

基于改进关系网络的小样本红外空中目标分类方法

金璐^{1,2,3}, 刘士建^{1,3}, 王霄^{1,2,3}, 李范鸣^{1,3*}

¹中国科学院上海技术物理研究所, 上海 200083;

²中国科学院大学, 北京 100049;

³中国科学院红外探测与成像技术重点实验室, 上海 200083

摘要 针对地对空红外空中目标识别任务中数据量严重不足的问题, 提出一种基于改进关系网络的小样本红外空中目标分类方法。该方法将关系网络模型、多尺度特征融合方法及元学习训练策略相结合, 首先构造多尺度特征提取模块提取输入图像的特征向量, 然后将支撑样本和预测样本的特征向量输入到关系模块中, 根据关系值得到预测样本的类别标签。mini-ImageNet 数据集上的实验结果表明: 所提模型的分类精度显著高于其他经典的小样本学习模型。Infra-aircraft dataset 上的实验结果表明: 所提方法在仅有个位数样本的情况下, 可完成多种机型的地对空红外图像分类任务。

关键词 成像系统; 红外图像; 空中目标分类; 小样本学习; 元学习

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/AOS202040.0811005

Infrared Aircraft Classification Method with Small Samples Based on Improved Relation Network

Jin Lu^{1,2,3}, Liu Shijian^{1,3}, Wang Xiao^{1,2,3}, Li Fanming^{1,3*}

¹Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

²University of Chinese Academy of Sciences, Beijing 100049, China;

³Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China

Abstract To resolve the problem that the available data on the ground-to-air infrared aircraft identification task is considerably scarce, the small samples infrared aircraft identification classification method is proposed on the basis of an improved relation network. This method combines the relation network model and the multi-scale feature fused method with the meta learning training strategy. First, a multi-scale feature extraction module is constructed to extract the feature tensors of input images. Then, the feature tensors of support samples and test samples are inputted into the relation module, and the category labels corresponding to test samples are predicted based on the relation value. The results of the proposed model on the mini-ImageNet dataset show that the classification accuracy of the proposed model is significantly higher than those of other conventional learning models using small samples. The experimental results based on the Infra-aircraft dataset verify that the proposed model can realize the ground-to-air infrared image classification task of various aircraft types even when the number of samples is limited.

Key words imaging systems; infrared image; aircraft classification; few-shot learning; meta learning

OCIS codes 110.3080; 100.5010; 150.1135

1 引 言

红外空中目标识别^[1]是电子搜索领域的重要组成部分。目前已有的目标识别方法可分为手动设计和深度学习两大方向。传统的目标识别方法强烈依

赖设计者的专业知识和数据的分布特征, 很难充分挖掘数据内部的内在关联。近几年来, 基于深度学习的人工智能技术发展迅猛, 这类方法可以通过学习的策略对数据提取出鲁棒性更高、表达性更强的特征。目前, 深度学习技术已在特定的工业领域得

收稿日期: 2019-12-03; 修回日期: 2019-12-29; 录用日期: 2020-01-14

基金项目: 国家十三五国防预研项目(Jzx2016-0404/Y72-2)、上海市现场物证重点实验室基金(2017xcwzk08)

* E-mail: lfmjws@163.com

到广泛应用,但是在地对空红外空中目标识别的问题上不具有迁移性,根本原因在于深度模型依赖大量的数据样本。在地对空红外空中目标的识别问题上,样本的获取代价昂贵,某些少见机型的样本数据仅为个位数,远远达不到深度学习对训练数据的要求。如果强行在小数据集上训练,则会产生严重的过拟合问题。

为了应对训练数据不足的问题,将当前的主流方法分为两个方向。其中数据增强技术^[2-4],如图像的几何变换、色彩空间变换、随机噪声叠加等,可以依托深度学习领域中的生成模型去仿照原始数据的分布特征来产生新的数据^[5-7]。然而当数据样本仅有几十个甚至几个时,此类方法依然无法构造出足够的数据样本。在这种情况下,小样本学习^[8]应运而生。小样本学习的具体任务为,对于训练过程中没有出现过的类别,即使仅给出极少量的标注样本,模型也能够正确识别。该方法不仅在电子侦察领域具有广阔的应用前景,在工业或医疗领域同样具有巨大的应用价值,如大型设备的故障检测^[9]和医疗辅助诊断中的罕见病识别^[10]等。

目前主流的小样本学习任务的解决方案包括元学习^[11-12]和度量学习^[13-14]。其中,元学习建立在多个相似的学习任务上,学习类别变化情况下模型的泛化能力。在训练过程中,每个训练周期通过采样得到不同的元任务,进而训练不同的类别组合,这种机制使得模型能够从不同的元任务中学到它们的共性部分。在测试阶段,即使面对训练过程中从未见过的数据类别,元学习也能较好地完成任务。而度量学习的思路更加简单高效,首先通过嵌入网络学习样本的嵌入向量,然后在嵌入空间中通过求解最近邻的方式得到预测物体的类别标签。在度量学习方法中,比较有代表性的研究成果包括匹配网络模型^[13],原型网络模型^[14]及关系网络模型^[15]。在匹配网络和原型网络模型中,嵌入向量

之间的距离度量采用手动设计的方法,如欧氏距离或余弦距离。而在关系网络模型^[15]中,除了通过学习的方式将训练数据映射到嵌入空间中,距离的度量方式也得到了改进,通过神经网络来学习嵌入向量间的距离度量模式,是一种端到端的学习策略。

针对红外空中目标识别任务中数据量严重不足的问题,本文提出一种基于改进关系网络的小样本红外空中目标分类方法。该方法将关系网络模型、多尺度特征融合方法及元学习训练策略相结合,实现了少量样本条件下对红外空中目标的有效分类。受关系网络模型的启发,首先构造嵌入模块来提取每一幅图像的特征信息,然后把待测试的图像特征和训练样本的图像特征连接起来并输入到关系模块中,最后根据关系值输出待测试图像的类别标签。在嵌入模块部分,对关系网络模型^[15]进行改进,使用不同尺寸的卷积核提取多尺度特征。在训练策略上使用元学习的训练策略,通过计算各个元任务的合成梯度来更新网络参数。如果把针对红外空中目标的分类任务作为一个具体的元任务,则需要构造大量类似的元任务集来对模型进行训练,而训练集可以利用易于获取的地面目标来构造,因此,可以在仅有少量样本的情况下实现地对空红外空中目标的类别判断。

2 基本原理

2.1 改进关系网络的系统模型

使用的网络框架如图1所示,整体框架可分为嵌入模块和关系模块两个部分。其中,嵌入模块可视为特征提取部分,它将元任务中的支撑样本和测试样本映射到某个特征空间,得到二者的嵌入向量。关系模块可视为距离度量部分,通过学习到的距离度量函数计算出支撑样本和测试样本对应的嵌入向量之间的距离。

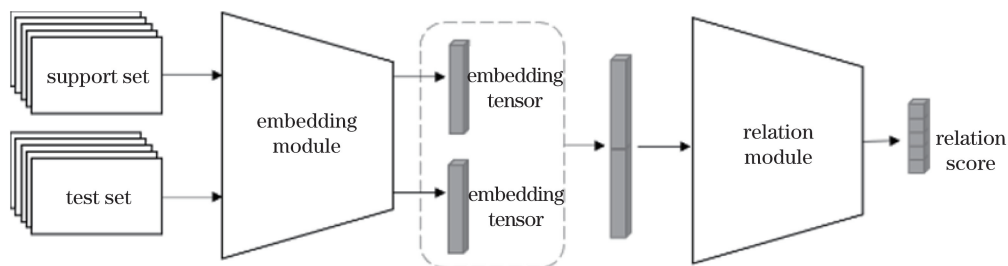


图1 小样本红外空中目标分类学习模型的整体架构

Fig. 1 Overall architecture of infrared aircraft classification learning model with small samples

在红外图像中,热源部位与目标全局关系紧密,在非热源区域,纹理特征较为明显。基于此

特性,受 InceptionNet^[16]的启发改进关系网络,设计图 2(a)所示的嵌入模块。在该结构中,通过设置不同尺寸的卷积核来得到多尺度的特征信息,进一步提高了特征的鲁棒性。其中,大尺寸卷积

核有助于捕获热源部位的全局信息,而小尺寸卷积核有助于获取非热源区域的纹理细节,由此得到的多尺度融合特征更加适用于红外图像的分类。

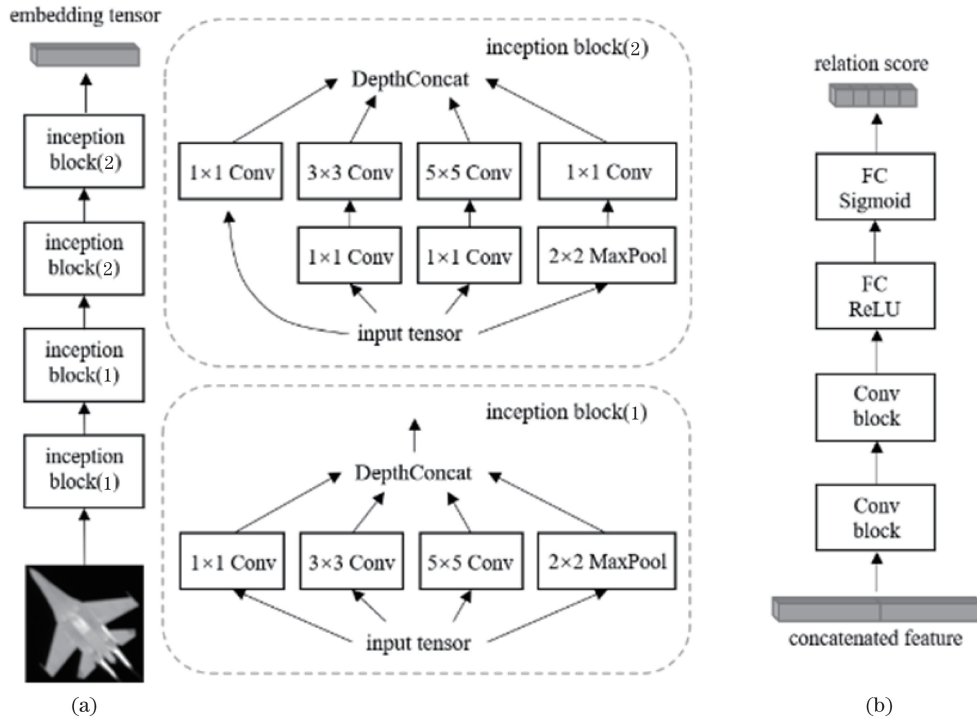


图 2 模块结构图。(a) 嵌入模块;(b) 关系模块

Fig. 2 Architecture of module. (a) Embedding module; (b) relation module

从图 2(a)可以看到,改进后的嵌入模块共包含四个卷积单元,每个卷积单元由三个不同尺度的卷积核及最大池化层并联而成。不同尺寸的卷积核提取的特征具有不同大小的感受野,卷积核的尺寸越大,感受野也越大,但是过大的感受野会忽略图像上的细节信息。为了探究卷积核尺寸与分类效果之间的关系,设计表 1 所示的 8 组对比实验,以 C -way K -shot 问题表示,其中, C 表示类别数量, K 表示每一个类别里的样本数量。在实验中,选用文献[15]提出的关系网络结构,以红外地面目标数据集作为训练集,仅改变嵌入模块中的卷积核大小,记录模型对红外空中目标的分类准确率。

表 1 卷积核尺寸评估的准确率

Table 1 Accuracy of convolution kernel size estimation %		
Convolution kernel size	5-way 1-shot	5-way 5-shot
3×3	81.23	91.28
5×5	81.41	90.92
7×7	79.33	87.07
9×9	75.98	82.30

从表 1 可知,当卷积核大小为 3×3 或 5×5 时,

红外空中目标的分类准确率相差不大,当卷积核尺寸继续增大至 7×7 和 9×9 时,准确率明显下降。因此改进后的嵌入模块选取尺寸为 3×3 和 5×5 的卷积核,同时加入 1×1 尺寸的卷积核,以更改通道数量,实现跨特征层的级联作用。但是并行分支会大大增加参数量和计算复杂度,因此在后两个卷积单元中加入 1×1 卷积,一方面 1×1 卷积可以减少通道数量,降低计算成本,另一方面 1×1 卷积可以起到线性修正的作用。

关系模块如图 2(b)所示,包含两个卷积单元和两个全连接层。两个卷积单元各包含 64 个 3×3 卷积层,其后接 ReLU 激活函数和 2×2 最大池化层,第一个全连接层后接 ReLU 激活函数,第二个全连接层后接 Sigmoid 激活函数。训练之前,将数据集分为两个部分,分别是用于泛化性能学习的元训练集 D_{train} 和用于测试的元测试集 D_{test} 。为了充分验证模型的泛化性能, D_{train} 和 D_{test} 所包含的类别不交集。

$$D = \langle D_{train}, D_{test} \rangle. \quad (1)$$

训练时,每个训练周期由若干个元任务构成,每

个元任务中的数据分为支撑集和预测集。对于C-way K-shot 问题,支撑集表示为 $S = \{(x_i, y_i)\}_{i=1}^m$, 其中 $m = K \times C$, x_i 表示数据, y_i 表示 x_i 对应的标签; 预测集表示为 $P = \{(x_j, y_j)\}_{j=1}^n$, 其中, n 表示预测集中样本数量, x_j 表示数据, y_j 表示 x_j 对应的标签。假设嵌入模块的映射函数为 f_φ , 关系模块的映射函数为 g_Φ , φ 和 Φ 分别为嵌入模块和关系模块的参数。对于每一个支撑数据 x_i 和预测数据 x_j , 它们通过嵌入模块 f_φ 后产生的嵌入向量分别为 $f_\varphi(x_i)$ 和 $f_\varphi(x_j)$, 将其首尾相连, 得到连接向量 $\text{Concat}[f_\varphi(x_i), f_\varphi(x_j)]$ 。然后将连接向量输入到关系模块 g_Φ 中, 得到 x_i 和 x_j 之间的关系值 $r_{i,j}$ 为

$$r_{i,j} = g_\Phi\{\text{Concat}[f_\varphi(x_i), f_\varphi(x_j)]\}, \quad i = 1, 2, \dots, C \times K. \quad (2)$$

经过 Sigmoid 函数后输出的关系值 $r_{i,j}$ 在区间 $(0, 1)$ 内, 对于同一类别的支撑样本和预测样本, 假设它们之间的关系值为 1, 对于不同类别的支撑样本和预测样本, 假设它们之间的关系值为 0。此时, 支撑样本和预测样本之间的关系值可视为回归问题, 如果用方均误差 (MSE) 来计算损失值, 则损失函数可表示为

$$L = \begin{cases} \sum_{i=1}^m \sum_{j=1}^n (r_{i,j} - 1)^2, & y_i = y_j \\ \sum_{i=1}^m \sum_{j=1}^n r_{i,j}^2, & \text{otherwise} \end{cases}. \quad (3)$$

训练的目标函数为

$$\varphi, \Phi \leftarrow \begin{cases} \underset{\varphi, \Phi}{\text{argmin}} \sum_{i=1}^m \sum_{j=1}^n (r_{i,j} - 1)^2, & y_i = y_j \\ \underset{\varphi, \Phi}{\text{argmin}} \sum_{i=1}^m \sum_{j=1}^n r_{i,j}^2, & \text{otherwise} \end{cases}. \quad (4)$$

2.2 元学习训练策略

经典的机器学习训练模式可概括如下: 获取特定任务的大型数据集, 然后利用该数据集开始训练, 反复迭代更新模型参数, 从而训练出有效的模型。但是这种学习策略和人类的学习方式相差甚远, 因为人类往往能够利用以往的经验, 仅仅通过少量样本就能迅速完成学习过程。为了使模型获得更好的泛化能力, 采用元学习的训练策略, 该策略的核心思想是并行地启用一批任务, 通过计算获得各个任务的合成梯度来更新网络参数, 这种训练模式首先需要接触大量的训练任务, 然后在测试阶段只通过少量的标签数据来考验元学习器对新任务的适应能

力。在本文所用的模型中, 这种训练策略如图 3 所示。

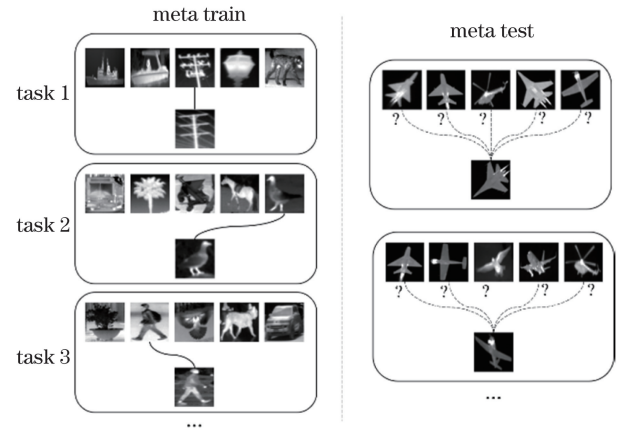


图 3 元学习训练模式下的少样本学习数据集示例

Fig. 3 Illustration of few-shot learning datasets under meta learning training mode

通常来说, 小样本学习被描述为 C-way K-shot 问题, 即每个元任务包含 C 个类别, 从每个类别中随机抽取 K 个样本, 共 $C \times K$ 个样本组成一个元任务中的支撑集; 再从这 C 个类别剩余的数据中随机抽取若干个样本作为此次元任务中的预测对象。图 3 描述的是一个典型的 5-way 1-shot 问题。在设计的学习模型中, 训练目标是找到合适的嵌入模块 f_φ 和关系模块 g_Φ , 为了描述方便, 下文将二者合并描述为参数为 θ 的学习器 f_θ 。训练之前, 从 D_{train} 中反复随机抽取大量的元任务 T_i 组成元任务集 T 作为训练集, 同理, 从 D_{test} 中反复随机抽取若干个元任务组成测试集。具体的学习算法伪代码描述如图 4 所示。

与经典的机器学习训练方式相比, 这种元学习训练策略并非直接使用所有训练样本的损失函数对参数进行更新, 而是使用单个元任务中的数据样本计算损失函数, 并求出梯度方向作为模型梯度下降的大方向。这种学习策略的目标是每个元任务都“稍微学一点”, 以找到合适的特征映射方式和特征的距离度量函数, 使得预测对象和支撑集中的同类数据的距离尽可能相近, 同时使预测对象和支撑集中的异类数据的距离尽可能远。

3 实验分析与讨论

3.1 实验环境和数据来源

实验均在 Intel (R) Core (TM) i7-6700HQ, 2.60 GHz, 64 GB 内存、NVIDIA GeForce GTX 1080 显卡、Ubuntu16.04 操作系统上, 通过 PyTorch 深度学

algorithm: meta-learning for improved relation network

```

input: training set  $D_{\text{train}}$ , testing set  $D_{\text{test}}$ , learning rate  $\alpha$ , episodetrain denotes training episode, episodetest denotes testing episode, best accuracy acc_best=0.
randomly initialize  $\theta$ 
while not done do
  for episode=1 to episodetrain do
    sample batch of tasks  $T = \{T_1, T_2, \dots, T_N\}$  from  $D_{\text{train}}$ 
    for  $T_i \in T$  do
      compute loss  $L$  for meta task  $T_i$  in equation (3)
      evaluate  $\nabla_{\theta} L$ 
      update parameters  $\theta \leftarrow \theta - \alpha \nabla_{\theta} L$ 
    end for
  end for
  for episode=1 to episodetest do
    Sample batch of tasks  $T_i$  from  $D_{\text{test}}$ 
    Evaluate accuracy acc $T_i$ 
  end for
  evaluate the average accuracy acc_avg =  $\frac{1}{\text{episode}_{\text{test}}} \sum_{i=1}^{\text{episode}_{\text{test}}} \text{acc}_{T_i}$ 
  if acc_avg > acc_best do
    update acc_best  $\leftarrow$  acc_avg
    save model parameters  $\theta$ 
  end if
end while

```

图 4 学习算法的伪代码

Fig. 4 Pseudo-code for learning algorithm

习框架实现实验。在训练时,每一个训练周期都是基于 C -way K -shot 的元任务进行的,即从训练数据中随机抽取 C 类,将每类 K 个标注样本组成支撑集,然后基于该支撑集输出预测样本的类别标签。测试阶段同样是基于 C -way K -shot 的元任务进行的。

共使用 3 个数据集进行验证,分别为 mini-

ImageNet 数据集、红外地面目标数据集 (Infra-object dataset)、地对空红外空中目标数据集 (Infra-aircraft dataset)。其中,mini-ImageNet 数据集在匹配网络^[13]中首次被提出,如今已成为小样本学习的一个基准数据集。三个数据集的部分图像示例如图 5 所示。

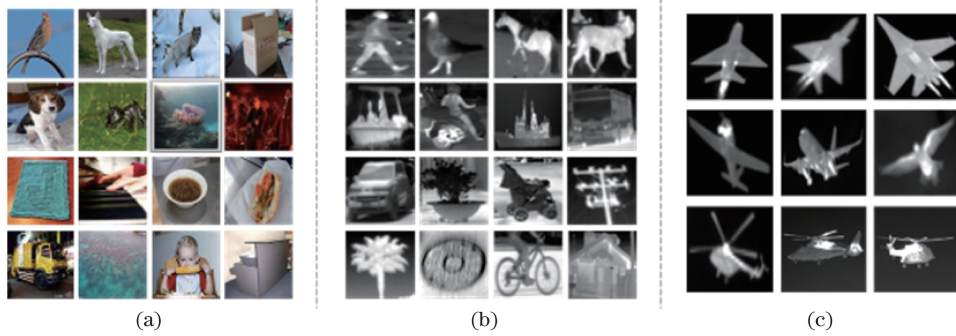


图 5 3 个数据集的部分示例。(a) mini-ImageNet dataset; (b) Infra-object dataset; (c) Infra-aircraft dataset

Fig. 5 Partial examples of three datasets. (a) mini-ImageNet dataset; (b) Infra-object dataset; (c) Infra-aircraft dataset

在所有实验组中,每个元任务除了包含 $C \times K$ 个支撑样本外,还包含 15 个预测样本。所有实验均使用 Adam 优化算法^[17]对模型进行参数更新,初始学习率设为 0.001。所有输入图像均被调整为 $84 \text{ pixel} \times 84 \text{ pixel}$ 的灰度图像,且分别旋转 90° 、 180° 、 270° ,以增强样本数量。每训练 500 个周期,就从测试集中随机抽取 1000 个元任务,验证模型对小样本数据的分类准确率。假设 1000 个元任务中

预测标签正确的测试样本总数为 N_{right} ,则准确率的计算公式为

$$E_{\text{acc}} = \frac{N_{\text{right}}}{1000 \times 15} \quad (5)$$

3.2 mini-ImageNet 数据集的小样本学习任务

Mini-ImageNet 数据集为 2012 年 ImageNet 大规模视觉识别的竞赛版本,该数据集包含从 ILSVRC-2012^[18]数据集中挑选的 100 类图像数据,

每类包含 600 幅图像, 共计 60000 幅图像数据。本文延续文献[19]的数据集制作方式, 使用其中 64 类作为训练集, 16 类作为验证集, 20 类作为测试集, 取 10 次实验结果并求波动范围, 并将结果与最近邻基准网络^[19]、线性基准网络^[20]、LSTM 模型元学习^[19]、未知模型元学习^[21]、匹配网络^[13]、原型网络^[14]、关系网络^[15]等近两年小样本学习领域的主流方法进行对比, 结果如表 2 所示。

表 2 Mini-ImageNet 数据集上不同模型的准确率

Table 2 Accuracy of each model on mini-ImageNet dataset %

Model	Fine-tune	5-way	
		1-shot	5-shot
Baseline-nearest-neighbor ^[19]	N	41.08±0.70	51.04±0.65
Baseline-linear ^[20]	Y	42.11±0.71	62.53±0.69
Meta-learner LSTM ^[19]	N	43.44±0.77	60.60±0.71
MAML ^[21]	Y	48.70±1.84	63.11±0.92
Matching network ^[13]	Y	42.40	58.00
Prototypical network ^[14]	F	49.42±0.78	68.20±0.66
RelationNet ^[15]	F	50.44±0.82	65.32±0.70
Improved relation network	F	54.89±1.02	69.87±0.75

从表 2 可以看出, 所提模型的准确率较其他经典的网络结构有一定幅度的提升。图 6 为改进关系网络与关系网络前 60000 次迭代的测试准确率和损失曲线。从图 6 可以看出, 与关系网络相比, 嵌入模块后的改进关系网络在 mini-ImageNet 上的分类效果显著提升。这是因为文献[15]提出的嵌入模块仅使用单一尺寸的卷积核对样本提取特征信息, 简单的 4 层卷积神经网络(CNN)结构运

算复杂度较低, 效果也相对较差; 而 mini-ImageNet 取自 ImageNet 数据集, 其数据分布比较复杂, 即便是同一类别的图像样本, 目标在图像中所占的位置和区域也相差较大, 在这种条件下, 使用本文设计的多尺度卷积核提取的特征具有更高的鲁棒性。

但是在红外空中目标识别任务中, 红外图像的样本数量远小于 mini-ImageNet 的数据总量, 即便使用红外地面目标数据集作为训练集, 样本的类别数量也只有 mini-ImageNet 训练类别的 1/3 左右。为了探究训练样本的充分度与模型分类能力之间的关系, 使用改进后的关系网络模型, 保持 mini-ImageNet 的测试集不变, 仅改变训练集中的类别数量和每一类中包含的样本数量, 记录训练过程, 绘制图 7 所示的准确率变化曲线。从 mini-ImageNet 训练集(64 类)中随机抽取相应的类别数量作为训练集, 当每一类保持原有的 600 个训练样本不变时, 前 60000 次迭代的测试准确率如图 7(a)、(b)所示。可以看到, 随着训练类别数量的增加, 模型的分准确率明显呈上升趋势, 但是当训练类别仅有 16 类时, 随着迭代次数超过 30000, 模型逐渐进入过拟合状态, 此时应提前终止训练过程。从 mini-ImageNet 训练集的每一类中随机抽取相应的样本数量作为训练集, 当类别总数仍保持 64 类时, 前 60000 次迭代的测试准确率如图 7(c)、(d)所示。可以看到, 当每一类仅包含 100 个训练样本时, 模型较早地进入到过拟合状态, 但是当训练样本超过 200 后, 模型的分能力对训练集中每一类的样本数量变化并不敏感。

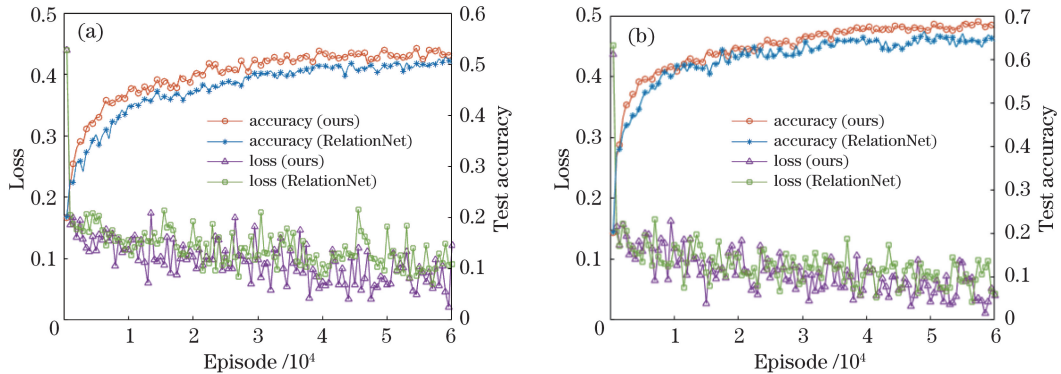


图 6 Mini-ImageNet 数据集上的损失曲线和测试准确率。(a) 5-way 1-shot; (b) 5-way 5-shot

Fig. 6 Test accuracy and loss curves on mini-ImageNet dataset.

(a) 5-way 1-shot; (b) 5-way 5-shot

进一步分析原因, 在元学习过程中, 将单个元任务的梯度方向作为模型参数更新的方向, 训练集类

别越多, 元任务集的分布越广, 不同元任务之间的梯度差异性也越大, 因此训练出的模型对新类别的泛

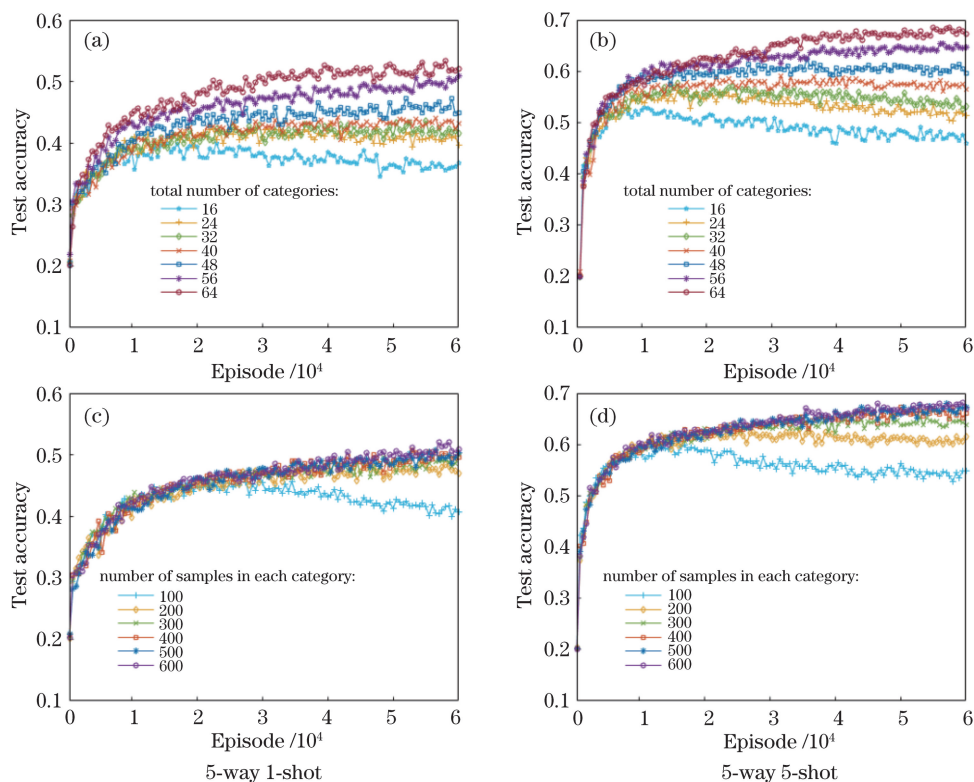


图 7 mini-ImageNet 数据集上测试准确率与训练样本充分度之间的关系

Fig. 7 Relationship between test accuracy and adequacy of training samples on mini-ImageNet dataset

化能力也越强。但是对每一类中的样本数量而言,只要该标签下的样本具有足够的代表性,能够有效覆盖样本空间,即使增加每一类中的样本数量,通过抽样得到的元任务集的分布也不会明显变广。因此,相比于每一类中的样本数量,训练集类别数量应具有更高的优先级。

3.3 Infra-aircraft 数据集的小样本学习任务

Infra-object dataset 共包含 20 类地面目标,每类包含 100~200 个 16 位灰度图像,图像类别包括房屋、行人、汽车、动物等。Infra-aircraft dataset 共包含 9 类,每类包含 100~200 个 16 位灰度图像数据,图像类别包含直 8、直 9、歼 7、歼 8、歼 11、教练机等。Infra-aircraft dataset 中的类别数目只有 9 类,不足以构成足够多样的元训练集,所以利用 mini-ImageNet dataset 和 Infra-object dataset 对模型进行训练,同时将 Infra-aircraft dataset 作为测试集,模拟验证真实场景下模型对少量红外空中目标的学习能力。然而这种训练策略有一个非常重要的前提条件——训练数据和测试数据在某些域间的分布应尽可能相似。如果二者的分布相差较大,将无法保证训练得到的特征映射函数和关系度量函数在测试任务集上是有效且稳定的。该理论涉及到领域自适应^[22-23]问题,受篇幅所限,本文对领域自适应问题不

进行更为深入的探讨,仅设置对照组实验讨论方法的合理性。在对照组实验中,将 Infra-aircraft dataset 与训练集组合成一个大数据集并打乱顺序,然后从该数据集中随机挑选 9 类作为测试集,剩余类别作为训练集。考虑到 mini-ImageNet 中的种类远多于 Infra-aircraft dataset,且可见光图像与红外图像之间存在较大的差异性,故对照组实验只将 Infra-object dataset 和 Infra-aircraft dataset 进行组合并打乱顺序。

综上所述,从训练集来源的角度看,共设 4 组实验。第 1 组:将 Infra-object dataset 和 Infra-aircraft dataset 组合成一个数据集并打乱顺序,从组合数据集中随机挑选 9 类作为测试集,剩余类别作为训练集,为了排除随机抽样的偶然性,本实验将重复进行 10 次。第 2 组:训练集为 Infra-object dataset,测试集为 Infra-aircraft dataset。第 3 组:训练集为 mini-ImageNet,测试集为 Infra-aircraft dataset。第 4 组:训练集为 mini-ImageNet 和 Infra-object dataset,首先在 mini-ImageNet dataset 上进行预训练,然后保留网络的权重参数作为初始值,利用 Infra-object dataset 对网络进行再训练,测试集依然为 Infra-aircraft dataset。同时,为了验证所提方法对于任务的多样性,分别设置了 5-way 1-shot 实验组、5-way

5-shot 实验组、8-way 1-shot 实验组及 8-way 5-shot 实验组。

实验结果如图 8 和表 3 所示。从表 3 可知,第 1 组的分类准确率略高于第 2 组。理想情况下,希望两组实验的结果完全相同,然而在现实场景中,地面目标的数据分布和空中目标会存在一定的差异。同时,在红外空中目标的分类任务中,歼 7、歼 8、歼 11 均属于喷气式飞机,直 5、直 8、直 9 同属于直升机,也就是说,在第 2 组实验中,测试任务集的分类粒度相比训练集更为精细。这些问题造成了两组实验在结果上的些许差别,也是下一步研究工作的重点。在第 3 组的训练过程中,使用可见光图像为训练数据,测试过程使用红外图像。这种方式的合理性在于,可见光图像的场景更加复杂,纹理细节更加丰富,而地对空红外图像的背景相对简单,纹理特征相对较少,形状特征更加明显,所以即使用可见光图像作为训练集,第 3 组实验的准确率结果仍全部大于 70%。但是对比第 2、3、4 组实验可以看到,使用红外图像作为训练数据对结果的影响更大。虽然模型能够通过可见光图像学到一些简单或复杂的形

状特征^[24],但红外图像中热源部位(如发动机、尾焰等)所存在的隐性特征是可见光图像不具备的。在元任务里,随着类别数量的增加,任务难度呈增大趋势,当类别数量一定时,支撑集数量越少,学习难度越大。在 4 组任务中,8-way 1-shot 的结果相对较低,且偏差最大;而在 5-way 5-shot 任务中,4 组实验的准确率均大于 90%。

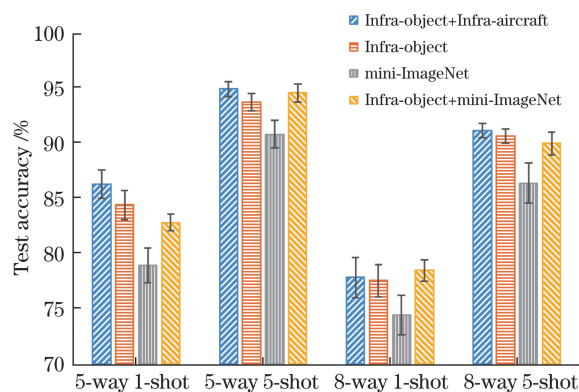


图 8 地对空红外空中目标分类的准确率比较
Fig. 8 Accuracy comparison of ground to air infrared aircraft classification

表 3 在不同训练数据集下,模型对红外空中目标分类的准确率

Table 3 Accuracy of model for infrared aircraft classification on different training datasets

%

Group	Training dataset	Test dataset	5-way 1-shot	5-way 5-shot	8-way 1-shot	8-way 5-shot
1	Infra-object + Infra-aircraft	Infra-object + Infra-aircraft	86.25 ± 1.25	94.84 ± 0.66	77.82 ± 1.79	91.11 ± 0.63
2	Infra-object	Infra-aircraft	84.37 ± 1.31	93.66 ± 0.76	77.56 ± 1.46	90.58 ± 0.64
3	mini-ImageNet	Infra-aircraft	78.92 ± 2.78	90.76 ± 1.24	74.44 ± 3.28	86.34 ± 1.95
4	Infra-object + mini-ImageNet	Infra-aircraft	82.79 ± 0.75	94.51 ± 0.82	78.47 ± 0.94	89.92 ± 1.02

4 结 论

针对地对空红外空中目标数据量严重不足的问题,提出一种基于改进关系网络的小样本红外空中目标分类方法。该方法将关系网络模型、多尺度特征融合方法及元学习训练策略相结合,首先构建多尺度特征提取模块提取支撑样本和预测样本的特征信息,然后把预测样本的图像特征和支撑样本的图像特征连接起来并输入到关系模块中,将预测样本和各类支撑样本之间的关系值作为类别的判断依据。即使某些类别的标注样本数量仅为个位数,该方法也能达到较好的分类效果。这种方法不仅在电子侦察领域具有广阔的应用前景,在一些仅有罕见样本的民用领域也有较大的应用潜力。但同时,这

种方法还存在一定的优化空间,面对具体的应用场景,仍有很多地方值得探索,具体体现在以下几个方面。

1) 领域漂移问题。小样本学习的提出源于标注样本的严重不足,但是训练过程却是需要大量标签样本的,这些标签样本可以通过一些公开的数据集获取。那么如何保证训练集和测试集在某些域间的分布尽可能相似,或者说,如何解决小样本学习中的领域自适应问题,是一个值得关注的方向。

2) 细粒度分类问题。目前的小样本学习模型大多基于 Omniglot 数据集和 mini-ImageNet 数据集来进行性能比对,其中 Omniglot 数据集在各大模型上的应用几乎进入瓶颈阶段,mini-ImageNet 数据集包含的数据种类跨度较大。然而在很多具体

的应用场景中,待分辨的类别间可能仅有细微差异,如本文面对的具体任务中,直5、直8及直9的类别判断属于细粒度分类问题。那么如何在小样本学习中使模型关注样本中细微的区分性区域,也是值得进一步探索的方向。

3) 更加复杂的元任务。在本文场景中,单个元任务仅针对5类或8类红外空中目标进行分类,但在实际场景中,单个元任务可能还涉及目标检测、定位、跟踪等问题。再者,当待分类类别增加至几十类或上百类时,小样本学习应该如何应对也是一个具有挑战性的研究方向。

参 考 文 献

- [1] Xie J R, Li F M, Wei H, et al. Enhancement of single shot multibox detector for aerial infrared target detection[J]. *Acta Optica Sinica*, 2019, 39(6): 0615001.
谢江荣, 李范鸣, 卫红, 等. 用于空中红外目标检测的增强单发多框检测器方法[J]. *光学学报*, 2019, 39(6): 0615001.
- [2] Ratner A J, Ehrenberg H, Hussain Z, et al. Learning to compose domain-specific transformations for data augmentation[C]//*Neural Information Processing Systems 2017*, December 4-9, 2017, Long Beach, California. New York: Curran Associates, 2017: 3236-3246.
- [3] Kulkarni T D, Whitney W F, Kohli P, et al. Deep convolutional inverse graphics network[C]//*Neural Information Processing Systems 2015*, December 7-12, 2015, Montreal, QC, Canada. New York: Curran Associates, 2015: 2539-2547.
- [4] Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning[J]. *Journal of Big Data*, 2019, 6: 60.
- [5] Shrivastava A, Pfister T, Tuzel O, et al. Learning from simulated and unsupervised images through adversarial training[C]//*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI. New York: IEEE, 2017: 2107-2116.
- [6] Antoniou A, Storkey A, Edwards H. Data augmentation generative adversarial networks [EB/OL]. (2018-03-21) [2019-12-03]. <https://arxiv.org/abs/1711.04340>.
- [7] Xie J R, Li F M, Wei H, et al. Infrared target simulation method based on generative adversarial neural networks[J]. *Acta Optica Sinica*, 2019, 39(3): 0311002.
谢江荣, 李范鸣, 卫红, 等. 基于生成对抗式神经网络
- 的红外目标仿真方法[J]. *光学学报*, 2019, 39(3): 0311002.
- [8] Li F F, Fergus R, Perona P. One-shot learning of object categories[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 594-611.
- [9] Jia X, Zhang J L, Wen X B. Infrared faults recognition for electrical equipments based on dual supervision signals deep learning[J]. *Infrared and Laser Engineering*, 2018, 47(7): 0703003.
贾鑫, 张惊雷, 温显斌. 双监督信号深度学习的电气设备红外故障识别[J]. *红外与激光工程*, 2018, 47(7): 0703003.
- [10] Khosravan N, Bagci U. S4ND: single-shot single-scale lung nodule detection[M]//Frangi A, Schnabel J, Davatzikos C, et al. *Medical image computing and computer assisted intervention-MICCAI 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11071: 794-802.
- [11] Jamal M A, Qi G J. Task agnostic meta-learning for few-shot learning[C]//*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 16-20, 2019, Long Beach, California. New York: IEEE, 2019: 11719-11727.
- [12] Sun Q R, Liu Y Y, Chua T S, et al. Meta-transfer learning for few-shot learning[C]//*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 16-20, 2019, Long Beach, California. New York: IEEE, 2019: 403-412.
- [13] Vinyals O, Blundell C, Lillicrap T, et al. Matching networks for one shot learning[C]//*Neural Information Processing Systems 2016*, December 5-10, 2016, Barcelona, Spain. New York: Curran Associates, 2016: 3630-3638.
- [14] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[C]//*Neural Information Processing Systems 2017*, December 4-9, 2017, Long Beach, California. New York: Curran Associates, 2017: 4077-4087.
- [15] Sung F, Yang Y X, Zhang L, et al. Learning to compare: relation network for few-shot learning[C]//*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 18-22, 2018, Salt Lake, UT, USA. New York: IEEE, 2018: 1199-1208.
- [16] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 1-9.
- [17] Kingma D P, Ba J. Adam: a method for stochastic

- optimization[EB/OL]. (2017-01-30)[2019-12-03]. <https://arxiv.xilesou.top/abs/1412.6980>.
- [18] Banerjee A, Merugu S, Dhillon I S, et al. Clustering with Bregman divergences[J]. *Journal of Machine Learning Research*, 2005, 6: 1705-1749.
- [19] Ravi S, Larochelle H. Optimization as a model for few-shot learning[C] // *International Conference on Learning Representations*, April 24-26, 2017, Toulon, France. [S.l.: s.n.], 2017.
- [20] Chen W Y, Liu Y C, Kira Z, et al. A closer look at few-shot classification[EB/OL]. (2020-01-12)[2019-12-03]. <https://arxiv.xilesou.top/abs/1904.04232>.
- [21] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks[C] // *Proceedings of the 34th International Conference on Machine Learning*, August 6-11, 2017, Sydney, NSW, Australia. [S.l.: s.n.], 2017, 70: 1126-1135.
- [22] Ben-David S, Blitzer J, Crammer K, et al. Analysis of representations for domain adaptation[C] // *Neural Information Processing Systems 2006*, December 4-7, 2006, Vancouver, BC, Canada. New York: Curran Associates, 2007: 137-144.
- [23] Wang M, Deng W H. Deep visual domain adaptation: a survey[J]. *Neurocomputing*, 2018, 312: 135-153.
- [24] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[M] // Fleet D, Pajdla T, Schiele B, et al. *Computer Vision-ECCV 2014*. Lecture Notes in Computer Science. Cham: Springer, 2014, 8689: 818-833.