(a) and (b)] methods only achieve the marginal distribution alignment of the extracted features of source and target domains. None of them handle class-conditional distributions of features well, resulting in the confounding of samples from various subclasses (especially in SAR domain), which hinders the improvement of the classification performance. MRAN [see Fig. 6(c)] utilizes CMMD to realize the class-conditional distribution alignment between the features of source and target domain, which relatively reduces the domain shift and greatly improves the classification accuracy of SAR ship images. That is, by aligning the class-conditional distribution of ship features extracted in the ORS domain and SAR domain, the transferable features will be class-discriminative, which is more beneficial to the fine-grained classification task of SAR ship images. When comparing with Fig. 4(a) and (d), it shows that DSAN and DSAN++ matches both global and local distributions through subdomain adaptation, thus greatly improving the classification performance.

V. CONCLUSION

This article proposes DSAN++ which improves the original DSAN [33] by designing a DBN embedding CBAM attention mechanism to extract more discriminative deep transferable features, thereby improving the performance of the subdomain adaptation. Extensive experiments on two SAR datasets

TABLE VI
FINE-GRAINED CLASSIFICATION PERFORMANCE OF SAR SHIPS WITH
DIFFERENT NUMBER OF SUBCLASSES

Methods	GF-SAR-4		GF-SAR-5	
	accuracy(%)	run time(s)	accuracy(%)	run time(s)
DSAN	77.17	0.067	66.00	0.071
DSAN++	81.83	0.070	69.60	0.073

demonstrate that the proposed DSAN++ is feasible and achieves remarkable improvement than the state-of-the-art methods.

APPENDIX

In the main part of this article, we have demonstrated the superiority of the proposed method from various aspects. In this section, we try to further illustrate that the proposed DSAN++ is also capable of handling more complex fine-grained classification tasks, i.e., subdividing more subclasses. We hope this additional study is useful for interested readers. To conduct this study, we further expand the GF-SAR dataset (target domain) and the ORS dataset (source domain) to include more subclasses. With no public dataset available for direct use, dataset expansion (especially for SAR dataset) is not a trivial task. By bringing together the samples from FUSAR-Ship dataset [43] and our self-collected, we add two new subclasses (50 samples per category), i.e., cargo and fishing ship, to the original GF-SAR dataset, which is introduced in Section III-A. The ORS dataset is also expanded accordingly by adding the fishing ship category to the original ORS dataset as introduced in Section III-A. The appended fishing ship samples come from two datasets, FGSC-23 [51] and ShipRSImageNet [52] with 102 and 171 samples, respectively.

Next, we evaluate the performance of the proposed DSAN++ in two multiclass SAR ship classification tasks and compare the results with those of DSAN (the best state-of-the-art method as described in Section IV-C) on the same tasks. The first task is to categorize four subclasses, which is conducted on GF-SAR-4 dataset (i.e., GF-SAR + cargo subclass), and the second for five subclasses classification conducted on GF-SAR-5 dataset (i.e., GF-SAR + both cargo and fishing ship subclasses). The comparison results and detailed confusion matrices are shown in Table VI and Fig. 7 in this Appendix.

As can be seen from Table V in the main part and Table VI in the Appendix, with the number of categories increases, the overall classification accuracy decreases gradually. For DSAN++, it is from 89.11% (three subclasses) to 81.83% (four subclasses) then to 69.60% (five subclasses). This result and trend are consistent with theoretical cognition. We also notice that the proposed DSAN++ performs better on any task than DSAN, which is the best state-of-the-art method, leading 6.44%, 4.66%, and 3.60%, respectively. It is also reasonable that this gap decreases as the number of categories increases. In-depth analysis of the confusion matrix (see Table VI), it can be found that DSAN++ and DSAN are slightly different in their ability to handle different subclasses on a specific task. DSAN performs better than DSAN++ for “cargo” subclass (85.33% versus

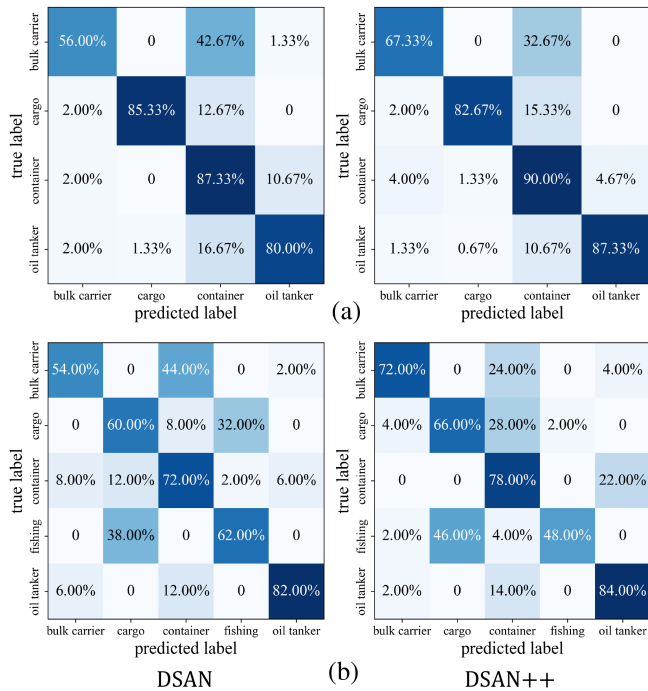


Fig. 7. Confusion matrix for (a) four-class and (b) five-class ship classification. Left column for DSAN and right column for DSAN++.

82.67%) on GF-SAR-4 dataset, and for “fishing ship” subclass (62.00% versus 48.00%) on GF-SAR-5 dataset, respectively. At the same time, because DSAN++ far outperforms DSAN in other more subclasses, making its overall performance better than that of the latter. These experiments demonstrate the effectiveness of the proposed DSAN++ method in handling the problem of fine-grained classification of multiclass ships.

ACKNOWLEDGMENT

The authors would like to thank the contributors to the FUSAR-Ship dataset, HR-SAR dataset, FGSC-23 dataset, and ShipRSImageNet dataset for providing data support for this research. And we also would like to thank all peer reviewers and associate editors for their constructive comments that significantly improve this article.

REFERENCES

- [1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, “A tutorial on synthetic aperture radar,” *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 1, pp. 6–43, Mar. 2013.
- [2] X. Xing, K. Ji, H. Zou, W. Chen, and J. Sun, “Ship classification in terraSAR-X images with feature space based sparse representation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1562–1566, Nov. 2013.
- [3] H. Lang, J. Zhang, X. Zhang, and J. Meng, “Ship classification in SAR image by joint feature and classifier selection,” *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 212–216, Feb. 2016.
- [4] C. Bentes, D. Velotto, and B. Tings, “Ship classification in terraSAR-X images with convolutional neural networks,” *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 258–266, Jan. 2018.
- [5] Y. Dong, H. Zhang, C. Wang, and Y. Wang, “Fine-grained ship classification based on deep residual learning for high-resolution SAR images,” *Remote Sens. Lett.*, vol. 10, no. 11, pp. 1095–1104, 2019.

- [6] G. Huang, X. Liu, J. Hui, Z. Wang, and Z. Zhang, “A novel group squeeze excitation sparsely connected convolutional networks for SAR target classification,” *Int. J. Remote Sens.*, vol. 40, no. 11, pp. 4346–4360, 2019.
- [7] Y. Xu and H. Lang, “Distribution shift metric learning for fine-grained ship classification in SAR images,” *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2276–2285, 2020.
- [8] J. He, Y. Wang, and H. Liu, “Ship classification in medium-resolution SAR images via densely connected triplet CNNs integrating fisher discrimination regularized metric learning,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3022–3039, Apr. 2021.
- [9] L. Zeng et al., “Dual-polarized SAR ship grained classification based on CNN with hybrid channel feature loss,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4011905.
- [10] J. Ai, Y. Mao, Q. Luo, L. Jia, and M. Xing, “SAR target classification using the multikernel-size feature fusion-based convolutional neural network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5214313.
- [11] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [12] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J. Big Data*, vol. 3, no. 1, pp. 1–40, 2016.
- [13] F. Zhuang et al., “A comprehensive survey on transfer learning,” *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021.
- [14] H. Lang, S. Wu, and Y. Xu, “Ship classification in SAR images improved by AIS knowledge transfer,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 439–443, Mar. 2018.
- [15] Y. Xu, H. Lang, L. Niu, and C. Ge, “Discriminative adaptation regularization framework-based transfer learning for ship classification in SAR images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 11, pp. 1786–1790, Nov. 2019.
- [16] Y. Xu and H. Lang, “Ship classification in SAR images with geometric transfer metric learning,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6799–6813, Aug. 2021.
- [17] M. Rostami, S. Kolouri, E. Eaton, and K. Kim, “Deep transfer learning for few-shot SAR image classification,” *Remote Sens.*, vol. 11, no. 11, 2019, Art. no. 1374.
- [18] J. Rabin, G. Peyré, J. Delon, and M. Bernot, “Wasserstein barycenter and its application to texture mixing,” in *Proc. Int. Conf. Scale Space Variational Methods Comput. Vis.*, 2011, pp. 435–446.
- [19] Y. Song, J. Li, P. Gao, L. Li, T. Tian, and J. Tian, “Two-stage cross-modality transfer learning method for military-civilian SAR ship recognition,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4506405.
- [20] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2242–2251.
- [21] G. Yang and H. Lang, “Semi-supervised heterogeneous domain adaptation via dynamic joint correlation alignment network for ship classification in SAR imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4508105.
- [22] H. Lang, G. Yang, C. Li, and J. Xu, “Multisource heterogeneous transfer learning via feature augmentation for ship classification in SAR imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5228814.
- [23] M. Long, Y. Cao, J. Wang, and M. Jordan, “Learning transferable features with deep adaptation networks,” in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 97–105.
- [24] B. Sun and K. Saenko, “Deep coral: Correlation alignment for deep domain adaptation,” in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 443–450.
- [25] Y. Ganin et al., “Domain-adversarial training of neural networks,” *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2096–2030, 2016.
- [26] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, “Adversarial discriminative domain adaptation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7167–7176.
- [27] M. Long, H. Zhu, J. Wang, and M. I. Jordan, “Deep transfer learning with joint adaptation networks,” in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2208–2217.
- [28] M. Long, Z. Cao, J. Wang, and M. I. Jordan, “Conditional adversarial domain adaptation,” *Adv. Neural Inf. Process. Syst.*, vol. 31, pp. 1640–1650, 2018.
- [29] S. Li, S. Song, G. Huang, Z. Ding, and C. Wu, “Domain invariant and class discriminative feature learning for visual domain adaptation,” *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4260–4273, Sep. 2018.

- [30] C. Chen, Z. Chen, B. Jiang, and X. Jin, "Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 01, 2019, pp. 3296–3303.
- [31] Y. Zhu et al., "Multi-representation adaptation network for cross-domain image classification," *Neural Netw.*, vol. 119, pp. 214–221, 2019.
- [32] J. Geng, X. Deng, X. Ma, and W. Jiang, "Transfer learning for SAR image classification via deep joint distribution adaptation networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5377–5392, Aug. 2020.
- [33] Y. Zhu et al., "Deep subdomain adaptation network for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1713–1722, Apr. 2021.
- [34] Y. Luo, Y. Wen, L.-Y. Duan, and D. Tao, "Transfer metric learning: Algorithms, applications and outlooks," 2018, *arXiv:1810.03944*.
- [35] B. Deng, S. Jia, and D. Shi, "Deep metric learning-based feature embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1422–1435, Feb. 2020.
- [36] Y. Dong, T. Liang, C. Yang, H. Luo, and Y. Zhang, "Joint distance transfer metric learning for remote-sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 6506205.
- [37] H. Yang, H. He, T. Li, Y. Bai, and W. Zhang, "Multi-metric domain adaptation for unsupervised transfer learning," *IET Image Process.*, vol. 14, no. 12, pp. 2780–2790, 2020.
- [38] S. Azarbarzin and F. Afsari, "Joint robust transfer metric and adaptive transfer function learning," *Neural Process. Lett.*, vol. 51, no. 2, pp. 1411–1443, 2020.
- [39] J. Wang, Z. Wei, T. Zhang, and W. Zeng, "Deeply-fused nets," 2016, *arXiv:1605.07716*.
- [40] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.
- [41] H. Gao, Y. Zhang, Z. Chen, and C. Li, "A multiscale dual-branch feature fusion and attention network for hyperspectral images classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 8180–8192, 2021.
- [42] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.
- [43] X. Hou, W. Ao, Q. Song, J. Lai, H. Wang, and F. Xu, "Fusar-ship: Building a high-resolution SAR-AIS matchup dataset of gaofen-3 for ship detection and recognition," *Sci. China Inf. Sci.*, vol. 63, no. 4, pp. 1–19, 2020.
- [44] C. Yu, J. Wang, Y. Chen, and M. Huang, "Transfer learning with dynamic adversarial adaptation network," in *Proc. IEEE Int. Conf. Data Mining*, 2019, pp. 778–786.
- [45] M. Long, J. Wang, G. Ding, S. J. Pan, and S. Y. Philip, "Adaptation regularization: A general framework for transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 5, pp. 1076–1089, May 2014.
- [46] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," 2014, *arXiv:1412.3474*.
- [47] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016, pp. 2058–2065.
- [48] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, 2008.
- [49] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
- [50] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-NET: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11531–11539.
- [51] X. Zhang, Y. Lv, L. Yao, W. Xiong, and C. Fu, "A new benchmark and an attribute-guided multilevel feature representation network for fine-grained ship classification in optical remote sensing images," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1271–1285, 2020.
- [52] Z. Zhang, L. Zhang, Y. Wang, P. Feng, and R. He, "Shipsimagenet: A large-scale fine-grained dataset for ship detection in high-resolution optical remote sensing images," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 8458–8472, 2021.



Shuangmei Zhao received the B.E. degree in electronic science and technology in 2019 from the Beijing University of Chemical Technology (BUCT), Beijing, China, where she is currently working toward the M.S. degree in physics with the College of Mathematics and Physics.

Her research interest is the fine-grained classification of SAR ships using transfer learning methods based on deep learning.



Haitao Lang (Member, IEEE) received the B.E. and M.S. degrees in optical engineering from the Ocean University of China, Qingdao, China, in 2000 and 2003, respectively, and the Ph.D. degrees in optical engineering from the Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, China, in 2006 and 2011, respectively.

He has authored or coauthored prolifically in some prestigious refereed journals and conference proceedings such as IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, Computer Vision and Pattern Recognition Conference, Asian Conference on Computer Vision, International Geoscience and Remote Sensing Symposium, SPIE Remote Sensing, etc. He is currently a Professor with the College of Mathematics and Physics, Beijing University of Chemical Technology, Beijing, China. His research interests include machine learning, pattern recognition, and optical engineering with a focus on developing advanced image analysis and interpretation techniques for maritime remote sensing applications.