

基于跨域小样本学习的SAR图像目标识别方法

史松昊, 王晓丹, 杨春晓, 王艺菲

引用本文

史松昊, 王晓丹, 杨春晓, 王艺菲. [基于跨域小样本学习的SAR图像目标识别方法](#)[J]. 计算机科学, 2024, 51(6A): 230800136-7.

SHI Songhao, WANG Xiaodan, YANG Chunxiao, WANG Yifei. [SAR Image Target Recognition Based on Cross Domain Few Shot Learning](#) [J]. Computer Science, 2024, 51(6A): 230800136-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[面向产线AI质检的少样本评测方法研究和验证](#)

Study and Verification on Few-shot Evaluation Methods for AI-based Quality Inspection in Production Lines

计算机科学, 2024, 51(6A): 230700086-8. <https://doi.org/10.11896/jsjcx.230700086>

[基于BERT和CNN的药物不良反应个案报道文献分类方法](#)

Literature Classification of Individual Reports of Adverse Drug Reactions Based on BERT and CNN

计算机科学, 2024, 51(6A): 230400049-6. <https://doi.org/10.11896/jsjcx.230400049>

[DUWe:动态未知词嵌入方法在Web异常检测中的应用](#)

DUWe:Dynamic Unknown Word Embedding Approach for Web Anomaly Detection

计算机科学, 2024, 51(6A): 230300191-5. <https://doi.org/10.11896/jsjcx.230300191>

[基于知识蒸馏的差分隐私联邦学习方法](#)

Differential Privacy Federated Learning Method Based on Knowledge Distillation

计算机科学, 2024, 51(6A): 230600002-8. <https://doi.org/10.11896/jsjcx.230600002>

[WiCare:一种非接触式的老人如厕跌倒监测模型](#)

WiCare:Non-contact Fall Monitoring Model for Elderly in Toilet

计算机科学, 2024, 51(6A): 230700044-8. <https://doi.org/10.11896/jsjcx.230700044>

基于跨域小样本学习的 SAR 图像目标识别方法

史松昊 王晓丹 杨春晓 王艺菲

空军工程大学防空反导学院 西安 710051

(cutee_squirrel@163.com)

摘要 由于 SAR 图像获取难度大,可供研究的样本数量较少,解决有限样本条件下 SAR 图像目标识别问题成为业界公认的挑战。随着深度学习在计算机视觉领域的发展,衍生出了多种小样本图像分类方法,因此考虑采用跨域小样本学习范式解决小样本 SAR 图像目标识别问题。具体地,先在多个源域中训练得到不同域的特征提取器,而后通过知识蒸馏的方法获取一个通用的特征提取器,这里采用中心核对齐的方法,将提取的特征映射到一个更高维的空间,从而更好地区分原特征之间的非线性相似性;通过上一阶段获得的通用特征提取器提取目标域图像特征,最后采用原型网络的方法预测样本的类别。实验证明,该方法在缩减模型参数的同时,获得了 88.61% 的准确率,为解决小样本 SAR 图像目标识别问题提供了新的思路。

关键词: 深度学习;元学习;跨域小样本学习;SAR 图像目标识别;知识蒸馏

中图分类号 TP18;TN957.52

SAR Image Target Recognition Based on Cross Domain Few Shot Learning

SHI Songhao, WANG Xiaodan, YANG Chunxiao and WANG Yifei

College of Air and Missile Defense, Air Force Engineering University, Xi'an 710051, China

Abstract Due to the difficulty in acquiring SAR images and the scarce number of samples available for research, solving the SAR image target recognition problem under few shot conditions has become a community-recognized challenge. With the development of deep learning in the field of computer vision, a variety of few-shot image classification methods have been derived, so a cross-domain few-shot learning paradigm is considered to solve the few-shot SAR image target recognition problem. Concretely, the feature extractors of different domains are first trained in multiple source domains, while a generalized feature extractor is obtained by knowledge distillation. In this stage, the central kernel alignment method is used to map the extracted features to a higher dimensional space, so as to better distinguish the nonlinear similarity between the original features. Then the target domain image features are extracted by the generalized feature extractor obtained in the previous stage. Finally, a prototype network approach to predict the class of the sample. The experiment proves that the method obtains 88.61% accuracy while reducing the model parameters, which provides a new method for solving the target recognition problem of SAR images with scarce samples.

Keywords Deep learning, Meta learning, Cross domain few shot learning, SAR image target recognition, Knowledge distillation

合成孔径雷达(Synthetic Aperture Radar, SAR)是一种主动式微波传感器,其依据目标对雷达电磁照射后的散射回波收集相位和幅度信息,形成类似于光学图像的遥感样本,可广泛应用于目标的侦察、监视和识别。随着 SAR 图像的获取,逐步产生了 SAR 自动目标识别技术(Automatic Target Recognition, ATR),其本质上是一种图像分类的过程,主要包括预处理、特征提取和分类识别 3 个阶段,利用深层神经网络解决 SAR 目标样本的分类检测问题。SAR-ATR 技术需要大量的样本进行训练,而在实际情况中这一条件却往往无法得到满足,原因有三:1)SAR 图像由专业设备产生,只有少部分的专业人士才能获得,一般研究人员无法获得大量的样本;2)SAR 图像样本通常无法公开,现有的公开数据集也较少,尤其是军事领域,很难获取质量高、数量多的 SAR 图像样本;3)SAR 图像成像机制与普通的光学图像不同,成像频率超出了可见光范围,需要专业人员进行标注,人工成本

较高。因此,SAR 图像识别可以被看作是一个小样本学习的过程,在小样本的条件下获得较好的识别率具有重要的研究价值。

针对 SAR 图像目标识别问题,Ying 等^[1]提出了一种自监督对比学习的方法,通过构建正负样本对增强网络提取特征的能力;Feng 等^[2]采用包括噪声抑制和边缘信息提取在内的数据融合的方法将样本融合为双通道的灰度图,再对特征进行学习;Wang 等^[3]以传统稀疏表示分类为基础,在全局字典上求解稀疏表示系数向量,通过比较不同类别的重构误差大小进行目标类别确认;Kang 等^[4]提出一种稀疏先验引导模块,对提取的目标特征内在低微子空间结构进行学习。以上方法都在 SAR 图像分类上获取了不错的效果,但都未在样本量稀少的条件下测试模型的性能。

最近的研究^[5]表明,经典的小样本学习^[6-12]是建立在源域和目标域数据集同分布或相似分布的假设之上,这样训练

基金项目:国家自然科学基金(61876189,61703426,61806219)

This work was supported by the National Natural Science Foundation of China(61876189,61703426,61806219).

通信作者:王晓丹(afeu_wang@163.com)

得到的模型性能最终会在跨域的情形下急剧下降。为了解决这一问题,研究人员提出了跨域小样本学习^[18]的概念,相关的研究很多,如 CNAPS^[14]和 Simple CNAPS^[15]通过调整建立在 FiLM^[16]层上的基于特定任务的适应模块,让模型的特征提取器适应目标域分布,只是后者使用了基于 Mahalanobis 距离的无参数分类器;SUR^[17]和 URT^[18]则对多个源域中提取出的不同特征进行组合,前者采用一种加权选择的方法,后者采用一种注意力机制,获得对目标域的通用表示方法;STARTUP^[19]放宽了跨域小样本学习条件限制,在训练阶段引入了大量无标签的目标域数据;ME-D2N^[20]则在训练阶段引入了少量有标签的目标域数据,并采用知识蒸馏的方式解决长尾识别的问题,消除网络对不同数据集的“偏见”;URL^[21]提出使用知识蒸馏的方式获得一个通用的表示,然后采用度量学习的方法在目标域进行适应。以上几种跨域方法都是基于自然元素的透视彩色图像进行研究^[22]的,如果直接应用于成像机制完全不同的 SAR 图像上会导致模型性能下降。本文针对这一问题对跨域小样本学习范式进行了改进,以便在 SAR 图像上获得更好的分类效果。

本文主要贡献如下:1)针对 SAR 图像样本量稀少的特点,考虑到光学图像较 SAR 图像采集获取更加容易,因此采用跨域小样本学习的方法来解决 SAR 图像识别的难题。也就是通过引入外部数据集^[23](普通光学图像),让模型在具有丰富样本的源域中学会学习,再迁移到有限样本条件下的目标域中(SAR 图像)进行分类识别。2)由于 SAR 图像分布与光学图像分布差异很大,对于经典的小样本学习方法,其性能在跨域时会被严重制约,因此采用二阶段的跨域元学习方法:第一阶段先在不同的光学图像源域上训练,得到各个域的特征提取器,再利用知识蒸馏的方法获得多个域的通用表示,使其具备提取不同类别图像高层语义特征的能力;第二阶段建立一个适应机制,让模型适应 SAR 图像的样本分布特点,从而提高模型在目标域的分类效果。本文提出的方法初步探索了跨域小样本学习方法在普通可见光学图像迁移到 SAR 图像的可行性,同时降低了巨大的域间隙对模型迁移造成的负面影响,较为有效地解决了 SAR 图像小样本条件下的分类问题,获得了 88.61% 的准确率,具有一定的借鉴意义。

1 问题描述

本文采用元学习的小样本学习方法,在这种情况下,构建一个标准的元学习设置—— N -way K -shot 任务来训练和测试网络,一个任务包含一个支持集 $S = \{x_i, y_i\}^{N \times K}$ 和一个查询集 $Q = \{x_i, y_i\}^{N \times M}$,其中 x 表示样本; y 表示对应的标签; N 表示从数据集中抽取的类别个数; K 表示指每个类别分别抽取的样本个数; M 表示指每个任务中用于测试的样本个数,每个任务都是让网络学会将查询集中的样本从支持集中寻找对应类别的过程,基本流程如图 1 所示。

所谓跨域小样本学习,指依托大量丰富的源域 $D_{src} = \{x_{src}, y_{src}\}$ 进行训练,使神经网络具备学会学习(Learn to Learn)的能力,而后在从未见过的只有少量带标签样本的目标域 $D_{tgt} = \{x_{tgt}, y_{tgt}\}$ 中进行识别分类。对于源域与目标域图像种类严格互斥的情况,其数据分布相差很大,因此存在很大的域间隙,传统的小样本学习和基于微调的方法已经无法在

这样的条件设置下获取令人满意的效果,而跨域小样本学习可以较好地解决该问题。

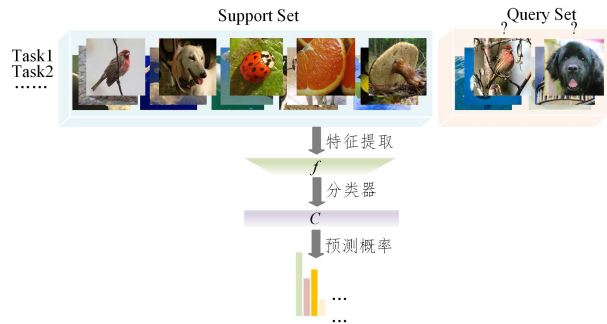


图 1 元学习基本流程示意图

Fig. 1 Basic process of meta-learning

2 模型介绍

这里构建一个两阶段的训练测试过程:一是元训练阶段获取一个基于多源域的通用特征提取网络;二是元测试阶段学习一个特征调整策略对目标域进行适应。

2.1 元训练阶段:获取通用的表示

假设 $D_{src}^1, D_{src}^2, \dots, D_{src}^N$ 代表从 N 个源域 D_{src} 中抽取的样本集合,根据一种经典的多域学习模型^[24],获得通用的特征提取网络的目标函数为:

$$L_m = \min_{\theta, \varphi_n} \sum_{n=1}^N \frac{1}{|D_{src}^n|} \sum_{(x, y) \in D_{src}^n} CE(h_{\varphi_n} \circ f_{\theta}(x), y) \quad (1)$$

其中, x 为样本 y 对应的标签; f 代表这 N 个源域上共享参数 θ 的通用特征提取网络,其输出为一个 d 维的特征; h 代表参数为 φ_n 的特定域的分类器,输出一个概率向量,即属于各个类别的概率;CE 代表交叉熵损失。由于不同域的特性相差较大,导致不同域上的小样本学习任务的侧重点不同,且同一个网络的参数在多个这样的域上迭代训练后会致灾难性遗忘^[25-26],因此优化式(1)最终得到的特征提取网络和分类器性能较差,在某个域上的表现甚至不如该域专用的提取器和分类器。

为了解决这一问题,让模型在下一阶段更好地提取 SAR 图像语义特征,首先需要让其能够提取多种光学图像特征^[27]。本文方法针对每个源域都训练一个特征提取网络 f_{θ_n} 和分类器 h_{φ_n} ,而后将多个源域网络作为教师网络,利用知识蒸馏的方法训练得到一个学生网络,即通用的特征提取网络。它继承了教师网络提取多种图像特征的能力,并且模型架构更为精简、参数规模更小,同时不损失在各个源域上提取特征的能力,也就意味着下阶段在 SAR 图像域上才能拥有更好的表现。

目标蒸馏(Logits 方法)作为知识蒸馏的主流方法之一,其目的是让教师网络和学生网络输出的概率分布结果尽可能相近^[28],也可以理解为通过最小化教师网络和学生网络 logits 的分布差异来得到学生网络的参数,这一过程等价于最小化式(2):

$$L_{\text{logits}} = CE^p(h_{\varphi_n} \circ f_{\theta}(x), h_{\varphi_n^*} \circ f_{\theta_n^*}(x)) \quad (2)$$

其中, CE^p 代表 KL 散度。

SAR 是通过收集主动发射的雷达信号在目标物上的回波来捕获信息的,其雷达回波的振幅信息被转化为像素的灰度值从而形成 SAR 图像。这一特殊的成像机制使得 SAR 图

像具有丰富的纹理特征、边缘特征等底层特征信息。一个自然的想法就是,在第一阶段训练得到的特征提取网络,也能够充分挖掘样本的底层特征,因此需要引入另一种主流的知识蒸馏方法——特征蒸馏,这种方法能够让学生模型充分学习到教师模型提取中间层特征的能力^[29],该过程等价于最小化式(3):

$$L_d = d^f(f_\theta(\mathbf{x}), f_{\varphi_n}(\mathbf{x})) \quad (3)$$

其中, d^f 代表特征空间内的距离度量函数。

对于特征蒸馏,给定同一个输入,教师网络和学生网络的输出特征仍旧可能存在很大差异^[30],直接使用距离度量函数对特征的相似度进行比较,这种方式局限性较大^[28],会使得模型丧失非线性度量的能力,尤其是下一阶段针对 SAR 图像目标域完全不同的分布而言,这种缺陷会被放大。受到 Kornblith 等^[31]工作的启发,这里采用一个更为复杂的非线性映射来度量不同的特征之间的关系——中心核对齐(Center Kernel Alignment,CKA),将提取的特征映射到同一个更高维的空间中进行比较。把这一映射模块记为 $C_{\tau_n} \in \mathbb{R}^{d \times d}$,其参数与通用的特征提取网络参数一起进行训练,记为 τ_n ,激活函数为高斯函数 $\phi(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right)$,能够将原特征编码成更高维度的表征,从而区分原特征之间的非线性相似性。

给定一个样本集 $\{x_1, x_2, \dots, x_i\}$,令

$$\mathbf{M} = [C_{\tau_n} \circ f_\theta(x_1), \dots, C_{\tau_n} \circ f_\theta(x_i)]^T \in \mathbb{R}^{i \times d} \quad (4)$$

$$\mathbf{Y} = [f_{\theta_n}(x_1), \dots, f_{\theta_n}(x_i)]^T \in \mathbb{R}^{i \times d} \quad (5)$$

\mathbf{M} 表示通用特征提取网络输出的特征被变换模块 C_{τ_n} 作用后组成的矩阵, \mathbf{Y} 表示各个单域网络直接提取的特征组成的矩阵。分别计算 \mathbf{M} 和 \mathbf{Y} 的 RBF 核矩阵 \mathbf{P} 和 \mathbf{T} :

$$\mathbf{P} = \begin{pmatrix} \phi(x_1, x_1) & \dots & \phi(x_1, x_i) \\ \vdots & \ddots & \vdots \\ \phi(x_i, x_1) & \dots & \phi(x_i, x_i) \end{pmatrix} \quad (6)$$

$$\mathbf{T} = \begin{pmatrix} \phi(x_1', x_1') & \dots & \phi(x_1', x_i') \\ \vdots & \ddots & \vdots \\ \phi(x_i', x_1') & \dots & \phi(x_i', x_i') \end{pmatrix} \quad (7)$$

然后用核矩阵 \mathbf{P} 和 \mathbf{T} 来度量 \mathbf{M} 和 \mathbf{Y} 之间的 CKA 相似度:

$$CKA(\mathbf{M}, \mathbf{Y}) = \frac{\text{tr}(\mathbf{P}\mathbf{H}\mathbf{T}\mathbf{H})}{\sqrt{\text{tr}(\mathbf{P}\mathbf{H}\mathbf{P}\mathbf{H})\text{tr}(\mathbf{T}\mathbf{H}\mathbf{T}\mathbf{H})}} \quad (8)$$

其中, $x_1 \dots x_i$ 和 $x_1' \dots x_i'$ 分别代表 \mathbf{M} 和 \mathbf{Y} 的列向量, tr 代表矩阵的迹,中心化矩阵 $\mathbf{H}_i = \mathbf{I} - \frac{1}{i}\mathbf{i}\mathbf{i}^T$, \mathbf{I} 为单位矩阵, \mathbf{i} 为全 1 向量,则损失函数 $d^f(\mathbf{M}, \mathbf{Y}) = 1 - CKA(\mathbf{M}, \mathbf{Y})$ 代表两种特征之间的非相似度。由于使用 CKA 相似性需要在整个数据集上计算核矩阵,这在体量庞大的数据集上不便操作,因此改为在每个 minibatch 中进行计算^[32]。

综上,学生网络不仅需要学会在多个源域中提取特征,并且其提取的特征以及最终的分类结果应与教师网络尽可能接近,因此在式(1)的基础上,可以得到该阶段的损失函数:

$$L_h = \min_{\theta, \varphi_n, \tau_n} \sum_{n=1}^N \frac{1}{|D_{src}^n|} \sum_{(x, y) \in D_{src}^n} (CE(h_{\varphi_n} \circ f_\theta(\mathbf{x}), y) + \kappa_n^h L_{\text{logits}} + \kappa_n^f d^f(C_{\tau_n} \circ f_\theta(\mathbf{x}), f_{\theta_n}(\mathbf{x}))) \quad (9)$$

其中, κ_n^h 和 κ_n^f 是超参数,表示针对任务 n 的域权重。优化式(2)的目的就是,获取一个通用特征提取网络 $\{f_\theta(x_i)\}_{i=1}^{|S|}$,在映射模块 C_{τ_n} 的作用下,其提取的特征需要和单源域网络提取的特征尽可能相近,并且能够预测真实的类分布 y_n ,具体流程如图 2 所示。

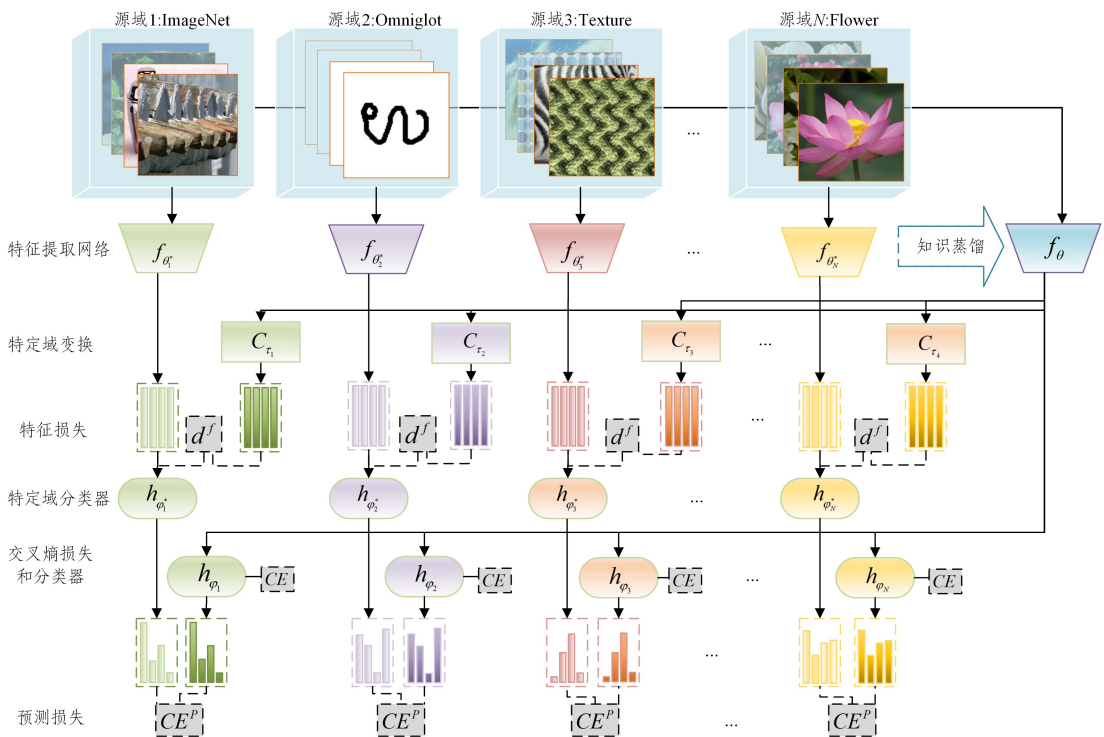


图 2 训练通用特征提取网络流程

Fig. 2 Flow chart of training general feature extraction network

2.2 元测试阶段:调整多域特征

由于 SAR 图像成像机制非常特殊,其具有目标姿态敏感

性强、随目标方位角变化的特点,因此即使是同一目标,所呈现出的视觉外观、散射中心和几何轮廓也并不完全相同,而不

同的目标却有可能具有非常相似的外观,这就造成了较大的类内差异和较小的类间差异性。在元测试阶段,使用上一阶段训练得到的通用特征提取网络 $\{f_\theta(\mathbf{x}_i)\}_{i=1}^{S_1}$ 来提取 SAR 图像目标域特征后,如果直接进行分类可能导致模型易混淆不同类别的图像,分类性能严重下降,因此需要针对这一问题加以改进。这里引进一个线性转换模块 $C_\tau: R^d \rightarrow R^d$,将输出的特征映射到一个更容易分辨的空间,而这一模块是在目标域的支持集上训练得到的,因此可以充分适应目标域的特点,从而实现更好的分类效果。

令 $\mathbf{z}_i = \{C_\tau \circ f_\theta(\mathbf{x}_i)\}_{i=1}^{S_1}$,即经过 C_τ 转换过后的特征, $\tau \in R^{d \times d}$ 。然后采用中心聚类^[11]的思想,得到转换后特征的类中心:

$$\mathbf{c}_j = \frac{1}{|S_j|} \sum_{\mathbf{z}_i \in S_j} \mathbf{z}_i, S_j = \{\mathbf{z}_k : y_k = j\}, j = 1, \dots, N \quad (10)$$

其中, N 表示支持集的样本类别个数,接下来得到支持集某一个样本 \mathbf{z} 属于某一个类别 l 的概率:

$$p(y=l|\mathbf{z}) = \frac{\exp(-d(\mathbf{z}, \mathbf{c}_l))}{\sum_{j=1}^C \exp(-d(\mathbf{z}, \mathbf{c}_j))} \quad (11)$$

其中, $d(\mathbf{z}, \mathbf{c}_l)$ 表示欧氏距离,并通过式(12)来优化 τ ,得到模块 C_τ 。

$$L_{\text{test}} = \min_{\theta} \frac{1}{|S|} \sum_{(\mathbf{x}_i, y_i) \in S} \log(p(y=y_i | \mathbf{x}_i)) \quad (12)$$

以上步骤能够使得样本类内相似度尽可能高、类间相似度尽可能低,克服了 SAR 图像姿态敏感性导致易混淆、难分类的问题。通过 τ 和式(9)得到特征中心点 \mathbf{c}_j ,继而评估查询集 Q 中的每个样本距离哪个中心点 \mathbf{c}_j 近,对其进行分类预测,具体流程如图3所示。

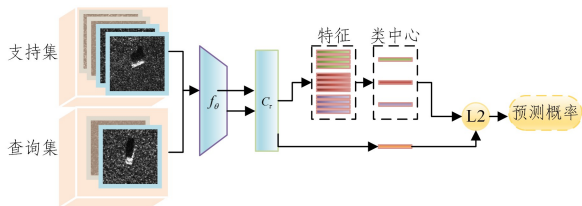


图3 元测试阶段示意图

Fig. 3 Schematic diagram of meta-testing stage

对于如何评估查询集样本与支持集类中心的距离,经典的方法包括最近邻法(NCC)、支持向量机(SVM)、逻辑回归(LR)、 K -近邻法(KNN)以及SCM法(Simple CNAPS^[14]提出的使用马氏距离进行度量)。各方法的评估详见实验部分,其中SCM方法的距离度量公式如下:

$$d(\mathbf{z}, c) = \frac{1}{2} (f_\theta(\mathbf{x}) - c')^T \mathbf{Q}^{-1} (f_\theta(\mathbf{x}) - c') \quad (13)$$

其中, \mathbf{Q} 为某一特定任务的协方差矩阵, c' 为特征的类中心。Bateni等^[14]指出,加入类的协方差矩阵能够让特征提取网络更好地适应目标任务。假设 \mathbf{Q}^{-1} 可分解为一个下三角矩阵和其共轭转置的乘积,即 $\mathbf{Q}^{-1} = \mathbf{L}\mathbf{L}^T$,因此可以考虑先将特征转换为 $\mathbf{z} = \mathbf{L}^T f_\theta(\mathbf{x})$,而后再计算 \mathbf{z} 与类中心的距离。

3 实验设置

3.1 数据集

Meta-dataset^[33]自2020年被提出后,就成为跨域小样本学习的主要基准。它包含10个子数据集:ILSVRC_2012^[34](ImageNet),Omniglot^[35],FGVC-Aircraft^[36](Aircraft),CUB-200-2011^[37](Birds),Describable Textures^[38](DTD),Quick-Draw^[39],FGVCx Fungi^[9],Traffic Signs^[40]以及MSCOCO^[41],通常将前8个数据集用于元训练阶段,并且每个数据集都被划分为训练集、测试集和验证集3个部分,划分详情见表1。Meta-dataset包含了丰富的图像种类,为跨域小样本学习提供了更加真实、大规模的源域。

用于目标域识别的SAR图像数据集为MSTAR(The Moving and Stationary Target Acquisition and Recognition),由美国桑迪亚国家实验室高分辨率的聚束式合成孔径雷达采集多种前苏联目标军事车辆的SAR图像形成,涵盖BMP2, BTR70, T72, BTR60, 2S1, BRDM2, D7, T62, ZIL131, ZSU23/4共计10种型号。成像角度主要分为 15° 和 17° 两个部分,在 $0^\circ \sim 360^\circ$ 方位角下,每隔 $1^\circ \sim 2^\circ$ 获取一次目标切片图像,每一类有200张左右的样本图片,都裁切为 84×84 的尺寸,且将每张图片的目标放在正中心。数据集划分: 15° 的10类目标作为训练集, 17° 的随机5类目标作为测试集,剩余5类目标作为验证集。划分详情如表1所列。

表1 数据集构成

Table 1 Dataset compositions

图像类型	数据集	内容	划分类别数量(训练/验证/测试)
普通光学 图像	ilsvrc_2012(ImageNet)	大型可视化数据集,种类繁多	1000(712/158/130)
	omniglot	手写字符	1623(883/81/659)
	aircraft	飞机	100(70/15/15)
	cu_birds	鸟类	200(140/30/30)
	dtd(Describable Textures)	纹理	47(33/7/7)
	quickdraw	手绘图	345(241/52/52)
	fungi	真菌类	1394(994/200/200)
vgg_flower	花类	102(71/15/16)	
SAR 图像	MSTAR	10类前苏联车辆(SOC)	20(10(15°)/5(17°)/5(17°))

3.2 实验环境

操作系统为Ubuntu 22.04.2版本,在CUDA 11.8环境下运行,使用Python3.7语言、tensorflow-gpu2.4.1和pytorch1.7.0深度学习框架,硬件配置为Nvidia GeForce RTX 3080Ti。

选取ResNet-18作为骨干网络,使用动量(Momentum)

随机梯度下降法(SGD)作为优化器,采用余弦退火的学习率迭代方法,迭代次数设为50000次,超参数 κ_n^l 和 κ_n^r 在分别设置为4和1。

3.3 性能评估指标

为验证模型性能,采用准确率(Accuracy)和损失值(Loss)来评估网络的分类情况。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (14)$$

$$P = \frac{TP}{TP + FP} \quad (15)$$

$$R = \frac{TP}{TP + FN} \quad (16)$$

$$F1 = \frac{2 \times P \times R}{R + P} \quad (17)$$

其中, TP 为真正例, FN 为假反例, FP 为假正例, TN 为真反例, P 为精确率, R 为召回率, M 为标签总数, T_i 为真实标签类别为 i 的样本中被正确分类的总数。

3.4 实验结果与分析

3.4.1 在源域、目标域上测试的准确度

为了检验模型在源域和目标域上的效果,本文在作为源域的 8 个数据集和目标域的 1 个数据集上都进行了测试,由

表 2 在源域、目标域上的平均准确率(95%置信率)

Table 2 Average accuracy(95% confidence) in source domain and target domain

域	数据集	SDL	MDL	Ours
源域	ilsvrc_2012	48.16±1.04	50.10±1.06	56.08±1.05
	omniglot	94.15±0.44	94.44±0.41	93.84±0.46
	aircraft	85.15±0.64	85.35±0.57	87.23±0.55
	cu_birds	73.75±0.86	74.47±0.81	80.58±0.75
	dtd	72.21±0.74	72.45±0.73	76.51±0.71
	quickdraw	76.59±0.73	77.58±0.71	81.33±0.66
	fungi	53.11±1.14	54.67±1.13	64.86±1.12
	vgg_flower	89.97±0.56	90.60±0.56	91.25±0.57
目标域	MSTAR	77.10±0.88	82.72±0.69	88.61±0.81

本文提出的模型在 9 个数据集上的识别准确率综合比较如图 4 所示,可以清晰地看到,在源域、目标域大部分数据集上,本文模型的分类准确度都高于前两种方法。

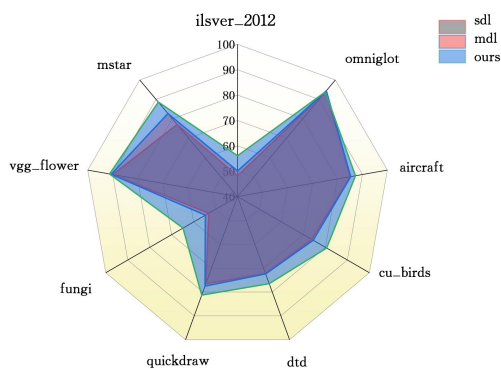


图 4 Sdl、Mdl 和本文方法在不同数据集上的表现

Fig. 4 Results of Sdl, Mdl and the proposed method on different datasets

3.4.2 MSTAR 数据集上小样本学习的准确度

该部分实验主要采取两种小样本的设置: K -way 5-shot 和 5-way 1-shot, 在 MSTAR 数据集上进行验证。对于每个任务, 每个类别只抽取 5 个或 1 个样本作为支持集, 从而验证模型在跨域小样本条件下对目标域分类的能力, 最终随机抽取 600 个任务, 取 95% 的置信度, 并采用经典小样本学习方法 Matching Network, MAML, Prototypical Network, Relation Network 和 Meta-Opt 作为对比方法进行测试。表 3

于 SAR 图像目标识别的相关文献没有采用与本文同样的方法, 因此这里主要与其他可行的方法进行对比: 1) 最好的单域网络(Single Domain Model), 即将 8 个源域中分类效果最好的网络直接用在测试中; 2) 采用 2.1 节所述的原始多域网络(Multi-Domain Model)进行测试。如表 2 所列, 所提方法在 7 个源域和 1 个目标域上都优于其余方法, 在 ImageNet, Aircraft, Birds, Texture, Quickdraw, Fungi, Flower 上分类准确度分别提升 7.92% 和 5.98%、2.08% 和 1.88%、6.83% 和 6.11%、4.3% 和 4.06%、4.74% 和 3.75%、11.75% 和 10.19%、1.28% 和 0.65%, 在 MSTAR 上提升 11.51% 和 5.89%, 证明通过第 2 章设计的模型具备同时从已见过的和未见过的域中提取特征的能力, 说明了从丰富的光学图像源域中将学习到的知识迁移到跨度较大的 SAR 图像目标域的可行性, 并且所提方法针对 SAR 图像特点所进行的改进是有效的。

说明, 在未考虑跨域因素影响下的的小样本学习方法无法获得较好的结果, 相比较为先进的 Meta-Opt 方法, 本文方法在 5-shot 和 5-way 1-shot 下准确度分别提高了 7.12% 和 1.15%。

表 3 MSTAR 数据集上的小样本学习结果比较

Table 3 Comparison of few shot learning methods on MSTAR

dataset		
对比方法	K-way 5-shot	5-way 1-shot
Matching Network	56.83	48.28
MAML	55.37	47.59
Prototypical Network	59.92	46.82
Relation Network	58.67	47.19
Meta-Opt	62.51	49.18
Ours	69.63	50.33

3.4.3 元测试阶段不同分类器带来的性能差异

根据 2.2 节, 在元测试阶段的距离度量函数可以使用最近邻分类器(NCC)、支持向量机(SVM)、逻辑回归(LR)、Mahalanobis 距离^[15](SCM)或 K 近邻(KNN)的分类方法, 以及搭配使用 L2 范数或余弦相似度计算特征向量之间相似度。从表 4 中可以发现, 对于 ImageNet 的图像而言, NCC+COS 的组合方式效果更好; 而对于 Omniglot 手写字符而言, SVM+COS 的分类准确度更高; 对于剩下大部分细粒度分类数据集而言, 加入 L2 正则化的 SVM+COS 分类性能更高; 对于 SAR 图像目标域, 则只有 NCC+L2 的组合获得了最好的效果。因此若将本文所提方法应用于 SAR 图像目标识别, 应采用 NCC+L2 的分类组合。

表 4 选择不同距离度量函数和特征相似度的结果比较

Table 4 Comparison of results with different distance metric functions and feature similarities

数据集	Ncc+Cos	Svm+Cos	Svm(L2)+Cos	Lr+Cos	Scm+Cos	Knn+Cos	Ncc+L2
ilsvrc_2012	57.93	57.64	56.89	53.84	52.17	49.17	56.08
omniglot	94.32	94.58	94.18	93.67	93.24	93.18	93.84
aircraft	88.14	88.19	88.25	88.22	86.54	85.37	87.23
cu_birds	80.61	80.32	80.40	79.06	77.35	73.35	80.58
dtd	76.38	76.44	76.33	74.64	74.36	72.05	76.51
quickdraw	82.38	81.81	83.07	80.42	81.43	77.72	81.33
fungi	67.94	68.67	69.27	64.17	58.72	59.94	64.86
vgg_flower	91.88	92.09	91.97	91.30	90.47	88.67	91.25
MSTAR	87.44	86.96	87.24	86.28	87.28	79.46	88.61

3.4.4 元训练阶段两种蒸馏损失的变化

根据 2.1 节所述,知识蒸馏是整个训练过程中最重要的环节,从 8 个源域的特征提取网络中获取的通用网络的好坏决定了下一步在目标域适应的性能,因此这里分别对元测试阶段目标蒸馏和特征蒸馏的两种损失进行了记录分析。从图 5 中可以看出,迭代到 8000 次左右时,对于 8 个教师网络分别计算的目标蒸馏损失基本已经达到平稳状态。

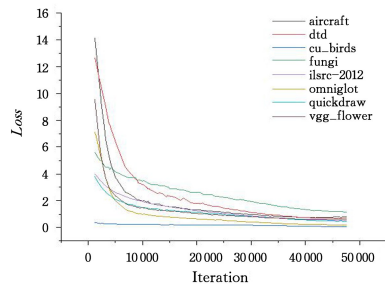


图 5 目标蒸馏损失变化图

Fig. 5 Loss of object distillation

从图 6 中可以看出,特征蒸馏损失在迭代到 45000 次左右才达到较平稳的状态,因此实验部分将迭代次数设置为 50000 次,能够获得性能较好的通用特征提取网络。

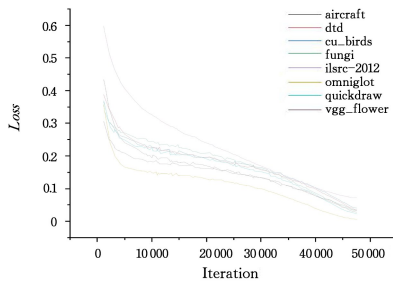


图 6 特征蒸馏损失变化图

Fig. 6 Loss of feature distillation

结束语 自跨域小样本学习的概念被提出以来,研究者们提出了很多不同的方法,但都很少涉及更具挑战性的、将分布差异更大的 SAR 图像作为目标域的情况;另外,针对 SAR 图像目标识别问题,极少有采用跨域小样本学习的方法来解决。本文针对以上两个挑战进行了部分尝试,论证了所提方法的可行性,克服了传统小样本学习要求源域和目标域分布一致的限制,大幅减少了网络参数数量,计算效率更高,为解决小样本的 SAR 图像目标识别问题提供了新的思路。

由于 SAR 图像与普通光学图像特性迥异,因此需要进一步研究如何通过设计合理的模型,有针对性地学习对 SAR 图像识别有帮助的信息,从而提高识别准确率。另外,MSTAR 数据集的目标类型有限,未来需要针对更多目标种类、更新的

SAR 数据集进行训练和测试。

参考文献

- [1] YING Z L, WANG W Q, XU Y, et al. Twin Self-Supervised Learning Method for Small Sample SAR Images Automatic Target Recognition[J/OL]. Signal Processing: 1-13. [2023-07-13]. <http://kns.cnki.net/kcms/detail/11.2406.TN.20230321.1926.010.html>.
- [2] FENG B D, YANG H T, WANG J N, et al. SAR Image Target Recognition Algorithm Based on Data Fusion[J]. Computer System Applications, 2022, 31(12): 342-349.
- [3] WANG Y Y. SAR Target Recognition Based on Modified Sparse Representation [J/OL]. Electronics Optics & Control: 1-7. [2023-08-21]. <http://kns.cnki.net/kcms/detail/41.1227.TN.20230727.1127.004.html>.
- [4] KANG Z Q, ZHANG S Q, FENG S J, et al. Sparse Prior-Guided CNN Learning for SAR Images Target Recognition[J]. Journal Of Signal Processing, 2023, 39(4): 737-750.
- [5] TANG H, LI Z, PENG Z, et al. Blockmix: meta regularization and self-calibrated inference for metric-based meta-learning [C]// Proceedings of the 28th ACM International Conference on Multimedia. 2020: 610-618.
- [6] ANTONIOU A, EDWARDS H, STORKEY A. How to train your MAML[J]. arXiv: 1810.09502, 2018.
- [7] SUNG F, YANG Y, ZHANG L, et al. Learning to compare: Relation network for few-shot learning [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 1199-1208.
- [8] XIE Y, FU Y, TAI Y, et al. Learning To Memorize Feature Hallucination for One-Shot Image Generation [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 9130-9139.
- [9] SCHROEDER B, CUI Y. Fgvex fungi classification challenge 2018[J/OL]. http://github.com/visipedia/fgvex_fungi_comp.
- [10] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks [C]// International Conference on Machine Learning. PMLR, 2017: 1126-1135.
- [11] SNELL J, SWERSKY K, ZEMEL R. Prototypical networks for few-shot learning [C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 4080-4090.
- [12] CHEN W Y, LIU Y C, KIRA Z, et al. A closer look at few-shot classification[J]. arXiv: 1904.04232, 2019.
- [13] REQUEIMA J, GORDON J, BRONSKILL J, et al. Fast and flexible multi-task classification using conditional neural adap-

- tive processes[C]// Proceedings of 33rd International Conference on Neural Information Processing Systems. 2019:7959-7970.
- [14] BATENI P,GOYAL R,MASRANI V,et al. Improved few-shot visual classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:14493-14502.
- [15] PEREZ E,STRUB F,DE VRIES H,et al. Film: Visual reasoning with a general conditioning layer[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2018.
- [16] DVORNIK N,SCHMID C,MAIRAL J. Selecting relevant features from a multi-domain representation for few-shot classification[C]// Computer Vision—ECCV 2020,16th European Conference, Glasgow, UK, Part X 16. Springer International Publishing,2020:769-786.
- [17] LIU L,HAMILTON W,LONG G,et al. A universal representation transformer layer for few-shot image classification[J]. arXiv:2006.11702,2020.
- [18] PHOO C P,HARIHARAN B. Self-training for few-shot transfer across extreme task differences[J]. arXiv:2010.07734,2020.
- [19] FU Y,XIE Y,FU Y,et al. Me-d2n: Multi-expert domain decomposition network for cross-domain few-shot learning[C]// Proceedings of the 30th ACM International Conference on Multimedia. 2022:6609-6617.
- [20] LI W H,LIU X,BILEN H. Universal representation learning from multiple domains for few-shot classification[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021:9526-9535.
- [21] GUO Y,CODELLA N C,KARLINSKY L,et al. A broader study of cross-domain few-shot learning[C]// Computer Vision—ECCV 2020,16th European Conference,Glasgow,UK,Part XX-VII 16. Springer International Publishing,2020:124-141.
- [22] YANG Y,ZHU W G,QU L L,et al. A Survey of Research on the Target Recognition via Limited SAR Sample Based on Transfer Learning. [J/OL]. Electronics Optics & Control;1-8. [2023-07-13]. <http://kns.cnki.net/kcms/detail/41.1227.TN.20230327.1159.002.html>.
- [23] CHEN W Y,LIU Y C,KIRA Z,et al. A closer look at few-shot classification[J]. arXiv:1904.04232,2019.
- [24] KENDALL A,GAL Y,CIPOLLA R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:7482-7491.
- [25] ZACARIAS A,ALEXANDRE L A. Sena-cnn: Overcoming catastrophic forgetting in convolutional neural networks by selective network augmentation[C]// IAPR Workshop on Artificial Neural Networks in Pattern Recognition. Cham: Springer International Publishing,2018:102-112.
- [26] ROY D,PANDA P,ROY K. Tree-CNN: a hierarchical deep convolutional neural network for incremental learning[J]. Neural Networks,2020,121:148-160.
- [27] XU H,ZHI S,SUN S,et al. Deep Learning for Cross-Domain Few-Shot Visual Recognition: A Survey[J]. arXiv:2303.08557,2023.
- [28] HINTON G,VINYALS O,DEAN J. Distilling the knowledge in a neural network[J]. arXiv:1503.02531,2015.
- [29] ROMERO A,BALLAS N,KAHOU S E,et al. Fitnets: Hints for thin deep nets[J]. arXiv:1412.6550,2014.
- [30] LI W H,BILEN H. Knowledge distillation for multi-task learning[C]// Computer Vision—ECCV 2020 Workshops,Glasgow, UK,Part VI 16. Springer International Publishing,2020:163-176.
- [31] KORNBLITH S,NOROUZI M,LEE H,et al. Similarity of neural network representations revisited[C]// International Conference on Machine Learning. PMLR,2019:3519-3529.
- [32] NGUYEN T,RAGHU M,KORNBLITH S. Do wide and deep networks learn the same things? uncovering how neural network representations vary with width and depth[J]. arXiv:2010.15327,2020.
- [33] TRIANTAFILLOU E,ZHU T,DUMOULIN V,et al. Meta-dataset: A dataset of datasets for learning to learn from few examples[J]. arXiv:1903.03096,2019.
- [34] RUSSAKOVSKY O,DENG J,SU H,et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision,2015,115:211-252.
- [35] LAKE B M,SALAKHUTDINOV R,TENENBAUM J B. Human-level concept learning through probabilistic program induction[J]. Science,2015,350(6266):1332-1338.
- [36] MAJI S,RAHTU E,KANNALA J,et al. Fine-grained visual classification of aircraft[J]. arXiv:1306.5151,2013.
- [37] WAH C,BRANSON S,WELINDER P,et al. The caltech-ucsd birds-200-2011 dataset [J/OL]. <https://www.docin.com/p-1472255882.html>.
- [38] CIMPOI M,MAJI S,KOKKINOS I,et al. Describing textures in the wild[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014:3606-3613.
- [39] JONGEJAN J,ROWLEY H,KAWASHIMA T,et al. The quick, draw!—AI experiment[J]. Mount View,2016,17(2018):4.
- [40] HOUBEN S,STALLKAMP J,SALMEN J,et al. Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark [C]// The 2013 International Joint Conference on Neural Networks(IJCNN). IEEE,2013:1-8.
- [41] LIN T Y,MAIRE M,BELONGIE S,et al. Microsoft coco: Common objects in context[C]// Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, Part V 13. Springer International Publishing,2014:740-755.



SHI Songhao, born in 1994, postgraduate. His main research interests include few shot learning, target recognition, etc.



WANG Xiaodan, born in 1966, Ph. D., professor, Ph. D supervisor. Her main research interests include intelligent information processing and target recognition.