

文章编号 1004-924X(2023)12-1816-11

基于度量元学习的铁路小样本入侵目标检测方法

郭保青^{1,2*}, 张德芬¹

(1. 北京交通大学 机械与电子控制工程学院, 北京 100044;

2. 北京交通大学 智慧高铁系统前沿科学中心, 北京 100044)

摘要: 异物入侵是导致铁路安全事故频发的主要原因之一, 传统深度学习需要大量训练样本进行网络训练, 但铁路场景中入侵样本很少且难于获取。本文提出了基于改进度量元学习的铁路小样本异物入侵检测方法。为了让入侵目标的特征表征在分类时发挥更大作用, 提出了基于通道注意力机制的特征提取网络; 为解决样本数量不足时个别样本在特征空间中产生偏离的问题, 提出了一种基于类中心微调的网络用于类别中心的修正; 同时, 基于 center loss 与交叉熵构建了中心相关损失函数用于小样本网络训练, 提升特征空间中同类别特征分布的紧凑性。在公共数据集 miniImageNet 上与经典小样本学习模型中最优的相比, 本文算法在 5-way 5-shot 设置下图像分类准确率提升了 7.31%。在铁路入侵小样本数据集的 5-way 5-shot 消融实验表明: 本文提出的通道注意力机制(Channel Attention Mechanism, CAM)和中心相关损失函数分别提升 0.86% 和 1.91% 的检测精度; 提出的类中心微调 and 预训练方法对检测精度的提升效果更明显, 分别达到 3.05% 和 6.70%, 上述模块综合应用的提升效果达到了 7.90%。

关键词: 小样本学习; 度量元学习; 铁路限界入侵; 目标检测; 注意力机制

中图分类号: U298; TP394.1 **文献标识码:** A **doi:** 10.37188/OPE.20233112.1816

Railway few-shot intruding objects detection method with metric meta learning

GUO Baoqing^{1,2*}, ZHANG Defen¹

(1. School of Mechanical, Electronic and Control Engineering,
Beijing Jiaotong University, Beijing 100044, China;

2. Frontiers Science Center for Smart High-speed Railway System,
Beijing Jiaotong University, Beijing 100044, China)

* Corresponding author, E-mail: bqguo@bjtu.edu.cn

Abstract: Object intrusion is among the primary causes of railway accidents. Typically, traditional deep-learning methods require numerous samples for network training; however, intrusion samples in railway settings are scarce and difficult to obtain. Thus, in this paper, a railway few-shot intruding-object detection method based on an improved metric meta-learning network is proposed. To better exploit the features of intruding objects during classification, a feature-extraction network based on the channel attention mechanism is proposed. A network based on fine-tuning of the class center is proposed for class-center cor-

收稿日期: 2022-06-14; 修订日期: 2022-08-06.

基金项目: 国家自然科学基金资助项目 (No. 52072026); 中国国家铁路集团有限公司系统性重大课题资助项目 (No. P2020T002)

rection to solve the problem of individual samples deviating in the feature space of insufficient samples. Additionally, a central correlation loss function based on the center loss and cross entropy is constructed for few-shot network training to improve the compactness of the same-class feature distribution in the feature space. In experiments on a public few-shot dataset called miniImageNet, the accuracy of the proposed method is 7.31% higher than the optimal accuracy of the classical few-shot learning model. In five-way five-shot ablation experiments using a railway dataset, the proposed channel attention mechanism and center-related loss function increase the mean average precision (mAP) by 0.86% and 1.91%, respectively. Additionally, the center fine-tuning and pretraining increase the mAP by 3.05% and 6.70%, respectively, and the total mAP improvement is 7.90%.

Key words: few-shot learning; metric meta learning; railway clearance intrusion; object detection; attention mechanism

1 引言

中国作为世界上高速铁路运营里程最长、运营速度最高、在建规模最大的国家,保障列车行车安全是铁路安全保障系统的重中之重。其中自然灾害或人为原因导致的人员和异物侵限等线路危情是造成铁路重大事故的重要原因之一,如何高效且准确地识别入侵铁路的异物对保障铁路运营安全具有重要意义。

目前,利用布拉格光纤光栅传感器^[1]或雷达^[2]进行异物侵限检测的方式大都存在部署安装困难、费用较高且报警准确率低等问题。随着深度学习的飞速发展,基于智能视频分析的方法被越来越多地应用在铁路异物入侵检测系统中。王瑞^[3]等人提出一种改进 Cascade Mask RCNN 的深度学习模型用于铁路周界入侵检测;陈莹^[4]提出模态自适应学习机制下的多光谱行人检测网络;郭保青^[5]利用卷积神经网络和传统的图像处理方法实现了对铁路中行人的分类;何文玉^[6]等人利用深度学习的 YOLO-V3 网络模型进行多次训练,在铁路数据集上验证了网络的可行性。但上述深度学习方法往往需要收集足够多的训练样本用于网络训练。而在实际中,一方面线路落石、泥石流等铁路危情往往无法收集到足够多的样本,另一方面随着社会不断发展,新的异物类别会不断出现^[7],而传统深度学习无法识别这些新类别的入侵异物。

为了解决上述问题,本文将小样本学习应用到铁路危情识别中。度量元学习^[8-11]是目前小样本学习方法中效果较好的方法,度量元学习重点

研究两部分内容:一是设计好的特征映射模块,使得同类别样本之间距离更近,不同类别样本之间距离更远,二是学习一种好的度量模型更好地描述不同类别样本之间的差距。

特征映射模块的性能直接决定了模型最终的检测精度。目前的特征映射模块大多是利用卷积块堆叠提取特征^[8-9],所有特征对最终分类决策的贡献一样。但实际上,待检测的特征才是网络需要重点关注的内容,对分类贡献应该最大。为解决这一问题,本文提出了一种改进的通道注意力模块,利用提取的特征生成每个通道上特征权重用作后续分类中的距离度量。

在类别中心初始化方面,SNELL J 等人^[8]采用支持集中所有样本特征均值作为所有类别中心,SUNG F 等人^[9]则直接将支持集中样本特征融合,但有时噪声样本可能会分布在偏离特征空间中较远的位置,并不适合用作类别中心的初始化,因此本文提出一种类别中心微调算法,让网络自己学习得到每个类别的中心。

度量元学习算法一般直接采用元训练方式对网络进行训练,并采用交叉熵损失函数优化网络参数;元训练可以让网络学会快速学习,但由于缺少先验知识,特征提取能力较弱;同时交叉熵损失只考虑能否分类正确,并没有考虑到同类样本特征在特征空间中的分布关系。为提高网络的特征提取能力,同时让同类别特征在样本空间中分布更加紧凑,本文引入模型预训练方式和中心相关损失函数来提升检测效果。

本文通过带有注意力机制的残差网络将样本映射到特征空间中,给予代表入侵目标的重要

特征更多的分类权重;提出类中心微调算法初始化类别中心,通过模型预训练获取更多先验知识,同时提出中心相关损失函数优化特征映射网络参数,使得同类别样本在特征空间中更加紧凑,便于后续模型分类。在铁路数据集和公共数据集 miniImageNet 的相关实验表明:与经典方法相比,本文算法在 5-way 5-shot 和 5-way 1-shot 的小样本任务上分类精度均有较大提升。

2 基于改进度量元学习的目标分类

度量元学习方法^[12]是目前解决小样本学习问题效果显著的一种方法。该方法利用距离函数作为目标函数,通过迭代训练获得一个特征映射网络,使得同类别样本在特征空间中相似度更高,不同类别的样本在特征空间中相似度更低。其特征映射网络和度量模块决定了度量元学习模型性能的好坏。因此,本文也从这两个方面着手进行研究。

本文设计的小样本分类网络整体框架如图 1 所示,主要分为三个部分:(1)特征映射网络 f_ϕ ;(2)类中心微调网络 g_ψ ;(3)距离度量模块 r 。该网络以支持集(Support set)和查询集(Query set)图片作为输入,首先用特征映射网络将其映射到同一个特征空间;接着,将支持集图像特征输入到类中心微调网络中,学习得到各类别的中心表示;最后,利用距离度量模块计算查询集图像特征与各类别中心之间的相似度,预测分类结果。

2.1 基于通道注意力的特征映射网络

由于不同的图像通道包含了不同特征,为了让样本图像中代表入侵目标的关键通道特征在

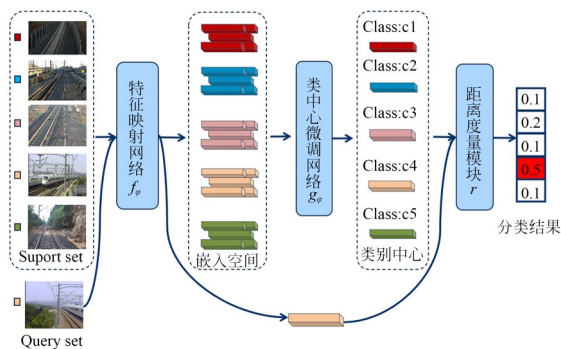


图 1 网络整体框架示意图

Fig. 1 Overall network framework

分类时发挥更大作用,本文提出图 2 所示的基于通道注意力机制的特征映射网络,主要包括主干网络和通道注意力模块。目前一些经典网络多采用四层卷积网络进行特征提取^[8-9],这种网络结构无法提取更深层次信息,因此本文选用具有残差结构^[13]的 Resnet12 作为主干网络,当样本量不足时,残差结构可以有效控制网络深度,避免过拟合,同时可以提取深层次信息,残差块的具体结构如图 3 所示。

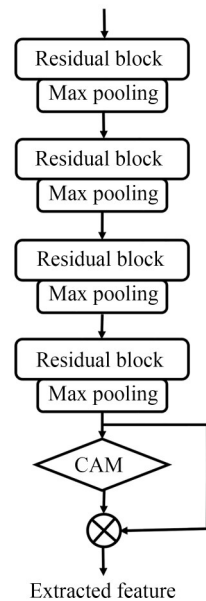


图 2 特征映射网络的结构

Fig. 2 Structure of feature extraction network

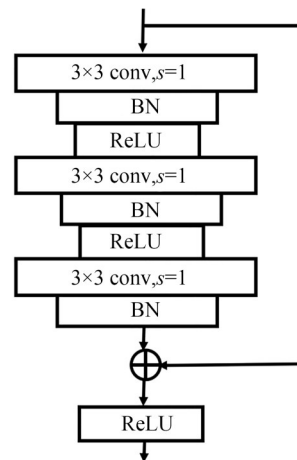


图 3 残差块结构

Fig. 3 Structure of residual block

在进行距离度量时,希望网络对重要特征给予更多关注^[14],与 CBAM (Convolutional Block

Attention Module)^[15]中的通道注意力模块相比,本文去掉了其由多层感知机和一个隐层组成的共享网络,减少了网络参数,防止在样本数量过少时产生过拟合;同时,在主干网络最后一层引入如图4所示的改进通道注意力模块。该模块由三个部分构成,第一部分为信息压缩,将卷积后的特征 $F \in \mathbf{R}^{H \times W \times C}$ 通过并行的全局最大池化层 $F_m(\cdot)$ 和全局平均池化层 $F_a(\cdot)$ 在空间维度上进行压缩;第二部分为信息激励 $F_g(\cdot)$,将上述两部分信息的特征相加融合,再利用激活函数将融合后的特征归一化到0~1之间得到通道注意力权重;第三部分为通道特征增强,将获得的通道权重值与原特征 F 相乘,从而实现对关键通道特征的注意力增强。

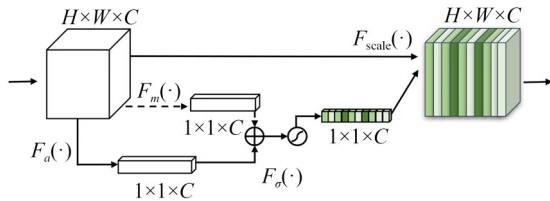


图4 改进的通道注意力模块

Fig. 4 Improved channel attention module

2.2 类中心微调网络

利用上述基于通道注意力的特征映射网络完成特征映射后,所有样本被映射到同一个特征空间。基于度量元学习的原型网络一般采用类别中心均值算法求取每个类的特征中心^[8],当在样本量较少时,该方法可以有效解决过拟合问题,但个别样本的偏离会导致均值类中心偏离真实的中心位置。而传统的分类网络一般采用全连接层作为分类层,训练过程中通过不断更新全连接层参数得到的最优解即为每个类别的初始化中心,克服了个别样本中心偏离过大的问题,但需要训练的参数量较大。本文综合上述全连接网络与类中心均值算法的优势,提出类中心微调网络对特征空间中的类别中心进行初始化,首先将类中心作为一组可学习的参数,利用每一类所有支持集中的图片特征均值作为其初始值,再让特征空间中带标签的图片

通过自主学习的方式去更新类中心参数,其伪代码如算法1所示。

算法1中,对于支持集 S 中类别为 k 的所有样本 $S_k = \{x_i | i = 1, 2, \dots, N\}$,经过特征映射网络后得到特征 $f_\phi^k(x_i)$,求得均值 c_k 作为类别 k 的可学习中心参数在特征空间的初始值,如式(1)所示:

$$c_k(\varphi) = \frac{1}{N} \sum_{x_i \in S_k} f_\phi^k(x_i). \quad (1)$$

根据式(1)依次得到支持集中所有类中心原型的初始值,利用 softmax 函数,归一化支持集中样本到各类别初始中心的距离,得到属于类 k 的概率,如式(2)所示:

$$p(\varphi)(y = k|x) = \frac{\exp(-d(f_\phi(x), c_k(\varphi)))}{\sum_k \exp(-d(f_\phi(x), c'_k(\varphi)))}, \quad (2)$$

其中 d 为两特征向量在特征空间的欧式距离,然后利用公式(3)计算交叉熵损失:

$$L(\varphi) = -\frac{1}{N} \sum_{i=1}^N \log(p(\varphi)(y = k|x_i)). \quad (3)$$

在类中心参数微调过程中,每次从支持集中抽取4张图片进行微调,直到遍历完整个数据集为一次微调,共循环100次,利用SGD优化器进行反向传播,完成类中心参数的更新。

算法1:类中心微调(Pseudo Code of center fine tuning)

输入:支持集 $S = \{(f(x_1), y_1), \dots, (f(x_N), y_N)\}$, 其中

$y_i \in \{1, \dots, k\}$, D_k 表示支持集 S 中所有标签为 $y_i = k$ 的样本集合

输出:更新得到的每个中心值 $\{c_1, \dots, c_k\}$

1. 初始化每个类中心值:

for k in $\{1, \dots, k\}$ do

$$c_k(\varphi) \leftarrow \frac{1}{N} \sum_{x_i \in S_k} f_\phi^k(x_i)$$

end for

2. 归一化支持集中样本 x_i 到每个初始类中心的距离:

for i in $\{1, \dots, N\}$ do

$$p(\varphi)(y = k|x_i) \leftarrow \frac{e^{-d(f_\phi(x_i), c_k(\varphi))}}{\sum_k e^{-d(f_\phi(x_i), c'_k(\varphi))}}$$

end for

3. 计算损失并更新类中心:

$$L(\varphi) \leftarrow -\frac{1}{N} \sum_{i=1}^N \log(p(\varphi)(y = k|x_i))$$

2.3 特征映射网络参数预训练

为了让特征映射网络具备更丰富的先验知识,本文引入预训练模型,将训练步骤分为如图5所示的两个阶段,第一阶段将整个训练集用标准的交叉熵损失训练一个N类分类器,如式(4)所示:

$$T = \arg \min_{\theta, \omega} \sum_{i=1}^N L(C_{\omega}(f_{\theta}(x_i)), y), \quad (4)$$

其中: L 为交叉熵损失, y 为特征所属标签, C_{ω} 则为softmax分类器;第二阶段,去除第一阶段训练模型的全连接层,将已经预训练过后的特征映射网络迁移到小样本学习中进行元训练。具体步骤如下:

(1)首先将特征映射网络放到标准的分类网

络中,与深度学习中一般的分类网络方式相同,利用训练集中所有图片训练网络;

(2)每个epoch后,将特征映射网络迁移到构建好的小样本学习模型,在验证集上采用元验证的方式,每个epoch都采用这种一边训练一边元验证的方式;

(3)整个训练过程结束后,保存在元验证过程中效果最好的特征映射网络参数;

(4)预训练结束后,将特征映射网络以及训练好的模型参数加载到构建好的小样本学习模型中,接着将已经被划分成多个元任务的训练集和验证集输入到模型,分别进行模型的元训练和元验证,训练结束后,保存检测精度最优的模型为最终训练模型。

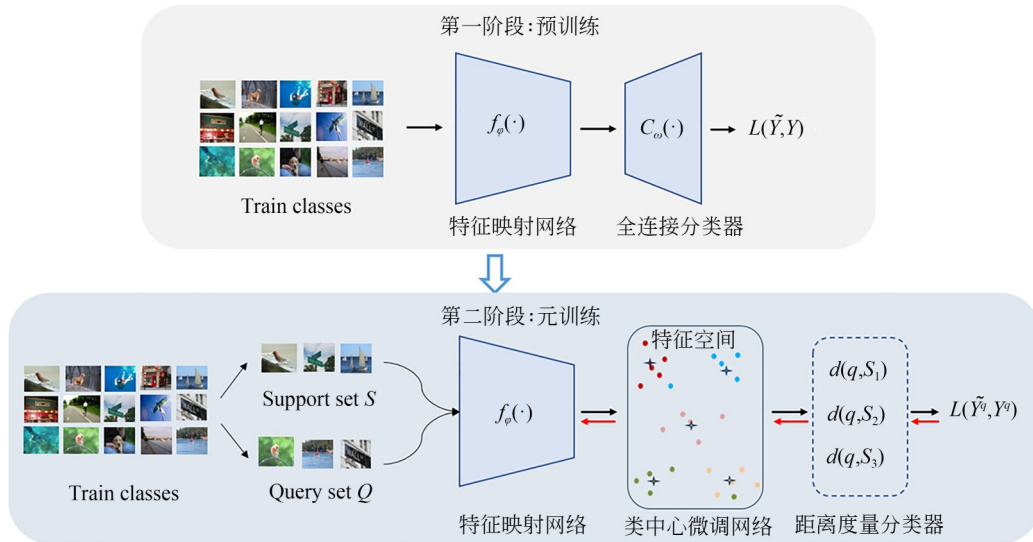


图5 网络预训练策略

Fig. 5 Network pre-training strategy

3 中心相关损失函数

为了让特征映射网络学习到一个更加具有判别能力的映射空间,即在该空间中使同类别样本特征分布更加紧凑,不同类别样本特征具有更好的区分性,本文将人脸识别领域中的Center loss^[16]损失函数引入到小样本学习中,结合交叉熵损失构建了中心相关损失函数。

首先Center loss损失函数如式(5)所示:

$$L_c = \lambda \sum_{i=1}^m \|f(x_i) - c_k\|^2, \quad (5)$$

其中: λ 是一个超参数,在训练的过程中用于控制类内的距离, $f(x_i)$ 表示某一图片 x_i 在特征空间中的特征值, c_k 表示该类的类别中心。与机器学习领域随机初始化类别中心不同,本文在元训练过程中 c_k 初始化如式(1)所示。

同时,在训练过程中,每个Episode都根据式(1)重新初始化类别中心。最后采用交叉熵损失和Center loss结合构建如式(6)所示的中心相关损失函数用于特征映射网络训练。

$$L_{\text{total}} = L_s + L_c = -\frac{1}{m} \sum_{i=1}^m \log \left(\frac{e^{-d(f_\theta(x_i), c_k)}}{\sum_{k'} e^{-d(f_\theta(x), c_{k'})}} \right) + \lambda \sum_{i=1}^m \|f(x_i) - c_k\|^2, \quad (6)$$

式中,交叉熵损失保证类别可分,Center loss使同类别样本特征在特征空间中的分布差异性更小,即缩小类间差距,算法细节如算法2所示。

算法2:特征映射网络学习算法 (Feature mapping network parameter learning algorithm)
输入:支持集 $S = \{(f(x_1), y_1), \dots, (f(x_N), y_N)\}$, 其中 $y_i \in \{1, \dots, k\}$, 查询集 $Q = \{(f(x_1), y_1), \dots, (f(x_m), y_m)\}$, 特征映射网络初始化参数 θ_c , 超参数 λ 和 α , 迭代次数 $t \leftarrow 0$ 。
输出:特征映射网络参数 θ_c
1: while not converge do:
2: $t \leftarrow t + 1$
3: 计算总的损失:
$L_{\text{total}} = L_s + L_c = -\frac{1}{m} \sum_{i=1}^m \log \left(\frac{e^{-d(f_\theta(x_i), c_k)}}{\sum_{k'} e^{-d(f_\theta(x), c_{k'})}} \right) + \lambda \sum_{i=1}^m \ x_i - c_k\ ^2$
4: 初始化类别中心:
for i in k:
$c_k = \frac{1}{ S_k } \sum_{(x_i, y_i) \in S_k} f(x_i)$
end for
5: 计算反向传播误差: $\frac{\partial L}{\partial x_i} = \frac{\partial L_s}{\partial x_i} + \lambda \cdot \frac{\partial L_c}{\partial x_i}$
6: 更新参数 $\theta_c: \theta_c = \theta_c - \alpha \sum_i \frac{\partial L}{\partial x_i} \cdot \frac{\partial x_i}{\partial \theta_c}$
7: end while

4 实验结果与分析

4.1 实验数据集

为了验证本文算法的效果,分别在小样本公共数据集 miniImageNet 和铁路数据集进行 5-way 5-shot, 5-way 1-shot 的实验,下面首先介绍这两个数据集:

miniImageNet: 该数据集包含从 ILSVRC-2012 中随机选择的 100 个类,每类包含 600 张 84×84 pixel 的图像,被分为 64 个训练类、16 个验证类和 20 个测试类^[14]。

铁路数据集: 通过整理高速铁路综合视频监

控相机获取的图像建立了铁路行人入侵图像,铁路空场景图像和铁路列车图像数据集,针对线路落石和泥石流两类难以收集样本的类别,分别采用人工模拟和网络收集等手段构建铁路落石与泥石流样本数据集,如表 1 所示,每类样本有 40 张图像。

表 1 铁路数据集构成

Tab. 1 Railway dataset composition

编号	场景类别	数量/张
1	空场景	40
2	行人入侵	40
3	泥石流入侵	40
4	落石入侵	40
5	列车经过	40

4.2 实验设置

在模型训练阶段,采用 Adam (Adaptive Moment Estimation) 优化算法对模型参数进行更新,初始学习率为 0.001, Center loss 损失函数中的超参数 λ 取 0.001, 采用等间隔调整的方式每 20 个 epoch 更新一次学习率,乘法因子 gamma = 0.5, 每个 epoch 包含 100 个元任务,共 200 个 epoch, 每个元任务支持集 (support set) 中有 5 个类别,每类有 5 张图片,即 5-way 5-shot 设置。在测试阶段,用于微调的分类器采用 SGD 优化算法对中心进行微调,初始学习率设为 0.1, 动量因子 momentum = 0.9, 动量抑制因子 dampening = 0.9。

4.3 对比实验

为了验证算法的有效性,分别在公共数据集 miniImageNet 和铁路数据集上进行了对比实验。

4.3.1 miniImageNet 数据集实验结果

为了验证本文算法的有效性,将本文设计的模型与经典的小样本学习算法 MAML^[17], ProtoNet^[8], RelationNet^[9] 等进行比较,所有的精度结果均以 95% 的置信区间平均 2 000 个元测试任务获得。实验结果如表 2 所示,在 5-way 5-shot 和 5-way 1-shot 设置下,本文算法相比于上述经典网络检测精度均有较大提升,尤其是 5-way 5-shot 下相比最优的 Protonet 提升 7.31%。

表 2 MiniImageNet数据集实验结果

Tab. 2 Experiment results on MiniImageNet dataset(%)

5-way Acc		
Model	1-shot	5-shot
MAM ^[17]	48.70	63.11
RelationNet ^[9]	50.44	65.32
ProtoNet ^[8]	48.51±0.40	67.09±0.36
MatchNet ^[18]	43.40±0.78	51.09±0.71
本文算法	50.65±0.45	74.40±0.33

4.3.2 模型迁移到铁路数据集实验效果

为了进一步分析模型在铁路数据集上的效果,根据文献^[19]跨域进行元训练和元测试进行模型迁移的思想,将在公共数据集 miniImageNet上元训练后的模型迁移至铁路数据集进行元测试。

将本文算法与经典网络 ProtoNet, RelationNet, MAML 和 MatchNet 在 5-way 1-shot 和 5-way 5-shot 设置下的实验检测结果进行了比较,实验结果如表 3 所示。在铁路小样本数据集下,本文算法在 5-way 5-shot 设置下,相比于在铁路数据集效果最好的经典网络 ProtoNet 检测精度有 3.61% 的提升,在 5-way 1-shot 设置下有 3.51% 的提升。

表 3 铁路数据集实验效果

Tab. 3 Experiment results on railway datasets (%)

5-way Acc		
Model	1-shot	5-shot
MAML	60.23±0.21	75.26±0.24
RelationNet	61.95±0.32	76.68±0.40
ProtoNet	62.40±0.36	81.83±0.24
MatchNet	58.45±0.23	65.23±0.35
本文算法	65.91±0.35	85.44±0.25

为了比较不同类型目标识别的效果,表 4 给出了不同类型目标组合的分类准确率,从中可以看出,背景、泥石流、列车三类大目标识别精度为 95.53%,在此基础上加入落石后识别精度降为 91.48%,加入落石、行人两类目标后识别精度降为 85.44%。由此可见,落石或行人等较小目标的识别准确率较低。

为了更加直观地看到网络加入通道注意力机制后的效果,利用 GradCAM 可视化技术,提取

表 4 不同类型目标的准确率

Tab. 4 mAp on different kinds of object (%)

目标类别	准确率
背景、泥石流、列车	95.53±0.18
背景、泥石流、列车、落石	91.48±0.21
背景、泥石流、列车、落石、行人	85.44±0.25

图像经过特征映射网络后最后一次卷积后的特征图,从最大概率分类类别的节点出发,反向传播求出特征图的梯度均值作为每张特征图的权重,将每张特征图乘以权重得到带权重的特征图,在第三维上求取均值,ReLU 激活后再经过归一化处理,将处理后的热力图缩放到图像尺寸大小,最后叠加到原图中,效果可视化如图 6 所示(彩图见期刊电子版)。

在图 6 中,左边图片代表输入网络的原始图,

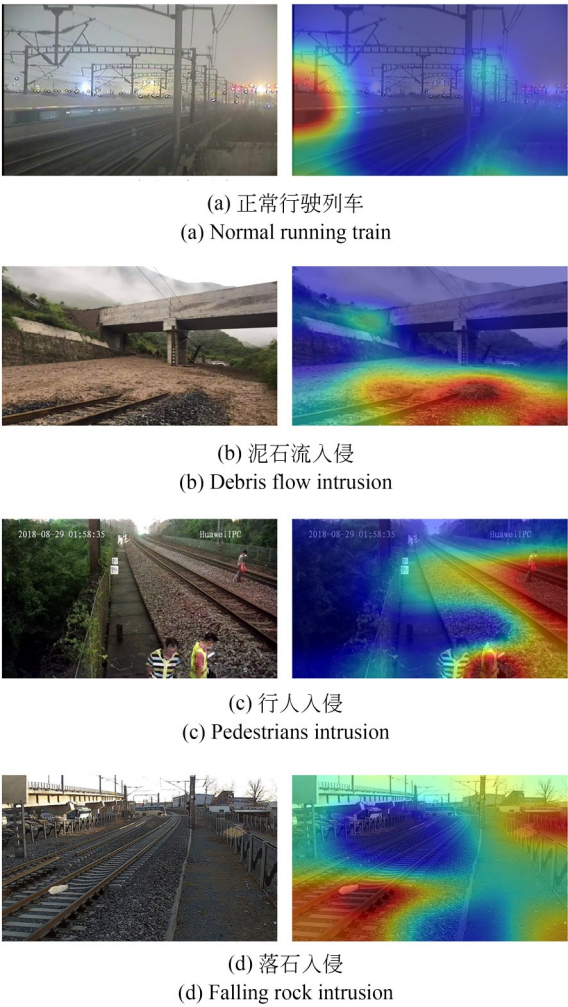


图 6 铁路数据集热力图效果

Fig. 6 Thermal map effect of railway datasets

右边图片是我们利用 GradCAM 可视化技术输出最终分类时图像中不同特征对分类的贡献量,图中颜色从蓝到红的变化代表贡献量逐渐增大,蓝色贡献量低,红色区域贡献量高。基于通道注意力机制的特征提取网络会自动对不同通道的特征进行提取并提升代表入侵目标的关键通道特征权重,如图 6 所示,经过通道注意力机制可以使入侵目标通道特征获得更高的关注度,入侵目标在热力图上表现为显著性更高的红色区域。对于图 6(a)所示的列车驶入以及图 6(b)所示的线路泥石流这类比较大的入侵目标,网络能够比较准确地关注到物体位置;对于如图 6(c)和图 6(d)所示的行人和落石这类比较小的入侵异物,这些异物都处于网络的关注区域内;但铁路场景中的防护网等目标也会对网络的检测产生一定的干扰,这可能是造成该类目标分类准确率较低的原因之一。

为了更加直观地观察使用模型预训练和中心相关损失前后度量空间的对比情况,在 5-way 5-shot 设置下,分别提取了训练结束后模型最后一层卷积神经网络的特征图,利用 TSNE 分析法将其降维至二维平面,效果如图 7 所示。

图 7(a)中,空场景和行人类别样本交织分布,这会导致距离度量模块很难将二者进行区分;同时,落石在特征空间中分布较为分散,从而导致分类错误率较高。而图 7(b)中,使用了模型预训练和中心相关损失函数后,行人和空场景分布趋于远离并且落石类目标的特征分布也趋于密集,说明中心相关损失函数可以有效地缩小同类样本之间的距离,在模型预训练之后,各类样本在特征空间内的分布质量明显提高。实验证明使用模型预训练和中心相关损失函数可以在铁路小样本数据集中取得不错的分类效果。

4.4 面向铁路数据集的消融试验

为了验证改进 CAM 模块的有效性,本文在铁路数据集进行了改进 CAM 与 CBAM 的对比实验,实验结果如表 5 所示。基本网络中加入 CBAM 后,网络检测精度在 5-way 1-shot 和 5-way 5-shot 设置下均出现了不同程度的下降;而加入改进 CAM 后,在 5-way 1-shot 和 5-way 5-shot 设置下,本文算法相比于基本网络+CBAM 检测精度分别提升了 4.64% 和 2.54%。

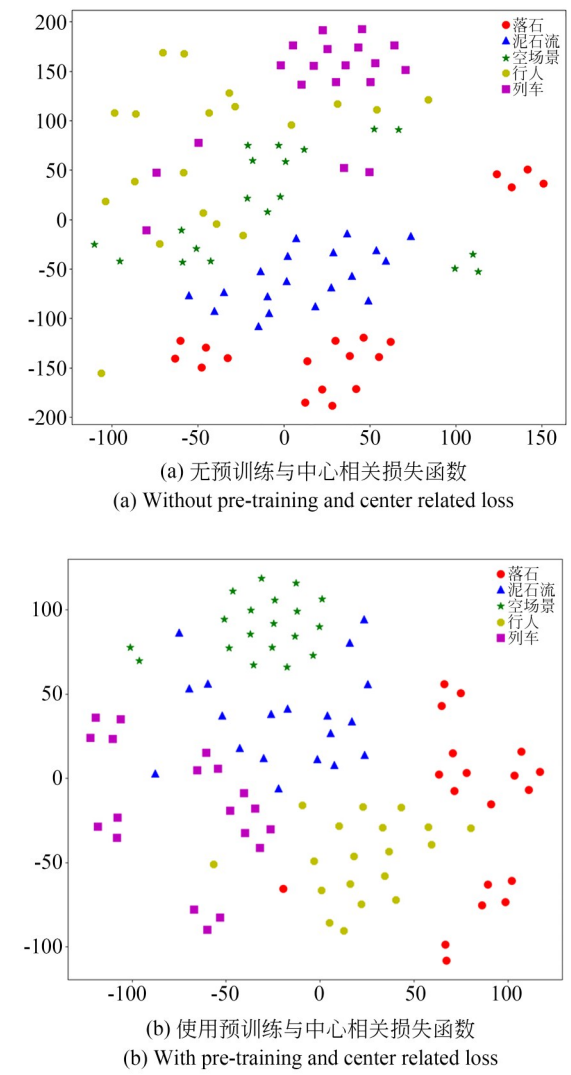


图 7 使用模型预训练和中心相关损失前后的样本分布
Fig. 7 Sample distribution with and without model pre-training and center related loss

表 5 不同注意力机制实验		
Tab. 5 Experiments on different attention mechanism(%)		
网络结构	5-way 1-shot	5-way 5-shot
基本网络	59.01±0.41	77.53±0.25
基本网络+CBAM	58.12±0.26	75.85±0.23
基本网络+改进 CAM	62.76±0.41	78.39±0.24

为了验证本文算法设计的通道注意力模块、类中心微调算法、模型预训练和中心相关损失函数在铁路数据集上的有效性,以本文构建的模型为基准,分别在网络中加入分别这四个模块进行消融实验。为了保证实验的公平性,在元训练过

程中均采用公共数据集下 20-way 5-shot 小样本学习任务设置,在铁路数据集中的元测试过程采用 5-way 5-shot 设置,其中距离函数采用曼哈顿距离,实验结果如表 6 所示。改进的注意力机制 CAM 模块在精度的提高上有显著作用,提高了 0.86% 的检测精度;中心相关损失函数可以使得同类别特征更加紧凑,提升了 1.91% 的检测精

度;类中心微调算法可以减小噪声样本对类中心初始化影响,提升了 3.05% 的检测精度;模型预训练后检测精度提升了 6.70%;四者结合起来检测精度共提升 7.90%。由此可见,CAM 与中心相关损失函数能够小幅提升模型检测精度,类中心微调与模型预训练对铁路数据集检测精度提升更明显,更适用于铁路场景。

表 6 铁路数据集消融实验
Tab. 6 Ablation experiments on railway datasets

基本网络	改进 CAM	类中心微调	中心相关损失	模型预训练	5-way 1-shot	5-way 5-shot
✓					59.01±0.41	77.53±0.25
✓	✓				62.76±0.41	78.39±0.24
✓		✓			—	80.58±0.24
✓			✓		64.90±0.39	79.44±0.22
✓				✓	64.16±0.40	84.23±0.23
✓	✓	✓	✓	✓	65.91±0.35	85.44±0.25

在 5-way 1-shot 设置下,即当支持集中每个类别的图片只有一张时,此时中心微调算法不再起作用,故在此不做讨论;网络中单独引入 CAM 模块时,检测精度可提升 3.75%;单独加入中心相关损失函数后,检测精度可提升 5.89%;加入模型预训练后,检测精度提升 5.15%。与主干网络 ResNet12 相比,当待测目标所占区域较小时,加入改进的注意力机制 CAM 模块可以有效关注重要特征,使其在分类时所占据权重更大;在类中心初始化的过程中,类中心微调可以有效滤除特征空间中分布偏离同类别特征较远的噪声样本的干扰;而中心相关损失函数可以使得同类别特征在特征空间中更加紧凑,更加有利于后续的分类;在元训练之前引入模型预训练可以使特征映射网络获得更加丰富的先验知识,特征提取能力更强。因此,注意力模块 CAM、类中心微调算法、中心相关损失函数和模型预训练可以有效提升网络的分类检测精度。

4.5 支持集样本数对模型准确率的影响

为了验证样本数量对小样本学习模型的影响,本文依次增加铁路支持集样本数量进行了相关实验,5-way 5-shot,5-way 10-shot,5-way 15-shot 和 5-way 20-shot 设置下的实验结果如表 7 所示,从中可以看出,随着支持集中样本数量的增加,本文设计的小样本学习模型的准确率会持续

提升,在 5-way 20-shot 设置下可以达到 93.51±0.21% 的准确率。由于实际铁路场景应用中对各类入侵目标收集几十张样本具有可操作性,因此采用 5-way 20-shots 的识别模式具有广泛的实际应用价值。此外,在 5-way 20-shots 设置下单张图像识别的平均耗时为 0.426 ms,能够满足铁路场景中非常见类型入侵目标的实时分类需求。

表 7 不同 K 值的模型准确率
Tab. 7 Model accuracy under different K-shot (%)

N-way K-shot	Accuracy
5-way 5-shot	85.44±0.25
5-way 10-shot	88.16±0.22
5-way 15-shot	91.21±0.21
5-way 20-shot	93.51±0.21

5 结 论

针对现有深度学习算法进行目标识别准确率严重依赖训练样本数量而铁路场景入侵目标样本获取困难的问题,本文提出了基于改进度量元学习的铁路小样本入侵目标识别方法。通过设计基于通道注意力机制的特征映射网络用于重要特征提取,提出基于类中心微调算法用于类

别中心修正,并通过模型预训练和中心相关分类损失函数优化网络训练过程。在公共小样本数据集 miniImageNet 与铁路小样本数据集上的识别结果及消融实验表明,本文算法提出的通道注意力机制(CAM)和中心相关损失函数对网络检测精度提升起到了重要作用。在 5-way 5-shots

设置下,与经典算法中最优的相比,本文的算法在 miniImageNet 公共数据集上提高了 7.31%,在铁路数据集上提高了 3.61%。在 5-way 20-shots 设置下,本文算法在铁路小样本数据集上可以达到 93.51% 的识别准确率,对于识别铁路行人、落石及泥石流入侵具有广泛的实际应用价值。

参考文献:

- [1] CATALANO A, BRUNO FA, GALLIANO C, *et al.* An optical fiber intrusion detection system for railway security[J]. *Sensors and Actuators A: Physical*, 2017, 253: 91-100.
- [2] ALVAREZ F J, URENA J, MAZO M, *et al.* Ultrasonic Sensor System for Detecting Falling Objects on Railways [C]. *IEEE Intelligent Vehicles Symposium*. 14-17, 2004, Parma, Italy. IEEE, 2004: 866-871.
- [3] 王瑞,李霄峰,史天运,等. 基于视频深度学习的铁路周界入侵检测算法研究[J]. *交通运输系统工程与信息*, 2020, 20(2): 61-68.
WANG R, LI X F, SHI T Y, *et al.* Railway perimeter intrusion detection algorithms based on video deep learning [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2020, 20(2): 61-68. (in Chinese)
- [4] 陈莹,朱宇. 模态自适应权值学习机制下的多光谱行人检测网络[J]. *光学精密工程*, 2020, 28(12): 2700-2709.
CHEN Y, ZHU Y. Multispectral pedestrian detection network under modal adaptive weight learning mechanism [J]. *Opt. Precision Eng.*, 2020, 28(12): 2700-2709. (in Chinese)
- [5] 郭保青,王宁. 基于改进深度卷积网络的铁路入侵行人分类算法[J]. *光学精密工程*, 2018, 26(12): 3040-3050.
GUO B Q, WANG N. Pedestrian intruding railway clearance classification algorithm based on improved deep convolutional network [J]. *Opt. Precision Eng.*, 2018, 26(12): 3040-3050. (in Chinese)
- [6] 何文玉,杨杰,张天露. 基于深度学习的轨道异物入侵检测算法[J]. *计算机工程与设计*, 2020, 41(12): 3376-3383.
HE W Y, YANG J, ZHANG T L. Orbital foreign object intrusion detection algorithm based on deep learning [J]. *Computer Engineering and Design*, 2020, 41(12): 3376-3383. (in Chinese)
- [7] 宋闯,赵佳佳,王康等. 面向智能感知的小样本学习研究综述[J]. *航空学报*, 2020, 41(S1): 15-28.
SONG CH, ZHAO J J, WANG K, *et al.* A survey of few shot learning based on intelligent perception [J]. *Acta Aeronautica et Astronautica Sinica*, 2020, 41(S1): 15-28.
- [8] SNELL J, SWERSKY Y K, ZEMEL R. Prototypical net-works for few-shot learning[C]. *Advances in Neural Information Processing Systems*, 2017: 4077-4087.
- [9] SUNG F, YANG Y, ZHANG L, *et al.* Learning to compare: Relation network for few-shot learning[C]. IEEE, 2018: 1199-1208.
- [10] XU C Y, SHEN J Z, DU X. A method of few-shot network intrusion detection based on meta-learning framework[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 3540-3552.
- [11] ALFASSY A, KARLINSKY L, AIDES A, *et al.* LaSO: Label-Set Operations Networks for Multi-Label Few-Shot Learning[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 6541-6550.
- [12] ZHANG C, CAI Y J, LIN G S, *et al.* Deep-EMD: Few-Shot Image Classification with Differentiable Earth Mover's Distance and Structured Classifiers [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 12200-12210.
- [13] HE K M, ZHANG X Y, REN S Q, *et al.* Deep Residual Learning for Image Recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [14] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]. 2018 *IEEE/CVF Conference on*

- Computer Vision and Pattern Recognition*. 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [15] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional Block Attention Module[M]. Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [16] WEN Y D, ZHANG K P, LI Z F, *et al.* A Discriminative Feature Learning Approach for Deep Face Recognition [M]. Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 499-515.
- [17] 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 - Volume 70. August 6 - 11, 2017, Sydney, NSW, Australia. New York: ACM*, 2017: 1126-1135.
- [18] VINYALS O, BLUNDELL C, LILLICRAP T, *et al.* Matching Networks for One Shot Learning [C]. *Proceedings of the 30th International Conference on Neural Information Processing Systems. New York: ACM*, 2016: 3637-3645.
- [19] CHEN W, LIU Y, KIRA Z, *et al.* A closer look at few-shot classification[C]. *arXiv preprint arXiv:1904.04232*. 2019.

作者简介:



郭保青(1978—),男,河北涿水人,教授,博士生导师,2004,2009年于北京交通大学分别获硕士和博士学位,主要从事铁路基础设施状态智能感知,机器视觉检测方面的研究工作。Email:bqguo@bjtu.edu.cn



张德芬(1997—),女,湖北仙桃人,研究生在读,2019于北方工业大学获得学士学位,现于北京交通大学攻读硕士学位,研究方向为小样本学习目标检测。E-mail:19121285@bjtu.edu.cn