Article

Domain Adaptive Few-Shot Learning for ISAR Aircraft Recognition with Transferred Attention and Weighting Importance

Binquan Li *, Yuan Yao and Qiao Wang

School of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, China
* Correspondence: lbq@haut.edu.cn; Tel.: +86-18623715003

Abstract: With the enhancement of air-based and space-based perception capabilities, space-aeronautics incorporation and integration is growing in importance. Full domain awareness is crucial for integrated perception systems, in which domain adaptation is one of the key problems in improving the performance of cross-domain perception. Deep learning is currently an advanced technique for complex inverse synthetic aperture radar (ISAR) object recognition. However, the training procedure needs many annotated samples, which is insufficient for certain targets, such as aircraft. Few-shot learning provides a new approach to solving the above problem by transferring useful knowledge from other domains, such as optical images from satellites. Nevertheless, it fails to fully consider the domain shift between the source and target domains, generally neglecting the transferability of training samples in the learning process. Consequently, it produces suboptimal recognition accuracy. To address the composite problems mentioned above, we propose a domain adaptive few-shot learning method from satellites to an ISAR called S2I-DAFSL for aircraft recognition tasks. Furthermore, unlike conventional domain adaptation methods that directly align the distributions, the attention transferred importance-weighting network (ATIN) is proposed to improve the transferability in the domain adaptation procedure. Compared with state-of-the-art methods, it shows that the proposed method achieves better performance, increasing the accuracy and effectiveness of classification, which is more suitable for cross-domain few-shot ISAR aircraft recognition tasks.

Keywords: inverse synthetic aperture radar; aircraft recognition; domain adaptation; few-shot learning

1. Introduction

Aircraft recognition is one of the important tasks in ISAR image identification [1–6]. In recent years, deep convolution neural networks have achieved better performance in ISAR image classification [7–14], which can be applied to aircraft recognition tasks. However, the lack of labeled samples in the training process often results in over-fitting under the condition of small training sets. Recently, a fast space target ISAR image recognition based on local and global structural feature fusion has been proposed [15]. Further, transfer learning also provides a solution by taking the deep CNN model as a feature extractor or for fine-tuning the weight in the new task to address the problem that the samples in the target classification task are too small to effectively train the deep CNN model [16]. Additionally, since humans can quickly learn transferable knowledge based on their experiences and a few observations [17–19], few-shot learning (FSL) approaches are proposed to deal with the problems mentioned above. Including meta-learning, also known as learning to learn, a meta-learner-based stacking network is proposed to realize the classification of space targets by ISAR images with small samples [20]. FSL methods often assume that the samples from the source domain and target domain are from the same distribution without fully considering the domain discrepancy. Domain adaptation (DA) can reduce the domain shift
by aligning the distributions between the source and target domains [21–26]. It is generally assumed that there are enough training samples when the categories from the source domain and target domain are consistent with each other, known as closed-set DA, which is different from the problem setting of FSL being able to identify new unseen categories.

The above scenarios are described in Figure 1.

![Figure 1. Different scenarios of visual recognition problems.](image)

From the perspective of DA methods, there exists a categorization of DA settings known as closed-set, partial, and open-set. This categorization is primarily based on the relationship between label spaces. Closed-set DA occurs when both domains possess identical label space, and the primary challenge lies in bridging the domain gap. In scenarios where the label set relationship is often unknown along with a wide domain gap, partial DA techniques are considered appropriate when the label space of the source domain is expansive enough to encompass the label space of the target domain. Conversely, open-set DA methods are more suitable when the label space of the source domain is restricted within the label space of the target domain.

Open-set DA can be seen as a kind of open-set problem, but focusing on the distribution shift between source and target domains. In open-set problems, it is assumed that there are known classes and unknown classes. It often learns with known classes and tests in all classes. The objective is to thoroughly investigate the capabilities of the model in distinguishing between images of known and unknown classes, with the aim of assessing its potential for open-set recognition based on the principle of “Openness”. The complexity of recognition can be determined by the number of target categories, as well as the training and testing categories utilized [27]. For example, an innovative approach based on proportional similarity is proposed in order to improve the accuracy and robustness of open-set recognition (OSR) in synthetic aperture radar (SAR) images [28].
As mentioned above, the open-set DA methodology offers a more practical approach to DA as it can incorporate novel classes that may emerge in the target domain. This approach involves dealing with unfamiliar categories in the target domain that did not appear in the source domain. To effectively overcome the challenges of open-set DA, it is crucial to reduce the domain shift in identical categories while simultaneously enhancing the discrimination among different categories.

Similar to open-set DA, FSL aims to identify novel target categories by leveraging samples from their source domain while also utilizing limited labeled samples from the target categories. In open-set DA, there are private categories within the target domain, in addition to the common categories between source and target domains. Open-set DA often assumes that the label space of the source domain is included within the label space of the target domain. Nevertheless, FSL often focuses on identifying novel categories when there is no intersection between source classes and target classes.

The different methodologies are further illustrated and presented in Figure 2.

**Figure 2. Illustration of different methodologies.**

As shown in Figure 2, FSL transfers knowledge from the source domain to the target domain: (1) The data distributions of the source domain are consistent with the target domain; (2) The intersection of the categories from the source and target domain is absent; and (3) The labeled samples for each category in the target domain are few.

Closed-set DA also involves transferring knowledge across domains: (1) There is a domain gap between the source and target domains; (2) The categories across domains are the same; and (3) The available samples for each category in the target domain are sufficient.

Open-set DA is similar to closed-set DA except that there are private categories in the target domain, and the source categories are included in the target categories.

DAFSL addresses FSL and DA simultaneously: (1) There is also a domain gap across domains; (2) The intersection of the categories from the source and target domain is absent; and (3) The labeled samples for each category in the target domain is few.

In general, due to the limited availability of labeled data, the field of FSL has recently garnered increased attention. The majority of visual recognition models rely on deep
CNN, which necessitates the collection and annotation of sufficient samples per class to enable effective training. However, this requirement can be impractical or unfeasible for infrequent categories. FSL is often treated as a transfer learning issue, where knowledge is transferred from the source to target classes. The primary focus has been on developing a classifier with minimal samples. A fundamental challenge has been neglected, as the target classes often suffer from insufficient representation due to few training samples and may belong to a different domain compared to the source classes. Thus, it is critical to accommodate both novel categories and domain discrepancies using minimal samples from the target categories.

In other words, while the classification of labeled categories from some datasets can be easily accessed, the identification of novel categories presents a twofold challenge. Firstly, imaging modalities utilized in the target domain may differ significantly from those employed in the source training data, resulting in domain discrepancy. Secondly, new categories may not be adequately represented in the training data, causing a serious category gap.

Directly combining the FSL and DA is a straightforward approach but not appropriate since the distribution alignment in DA is harmful to the distinctiveness of new unseen categories in the target domain. Therefore, aligning the distributions of the source and target domains while maintaining certain independence is fundamentally important.

To address the above problems, in this paper, the attention transferred importance-weighting network (ATIN) is proposed, and FSL is implemented simultaneously in both the source and target categories, allowing it to build a discriminative embedding model as well as explore the transferable information. The major contributions of our work are summarized as follows:

1. A cross-domain FSL method from satellites to ISAR called S2I-DAFSL is proposed to implement the aircraft recognition task, addressing the DA and FSL simultaneously.
2. We propose the ATIN to further improve the transferability and effectiveness in the DA procedure, in which the attention-transferred module focuses on more informative regions, and the importance-weighting module helps choose more appropriate training samples.
3. Extensive experimental results demonstrate that the proposed method can improve the accuracy of ISAR aircraft image classification with more efficient implementation.

The remainder of this paper is organized as follows. Section 2 introduces the overall architecture of the proposed method. Section 3 provides the details about each component of the proposed method. In Section 4, we analyze the experiments, and Section 5 is the paper’s conclusion.

2. Overall Architecture

In the conventional problem settings of DA, the data distributions of the source and target domains are different. Formally, the target domain is indicated by $D_t$, and the source domain is denoted by $D_s$. $n_s$ and $n_t$ are the numbers of samples in the source and target domains, respectively. Due to the domain changes, $P_s \neq P_t$ in the DA settings. Therefore, deep DA approaches are utilized to minimize the distribution discrepancy across domains. As described above, the problem settings of domain adaptive few-shot learning are more challenging than DA or FSL. The proposed method leverages a comprehensive dataset $D_s$ that includes samples from categories $C_s$ in the source domain and few-shot samples $D_d$ from classes $C_d$ in the target domain. In addition, we have access to a separate test set $T$ that represents a different target domain with classes $C_t$, ensuring the absence of overlaps between these classes. The primary objective of our approach is to teach an effectively generalized model to $T$ by utilizing $D_t$ and $D_d$. Obviously, the dissimilarities between the source and target domains mean that $P_s(x)$ of source classes $C_s$ differs from that of the target classes $C_t \cup C_d$. The overall architecture of the proposed method is shown in Figure 3. The prototypical network is utilized to implement FSL while the ATIN performs the DA procedure, containing the attention transferred module and importance-weighting module.
overlaps between these classes. The primary objective of our approach is to teach an ef-
tilling it into two separate abstract feature maps: $M_1$ and $M_2$, with the dimensions $C \times H \times W$. Next, we rearrange them into $C \times N$ and conduct a matrix multiplication operation to transpose them. The self-attention map $M^{N \times N}$ may be calculated using the softmax operation.

$$a_{ij} = \frac{\exp(M_{1i} \cdot M_{2j})}{\sum_{i=1}^{N} \exp(M_{1i} \cdot M_{2j})}$$

where $a_{ij}$ calculates the influence of location $i$ on position $j$. The similarities between the two places and their correlation increase as they become closer. Further, a convolution layer is also used to obtain $M_3$ with the dimensions $C \times H \times W$ and then reshaped into $C \times N$. Then, it is reshaped to $C \times H \times W$ using the feature map $M_3$ and $M^{N \times N}$.

3. The ATIN Approach

The ATIN is thoroughly proposed in this section. The majority of adversarial-based DA techniques have three parts: a feature extractor $G_f$, a classifier $G_y$, and a domain discriminator $G_d$ [29]. The objective functions are shown as

$$\min_{G_f, G_y} L_{cls}(G_f, G_y) = \frac{1}{n_y} \sum_{i=1}^{n_y} L(G_y(G_f(x_i^s)), y_i^s)$$

(1)

$$\min_{G_f} \max_{G_d} L_{adv}(G_f, G_d) = -\frac{1}{n_x} \sum_{i=1}^{n_x} \log(G_d(G_f(x_i^s)))$$

$$-\frac{1}{n_i} \sum_{j=1}^{n_i} \log(1 - G_d(G_f(x_i^s)))$$

(2)

where $L_{cls}$ is the cross-entropy loss for the classification task of the source-domain; $L_{adv}$ is the domain adversarial loss.

However, the current adversarial DA techniques may not take into account complicated multi-modal structures in domain distributions [30]. If inappropriate samples and untransferable features are forced to match, this may result in a negative transfer. Therefore, the ATIN is proposed to address the above problems, which include the attention transferred module and importance-weighting module.

3.1. The Attention Transferred Module

Learning transferable representations across domains is crucial in adversarial DA. However, it is difficult to learn features that are actually both transferable and discriminative. To overcome this problem, a self-attention mechanism for conditional adversarial learning is added. We initially input features from $G_f$ into a convolution layer before distilling it into two separate abstract feature maps: $M_1$ and $M_2$, with the dimensions $C \times H \times W$. Next, we rearrange them into $C \times N$ and conduct a matrix multiplication operation to transpose them. The self-attention map $M^{N \times N}$ may be calculated using the softmax operation.
Next, the weighted feature representations are obtained, which are expressed as

\[ M' = \mu \sum_{i=1}^{N} a_i M_{3i} + M \]  

(4)

The parameter \( \mu \) with an initial value of 0 is progressively increased in the training process [31]. Singular value decomposition (SVD) is employed to investigate the spectrum characteristics of the attention feature \( M' \) in batches, further enhancing the transferability and discriminability [32]. Moreover, batch spectral penalization (BSP) is used to improve the discriminability of the feature representation [33]. Thus, using the greatest \( K \) singular values as a regularization term, the final calculation for self-attention transfer loss is obtained by

\[ L_{BSP}(M') = \sum_{i=1}^{K} (\rho_{s,i}^2 + \rho_{t,i}^2) \]  

(5)

where \( \rho_{s,i} \) and \( \rho_{t,i} \) are the \( i \)-th largest singular values in matrix \( \Sigma_s \) and \( \Sigma_t \).

3.2. The Importance-Weighting Module

In order to determine the significance of each source sample to the target domain, we utilize the output of \( G_d \) and the classifier predictions \( g \). In this approach, the proposed network may lessen the impact of outliers in the source domain.

Let \( h = (f, g) \) represent the joint variable of the classifier prediction \( g \) and feature representation \( f \). The conditioning strategy can be expressed as

\[ T(h) = \begin{cases} 
\otimes(f, g), & d_f \times d_g \leq d_{feat} \\
\odot(f, g), & \text{otherwise}
\end{cases} \]  

(6)

where \( \otimes \) and \( \odot \) stand for a multi-linear map and an explicit randomized multi-linear map. \( d_f \) and \( d_g \) are the dimensions of \( f \) and \( g \), and \( d_{feat} \) denotes the dimension of the output.

As seen in Equation (7), the domain adversarial loss may be expressed as follows:

\[ \min_{G_f, G_y} \max_{G_d} L_{adv} \left( G_f, G_y, G_d \right) = -\frac{1}{n} \sum_{i=1}^{n} \log \left( G_d \left( T(h)_i^s \right) \right) + \frac{1}{n} \sum_{i=1}^{n} \log \left( 1 - G_d \left( T(h)_i^t \right) \right) \]  

(7)

Equation (7) assigns all cases the same weight, which might be problematic since certain samples with ambiguous predictions could make the adversarial adaptation process worse. In most cases, the entropy criteria can be described as

\[ H(g) = -\sum_{i=1}^{C} g_i \log g_i \]  

(8)

where \( g_i \) is the probability of correct prediction of category \( i \) and may be used to quantify the uncertainty of classifier predictions. \( C \) represents the total classes in the source domain. The prediction is more definite when the \( H(g) \) is smaller. As a result, it is possible to condition the domain discriminator by weight \( w(H(g)) = e^{-H(g)} \), and the domain adversarial loss is defined as follows:

\[ \min_{G_f, G_y} \max_{G_d} L_{adv} \left( G_f, G_y, G_d \right) = -\frac{1}{n} \sum_{i=1}^{n} w(H(g_i^s)) \log \left( G_d \left( T(h)_i^s \right) \right) + \frac{1}{n} \sum_{i=1}^{n} w(H(g_i^t)) \log \left( 1 - G_d \left( T(h)_i^t \right) \right) \]  

(9)

However, obviously, in the minimax game represented by Equation (9), \( w(H(g_i^t)) \) soon approaches 1 as training goes on [34]. As each source sample is, thus, equally significant, it will result in inappropriate adaptation.

Intuitively, the transferability of samples can be directly reflected in the output result \( d_i \) of the domain discrimination model. The more uncertain the sample is, the better the transferability is. Thus, the entropy of domain prediction \( d_i \) can be utilized to calculate the transferability of the source domain samples, which is expressed as \( \hat{\omega} = H(d_i) \). The former
G_d is only used to generate the transferability \( \hat{w} \) of the source domain samples, which is described as

\[
\min_{G_d} L_d(G_d) = -\frac{1}{n_s} \sum_{i=1}^{n_s} w(H(g_s^i)) \log \left( G_d(T(h_s^i)) \right) - \frac{1}{n_t} \sum_{j=1}^{n_t} w(H(g_t^j)) \log \left( 1 - G_d(T(h_t^j)) \right)
\]  

(10)

Meanwhile, another adversarial domain discriminator \( \hat{G}_d \) is utilized, which is shown as

\[
\min_{G_f, G_y} \max_{\hat{G}_d} L_{adv}(G_f, G_y, \hat{G}_d) = -\frac{1}{n_s} \sum_{i=1}^{n_s} \hat{w} \cdot w(H(g_s^i)) \log \left( \hat{G}_d(T(h_s^i)) \right) - \frac{1}{m} \sum_{j=1}^{n_t} w(H(g_t^j)) \log \left( 1 - \hat{G}_d(T(h_t^j)) \right)
\]  

(11)

Therefore, the objective of the importance-weighting module consists of Equations (10) and (11).

4. The FSL Procedure

During the episode training phase, we aim to emulate the few-shot testing process by generating small episodic training sets from both the source dataset \( D_s \) and the few-shot dataset \( D_d \). To achieve this, we randomly choose a subset of \( N \) classes from \( D_s \) and divide them into support and query sets, creating training episodes. Nevertheless, owing to the restricted number of samples in \( D_d \), we augment the data with standard techniques such as horizontal flips and random crops. To develop our FSL framework, we have leveraged the widely acclaimed prototypical network. This choice is motivated by its remarkable performance and simplicity, making it an ideal cornerstone for our study. The network operates by encoding a prototype and subsequently utilizing a nearest neighbor classifier to classify samples in the query set. It projects samples from visual space into feature space where samples belonging to the same class are closely clustered, while those from different classes are distinctly separated.

During the training phase, the network parameters undergo an updating process utilizing a nonparametric softmax technique that draws upon the outcomes of a distance metric. This approach effectively determines the distribution of class membership for a query sample \( x \) residing within the query set \( Q \).

\[
P(y = k \mid x \in Q) = \frac{\exp(-\text{dis}(f_{\varphi}(x), c_k))}{\sum_{k=1}^{C} \exp(-\text{dis}(f_{\varphi}(x), c_k))}
\]  

(12)

where \( \text{dis}() \) represents the Euclidean distance function. Next, the FSL loss is obtained using support set \( S \) and query set \( Q \), where \( L_{fss} \) for the source domain and \( L_{fst} \) for the target domain can be described as

\[
L_{fss} = E_{S, Q_s} \left[ - \sum_{(x,y) \in Q_s} \log p_{\varphi}(y = k \mid x) \right]
\]  

(13)

\[
L_{fst} = E_{S, Q_t} \left[ - \sum_{(x,y) \in Q_t} \log p_{\varphi}(y = k \mid x) \right]
\]  

(14)

Therefore, the integrated objective functions of the proposed method are Equations (5), (10), (11), (13) and (14).

5. Experiments

As shown in Figure 3, there are 15 classes of optical aircraft in the source domain [35,36]. To increase the diversity of the tasks and verify the robustness of the proposed method,
10 classes of ISAR images are selected in the target domain, in which 5 classes are synthesized and others are obtained from an anechoic chamber [5]. Next, 5-way 1-shot and 5-way 5-shot settings are utilized to evaluate the test set. The study reports the mean accuracy of 600 testing episodes, which were generated at random, coupled with their corresponding 95% confidence intervals.

ADDA [24], CDAN [22], RelationNet [19], MatchingNet [37], ProtoNet [18], and CDAN + ProtoNet (directly combining CDAN and ProtoNet) are selected as baselines. We utilize the ResNet50 model [38] as the backbone architecture. The backbone is meticulously pre-trained from scratch on the training datasets, and subsequently fine-tuned to address the challenging domain adaptation few-shot learning task. The aim of this study is to investigate the efficacy of a novel approach to address the challenges posed by FSL and DA in cross-domain ISAR aircraft recognition. The results are presented in Table 1, highlighting the following noteworthy findings.

<table>
<thead>
<tr>
<th>Model</th>
<th>5-Way 1-Shot</th>
<th>5-Way 5-Shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADDA [24]</td>
<td>73.25 ± 0.55</td>
<td>83.98 ± 0.39</td>
</tr>
<tr>
<td>CDAN [22]</td>
<td>73.72 ± 0.43</td>
<td>84.11 ± 0.51</td>
</tr>
<tr>
<td>RelationNet [19]</td>
<td>72.58 ± 0.46</td>
<td>82.47 ± 0.28</td>
</tr>
<tr>
<td>MatchingNet [37]</td>
<td>73.06 ± 0.53</td>
<td>83.62 ± 0.40</td>
</tr>
<tr>
<td>ProtoNet [18]</td>
<td>71.79 ± 0.26</td>
<td>81.88 ± 0.31</td>
</tr>
<tr>
<td>CDAN + ProtoNet</td>
<td>74.02 ± 0.41</td>
<td>84.35 ± 0.29</td>
</tr>
<tr>
<td>S2I-DAFSL</td>
<td>75.97 ± 0.37</td>
<td>86.93 ± 0.26</td>
</tr>
</tbody>
</table>

Our proposed method outperforms FSL or DA methods, demonstrating its potential as a promising solution to address the composite problems. Furthermore, our comparative analysis with DA + FSL baselines emphasizes the need for more sophisticated models as the naive combination of strategies has limitations.

The proposed method addresses the challenges associated with bridging the gap between different domains by combining both FSL and DA effectively. To evaluate the robustness of the proposed method, we change the number of $n$ and $k$ in the $n$-way $k$-shot settings.

As shown in Figure 4, when the number of $n$ decreases or the number of $k$ increases, the accuracy of recognition is improved. When $n$ is 5, as $k$ increases, the recognition accuracy improves significantly. When $k$ is 9, it does not matter if $n$ equals 3 or 5, the recognition accuracy is more than 95%. As $n$ equals 3, it reaches 99.1%. The above results prove the robustness and effectiveness of the proposed method.

In conclusion, the proposed method presents a promising solution for addressing the challenges associated with DA in FSL for ISAR aircraft recognition.
Our proposed method outperforms FSL or DA methods, demonstrating its potential as a promising solution to address the composite problems. Furthermore, our comparative analysis with DA + FSL baselines emphasizes the need for more sophisticated models as the naive combination of strategies has limitations.

The proposed method addresses the challenges associated with bridging the gap between different domains by combining both FSL and DA effectively. To evaluate the robustness of the proposed method, we change the number of \( n \) and \( k \) in the \( n \)-way \( k \)-shot settings.

As shown in Figure 4, when the number of \( n \) decreases or the number of \( k \) increases, the accuracy of recognition is improved. When \( n \) is 5, as \( k \) increases, the recognition accuracy improves significantly. When \( k \) is 9, it does not matter if \( n \) equals 3 or 5, the recognition accuracy is more than 95%. As \( n \) equals 3, it reaches 99.1%. The above results prove the robustness and effectiveness of the proposed method.

Figure 4. Recognition accuracy under varied \( n \)-way \( k \)-shot settings.

In addition, we implemented an ATIN approach to enhance our full model. To evaluate the efficacy of each module, we conducted an ablation study by comparing the performance of our complete model against four simplified versions, i.e., FSL, ATIN, ATIN-I (using only the importance-weighting module), and ATIN-A (using only the attention transferred module) in 5-way 1-shot and 5-way 5-shot settings.

The results, as shown in Figure 5, demonstrate that the integration of each module improves the performance, proving the contribution of each module in resolving the ISAR image classification. Our study further emphasizes the crucial role of the ATIN, as demonstrated by the performance improvements achieved by ATIN over classical FSL. In conclusion, the proposed method presents a promising solution for addressing the challenges associated with DA in FSL for ISAR aircraft recognition.

Figure 5. Recognition accuracy using each module.

Furthermore, to analyze the recognition performance of ISAR images that are synthesized or from an anechoic chamber, we only take synthesized images, those from an anechoic chamber and their mixing as a test set, respectively. As shown in Figures 6–8, the confusion matrix is obtained under the 5-way 5-shot setting, and the per class indicators for all confusion matrices are presented in Tables 2–4.

As shown in Figures 6 and 7, the recognition accuracy of synthesized samples is affected by the appearance of the aircraft. When the samples are only selected from the anechoic chamber, the confusion of each category is increased.
It also can be seen in Figure 8 that there is no confusion between synthesized samples and those from the anechoic chamber, proving the effectiveness of the proposed method.

Further, the t-SNE [39] is used to visualize the tested target features with corresponding category labels according to the above settings, where each category contains 31 samples in the 5-way 1-shot setting and 35 samples in the 5-way 5-shot setting.

![Confusion matrix of synthesized images.](image1)

**Figure 6.** Confusion matrix of synthesized images.

![Confusion matrix of images obtained from anechoic chamber.](image2)

**Figure 7.** Confusion matrix of images obtained from anechoic chamber.

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.80</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>2</td>
<td>0.91</td>
<td>0.67</td>
<td>0.77</td>
</tr>
<tr>
<td>3</td>
<td>0.71</td>
<td>1.00</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2. Per class indicators for confusion matrix of synthesized images.
As shown in Figure 9, the distributions of features from the anechoic chamber are different compared with those from the synthesized samples. The recognition accuracy under the 5-way 5-shot setting is higher than the 5-way 1-shot setting. Each category can be generally distinguished in each case. When the categories are mixed, the features of the synthesized samples and those from the anechoic chamber (viewed in green and yellow...
in Figure 9c) can also be well separated, proving the robustness of the proposed method, which is suitable for ISAR aircraft recognition.

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95</td>
<td>0.60</td>
<td>0.74</td>
</tr>
<tr>
<td>2</td>
<td>0.70</td>
<td>0.47</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>0.53</td>
<td>0.80</td>
<td>0.64</td>
</tr>
<tr>
<td>4</td>
<td>0.78</td>
<td>0.97</td>
<td>0.86</td>
</tr>
<tr>
<td>5</td>
<td>0.93</td>
<td>0.90</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 3. Per class indicators for confusion matrix of images obtained from anechoic chamber.

As shown in Figures 6 and 7, the recognition accuracy of synthesized samples is affected by the appearance of the aircraft. When the samples are only selected from the anechoic chamber, the confusion of each category is increased.

It also can be seen in Figure 8 that there is no confusion between synthesized samples and those from the anechoic chamber, proving the effectiveness of the proposed method.

Further, the t-SNE [39] is used to visualize the tested target features with corresponding category labels according to the above settings, where each category contains 31 samples in the 5-way 1-shot setting and 35 samples in the 5-way 5-shot setting.

As shown in Figure 9, the distributions of features from the anechoic chamber are different compared with those from the synthesized samples. The recognition accuracy under the 5-way 5-shot setting is higher than the 5-way 1-shot setting. Each category can be generally distinguished in each case. When the categories are mixed, the features of the synthesized samples and those from the anechoic chamber (viewed in green and yellow in Figure 9c) can also be well separated, proving the robustness of the proposed method, which is suitable for ISAR aircraft recognition.

Figure 9. Feature visualization in different combinations of categories in 5-way 1-shot and 5-way 5-shot settings (best viewed in color): (a) Features of synthesized samples; (b) Features of samples from anechoic chamber; (c) Features of samples from mixed categories.
6. Conclusions

We present a novel domain adaptive few-shot learning method from satellites to an ISAR called S2I-DAFSL for aircraft recognition. Rather than aligning distributions directly, our approach leverages an ATIN to further enhance transferability during the domain adaptation process. Compared with the state-of-the-art approaches, the proposed method achieves better classification accuracy and effectiveness, making it more suitable for ISAR aircraft recognition tasks. It is worth noting that although our work suggests a promising solution for ISAR aircraft recognition, there is still room for improvements. For example, the proposed method is evaluated on a specific task, and the availability of labeled data in the source domain is assumed, which could be further extended to other tasks and domains. In the future, we can also consider extending the framework to implement open-set recognition tasks as well as zero-shot learning under the unsupervised learning paradigm.

Author Contributions: Conceptualization, B.L., Y.Y. and Q.W.; methodology, B.L., Y.Y. and Q.W.; validation, B.L.; data curation, Q.W.; investigation, Y.Y.; resources, Q.W.; writing—original draft preparation, B.L. and Y.Y.; writing—review and editing, Q.W.; supervision, B.L. and Y.Y.; project administration, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Henan Provincial Science and Technology Project (grant numbers 202102210136, 2210222009, 222102220080, and 222103810083), the Henan Province Education Department Natural Science Project (grant number 20A413003), the R&D Project of Zhengzhou (grant number 22ZZRDZX06), and the Natural Science Project of Henan University of Technology (grant numbers 2019BS018 and 2019BS055).

Data Availability Statement: The data used in this study are collected according to references [5,35,36].

Acknowledgments: All authors would sincerely thank the reviewers and editors for their suggestions and opinions for improving this article.

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

References


15. Yang, H.; Zhang, Y.S.; Ding, W.Z. A fast recognition method for space targets in ISAR images based on local and global structural fusion features with lower dimensions. *Int. J. Aerosp. Eng.* 2020, 2020, 3412582. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.