

引用格式: 曹哲骁,傅瑶,王丽,等. 基于深度学习网络的遥感图像异常检测方法研究[J]. 空间控制技术与应用, 2023, 49(6): 77-85. CAO Z X, FU Y, WANG L, et al. An anomaly detection method for remote sensing image based on deep learning network[J]. Aerospace Control and Application, 2023, 49(6): 77-85 (in Chinese). doi: 10.3969/j.issn.1674-1579.2023.06.008

基于深度学习网络的遥感图像异常检测方法研究

曹哲骁^{1,2}, 傅瑶^{3*}, 王丽⁴, 苏盈⁴, 郭云翔⁴, 王田^{1,2,5}

1. 北京航空航天大学, 北京 100191
2. 复杂关键软件环境全国重点实验室, 北京 100191
3. 中国科学院长春光学精密机械与物理研究所, 长春 130033
4. 武汉高德红外股份有限公司, 武汉 430223
5. 中关村实验室, 北京 100191

摘要: 针对现实世界中异常图像数据稀少的数据不均衡问题, 构建了一个高性能的异常检测模型. 仅使用正常训练数据和小部分仿真异常数据, 构建了两阶段框架的异常检测模型. 通过对正常数据和模拟生成的异常数据进行分类训练, 得到提取特征的 ResNet-18 编码器模型, 通过高斯密度估计对正常数据的特征建模, 构建异常图像的单分类器. Grad-CAM 扩展了模型, 使得异常检测模型可以在没有标签的情况下定位异常区域. 通过仿真异常检测数据集上进行的实验证明, 提出的算法能够检测现实世界遥感图像中人类肉眼难以发现的异常样本, 并给出定位结果.

关键词: 异常检测; 遥感图像; 深度学习; 卷积神经网络

中图分类号: TP39 文献标志码: A 文章编号: 1674-1579(2023)06-0077-09

0 引言

随着人工智能的迅速进步, 尤其是深度学习技术的不断演进, 航天高光谱遥感数据正迎来广阔的应用前景. 这类数据具备成像覆盖范围广、光谱分辨率高、图谱合一等特性, 是图像和光谱信息的有机融合^[1]. 基于深度学习的人工智能技术方法的迅猛发展, 有助于这些数据中的关键信息提取, 应用于场景分类、目标检测等应用场景^[2-3].

传统的目标检测和图像分割算法往往使用监督学习的范式, 需要大量人力资源用于收集充分数量的目标类别图像样本, 并对图像类别、目标位置或像素类别进行精确标注. 然而, 在许多应用场景下, 目标样本的获取十分艰难, 存在正负样本不均

衡的问题. 例如, 对于表面缺陷检测任务, 可供使用的数据集图像大部分表面并无缺陷, 属于正常样本, 而缺陷样本仅占数据集中的少数比例, 同时, 需要检测的缺陷种类却种类繁多, 更可能有许多数据集中未包含但现实中存在的缺陷种类, 使得可供如目标检测、图像分割等有监督训练范式的缺陷标签和对应样本严重受限^[4]. 在这些情况下, 由于目标类别样本的缺乏, 传统的监督学习方法遇到瓶颈. 与监督学习相比, 异常检测方法由于其无监督或半监督的范式, 构建检测模型的方法对样本标签的需求小, 对于数据集正负样本不均衡的情况十分适用. 在制造缺陷检测、医学图像分析、视频监控、高光谱图像处理等领域有较高的研究意义和应用价值^[5].

图像异常检测与典型的监督分类或检测问题不同之处在于对异常数据及其标签的获取受限. 因

收稿日期: 2023-02-22; 录用日期: 2023-06-27

基金项目: 国家自然科学基金资助项目(61972016 和 62032016) 和北京市科技新星资助项目(20220484106 和 202304844451)

* 通信作者: E-mail: fuy_cas@foxmail.com

此异常检测器通常是在半监督或仅使用正常数据的单分类无监督设置下构建的. 由于异常模式的分布在模型学习过程中的未知性, 模型在训练阶段学习正常实例的模式, 并在测试阶段不能较好表示测试实例的情况下确定异常.

在图像异常检测问题中, 传统的机器学习方法已经取得了较好的效果, 并得到了广泛应用. 随着深度学习技术的不断发展, 前沿的研究进一步将深度卷积神经网络与图像异常检测相结合, 以提高检测效果. 根据机器学习模型是否由神经网络组成, 目前主流的图像异常检测方法可分为两大类: 传统的机器学习方法和深度学习方法.

基于传统机器学习方法的异常检测技术包含

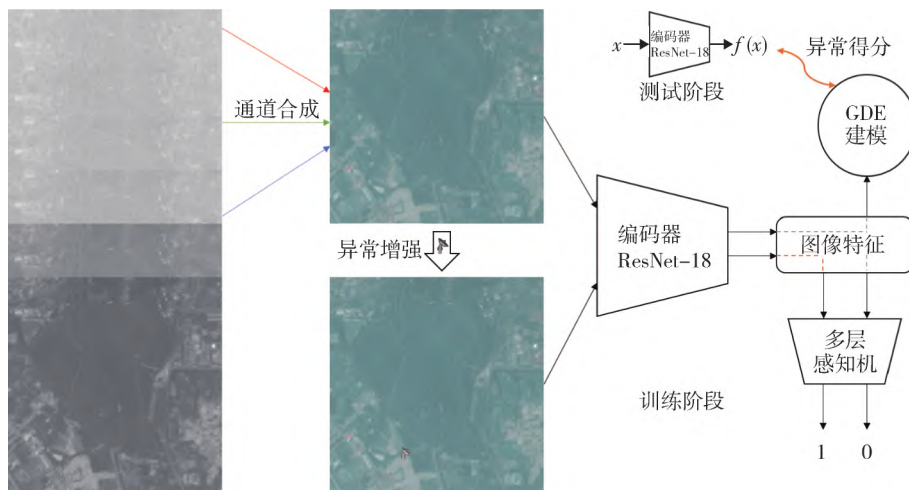


图 1 异常检测算法训练与测试框架

Fig. 1 Training and testing framework of the proposed anomaly detection algorithms

基于深度学习方法的异常检测技术通常采用基于图像重构的异常检测方法, 这种方法往往使用自编码器学习训练阶段正常输入图像的分布^[6-7], 在检测阶段通过重构图像和真实图像的重构误差作为异常评分, 从而进行异常检测.

基于自编码器的方法被认为是过于完备的, 存在着无法保证完全消除输入图像中的异常区域的问题^[8-10]. 即在输入异常样本时, 即使编码器得到的特征向量和正常分布相差较大, 解码器得到的重构图像依然近似于输入, 与输入图像的差异并不明显, 导致这些异常样本存在被误判成正常样本的可能性. 为了解决异常样本重构效果不明显的问题, GONG 等^[11]提出的记忆强化自编码器在自编码器的特征空间上增加了记忆模块, 将训练过程中最具有代表性的特征向量存储在模块中, 并在重建时通过注意力机制的方式对特征向量进行处理, 使重构

基于统计模型的异常检测方法, 通常利用统计模型来构建图像中像素值或特征向量的正常分布模式, 对于测试图像则通过与分布的距离判定异常与否. 使用高斯分布模型进行描述是较为常见的方法. 在高光谱图像处理应用中, REED 等^[6]将背景建模为高斯分布, 并从图像场景中随机选择像素来代表背景统计数据从而检测异常. 然而, 单一高斯分布模型对更为复杂的场景建模能力有限, 为了避免这一劣势, VERACINI 等^[7]通过高斯混合模型对高光谱图像的背景像素进行正常模型建模, 并通过在贝叶斯框架内开发的期望最大化算法, 在参数估计过程中自动评估组件数量, 最后使用广义似然比检验来检测异常.

图像的更加接近学习的正常样本分布; 在高光谱图像处理应用中, HADGAN 等^[12]引入了潜在对抗学习、一致性增强表示约束和收缩约束, 使得编码器专注于重构大多数类别(即背景像素), 在原始图像和重构图像的残差图上应用基于能量的空间和基于距离的光谱联合异常检测器, 以获得最终的检测图, 并提出了一种新的基于最小能量的波段选择方法, 用于降低维度并辅助空间检测器进行检测.

近年来, 基于深度学习方法的异常检测技术也采用了基于特征相似度的方法. RUFF 等^[13]提出的深度支持向量数据描述(deep support vector data description, Deep-SVDD) 是这一类型的代表性方法. SVDD 将正常图像样本的特征聚集在人为指定的特征中心点附近, 而由于缺乏对异常样本的学习, 异常样本在特征空间中的特征并不符合正常样本的分布, 往往远离正常样本的特征中心点, 由此根据

检测阶段样本与正常样本特征中心点的距离判断样本是否异常. 文献 [14-15] 结合传统方法和深度学习来进行异常检测, 使用神经网络提取特征向量, 用于描述异常检测的整个图像或用于异常定位的图像子块, 在异常检测阶段利用传统的机器学习方法进行异常检测, 产生了较好的异常检测和定位结果, 但在高光谱处理方面应用仍存在空白.

本文采用了基于特征相似度的方法, 将深度学习和传统方法结合得到异常检测方法. 遵循两阶段框架, 使用自监督预训练的 ResNet-18 网络作为特征提取器, 然后借助高斯密度估计实现对异常图像的分类, 如图 1 所示. 并利用一种基于梯度定位的深度网络可视觉解释方法 GradCAM^[16] 解决了异常图像定位和算法可解释性问题.

1 特征提取模型

1.1 自监督学习方法

在第一阶段, 通过自监督学习从普通数据中学习深度表示.

代理任务对于自监督表示学习至关重要. 虽然在现在已有许多流行的方法构建代理任务, 包括旋转预测^[17] 和对比学习^[18-19], 但简单地应用现有的方法对于检测局部异常帮助较小, 因此需要设计一种增强策略, 创建局部不规则模式, 然后训练模型识别这些局部不规则性, 并希望它能在测试时推广到未知的真实异常. 一种常用的增强方法是 Cut-out^[20], 它可以擦除图像中随机选择的一个小矩形区域. 这个任务似乎很容易通过精心设计的低级图像过滤器来解决, 但由于缺乏先验信息, 深度卷积网络并不会学习这些捷径, 因此可以学习局部异常信息. LI 等^[15] 针对实际缺陷检测任务缺陷一般含有拉伸变形, 特殊的纹理构造, 通过 CutPaste 模拟异常样本.

为了进一步防止学习用于区分增强图像和鼓励模型学习检测不规则性, 如表 1 所示, 模仿 Cut-Paste 的代理任务, 使用如下所示的异常数据生成方法用于自监督训练:

- 1) 从仿真异常图像中随机选择一个样本, 并可选地对其进行旋转或像素值抖动;
- 2) 从正常训练图像中随机选择一个和样本大小相当的小矩形区域;
- 3) 将异常样本粘贴在图像中.

对于生成的样本, 通过特征提取模型进行特征提取, 并连接一层 MLP (multilayer perceptron), 对特征进行二分类. 特征提取和分类模型记为 $g()$. 正常样本 x 的目标标签为 0, 异常样本 x_a 的目标标签为 1, 使用交叉熵 (cross entropy) 损失函数进行训练. 交叉熵函数和损失函数的定义如下:

$$CE(P, Q) = - \sum P(x) \log Q(x) \quad (1)$$

$$L = E\{ CE(g(x), 0) + CE(g(x_a), 1) \} \quad (2)$$

表 1 训练过程的 PyTorch-style 伪代码

Tab. 1 Pseudo code for training in Pytorch style

```
# cross_entropy: 交叉熵损失函数
for x_n in loader:
    # 对输入数据进行数据增强, 构造异常样本
    x_a = augment(x_n)
    # 计算特征
    z_a = f(x_a)
    z_n = f(x_n)
    # 计算损失
    loss = cross_entropy(z_a, 1)
           + cross_entropy(z_n, 0)
    # 优化步
    loss.backward()
    optimizer.step()
```

1.2 Resnet-18 特征提取网络模型

ResNet 是一种用于图像分类的深度卷积神经网络结构^[21]. 其主要特点是使用了残差块 (residual block) 来解决深度神经网络中的梯度消失问题, 从而使得网络可以更深, 更容易训练.

假设输入数据 x 时, $H(x)$ 是神经网络期望的期望输出, 由于学习难度大, 直接把输入 x 作为初始结果, 通过“捷径”传到输出, 使学习目标变为残差 $F(x) = H(x) - x$ 相比 $H(x)$ 更易于学习, 则可以使网络稳定. 残差块的计算方法如式 (2) 和图 2 所示.

$$H(x) = F(x) + x \quad (3)$$

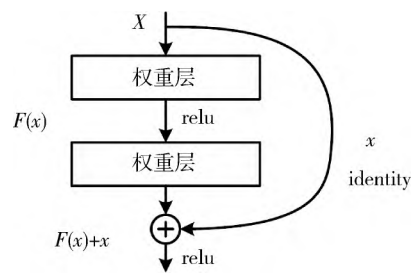


图 2 残差块

Fig. 2 Residual block

近年 Resnet 作为计算机视觉的里程碑工作, 广泛地应用于图像分类、目标检测和予以分割等任务. ResNet18 是 ResNet 系列的一种轻量级结构, 它

是由 18 个卷积层和全连接层组成的深度神经网络，如图 2 所示。相对于 VGG16^[22] 和 AlexNet^[23] 等模

型，它具有更少的参数，更快的训练速度和更好的性能。适用于训练数据较少的异常检测任务。

层名称	输出维度	18-层	34-层	50-层	101-层	152-层
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	平均池化, 1000-d fc, softmax				
FLOPs		1.8×10 ⁹	3.6×10 ⁹	3.8×10 ⁹	7.6×10 ⁹	11.3×10 ⁹

图 3 ResNet 模型网络结构

Fig. 3 Network architecture of Resnet models

2 异常评分和定位方法

2.1 异常评分方法

在第二阶段，评估输入数据的深度表示与正常数据的离群程度来得到异常评分。

目前存在各种通过单类分类器计算异常分数的方法，生成式分类器包括 KDE、GDE 等。在本文

中，在特征 f 上构建生成高斯密度估计器 GDE^[24] 以作为分类器，其对数密度计算如下：

$$\log p_{\text{GDE}}(x) \propto \left\{ -\frac{1}{2} (f(x) - \mu)^T \Sigma^{-1} (f(x) - \mu) \right\} \quad (4)$$

其中 μ 是训练集特征向量的中心， Σ 是训练集特征向量分布的对称协方差矩阵，二者均从正常数据中建模得到。

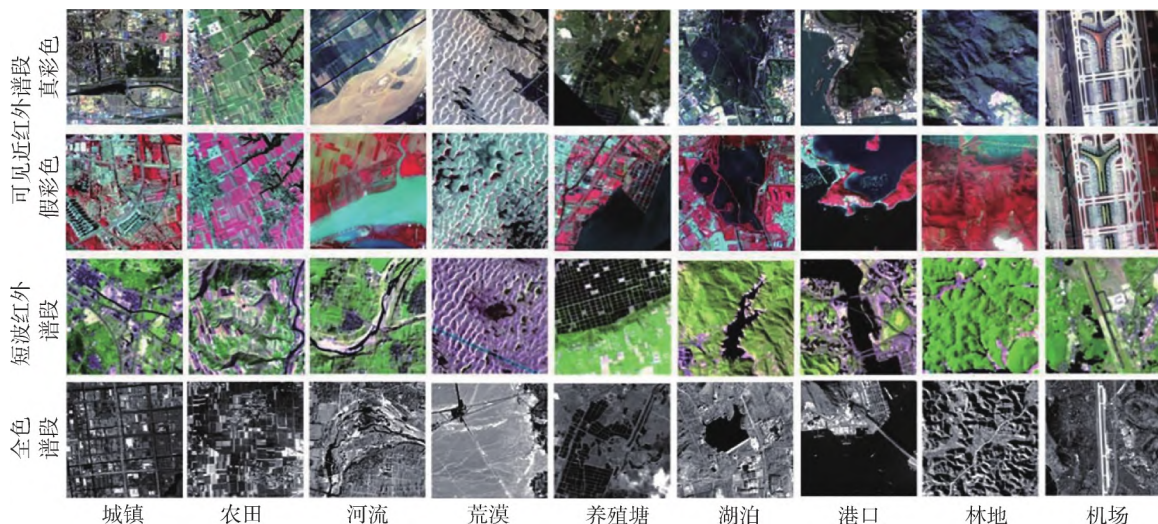


图 4 TG1HRSSC 数据集样本数据可视化展示

Fig. 4 Sample display of TG1HRSSC dataset

计算样本的密度函数后，可以通过设置分类阈值的方法来判断样本是否异常。如图 5 所示，在蓝色圆形分类阈值内的样本被判定为正常，标记为蓝色圆点，而在分类阈值外的样本被判定为异常，标记

为红色 × 号。

2.2 异常定位方法

在异常检测中，当模型通过预测任务学习图像的整体表征时，很难像生成模型那样通过重构误差

进行异常定位. 因此, 使用一种基于梯度定位的深度学习网络可视觉解释方法 Grad-CAM^[16] 来突出显示影响异常检测模型决策的区域.

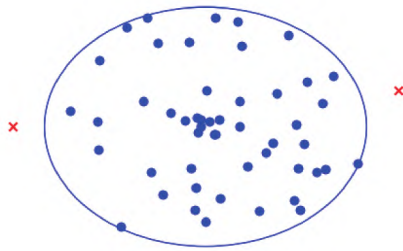


图 5 异常评分方法可视化展示

Fig. 5 Display of anomaly score method

Grad-CAM 使用 CNN 模型的最后一个卷积层的梯度信息为每个神经元分配权重, 以进行定位决策.

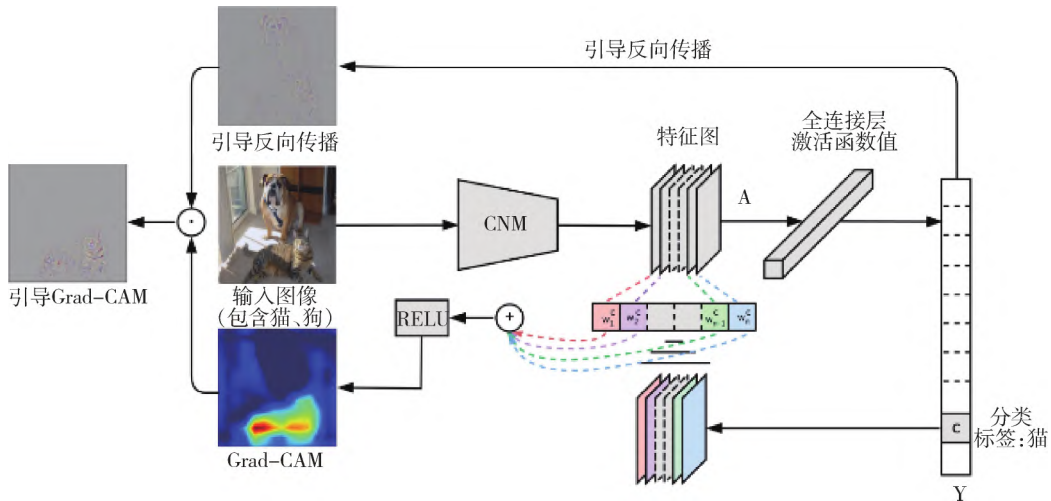


图 6 Grad-CAM 算法流程

Fig. 6 Overview of Grad-CAM algorithm

3 实验及结果

3.1 数据集

本文基于天宫一号高光谱遥感场景分类数据集^[25] 在测试集构建仿真异常数据, 对异常检测模型性能进行了训练和评估.

天宫一号高光谱遥感场景分类数据集 (TG1HRSSC) 包括单波段的 5 m 分辨率的全光谱段 54 个有效波段的 10 m 分辨率可见近红外波段和 52 个有效波段的 20 m 分辨率短波红外波段. 数据集包括 9 个典型场景, 包括城镇、机场、养殖塘、河流、港口、湖泊、农田、林地和荒漠. 数据集展示图像如图 4 所示, 其中彩色图像由选择波段作为 RGB 通道的图像合成, 灰度图像由单波段图像构成. 可见

其算法流程如图 6 所示, 记特征图的高度为 u , 宽度为 v , 类别为 c , 为计算定位结果, 首先计算 c 类分类得分 y^c 的梯度 $\partial y^c / \partial A_k$, 其中 A_k 为特征图 k 通道的激活值. 对梯度 $\partial y^c / \partial A_k$ 在宽度和高度上平均池化得到权重 α_k^c , 用其对最后一层特征图加权线性组合后, 进行 ReLU 激活函数处理, 得到

$$L_{\text{Grad-CAM}}^c \in R^{u \times v} = \text{RELU}(\sum_k \alpha_k^c A_k) \quad (5)$$

在计算梯度反向传播时, 计算 α_k^c 的过程涉及对激活函数的梯度进行连续的权重矩阵与激活函数梯度之间的矩阵乘积, 一直到梯度传播到最后一个卷积层为止. 因此, 权重 α_k^c 代表了从 A 开始的深度网络的部分线性化, 并获取了特征图 k 对于目标类别 c 的“重要性”.

近红外波段中的近似真彩色图像使用 29、19 和 6 波段合成, 假彩色图像使用 52、29 和 19 波段合成; 短波红外波段的假彩色图像使用 23、11 和 37 波段合成; 使用全光谱段的灰度图像使用其单波段构成. 实验中, 仿照数据集可视化方法, 选用可见近红外波段数据真彩色波段通道 29、19 和 6 组合, 选用短波红外波段数据假彩色波段通道 23、11 和 37 组合, 分别用于合成 RGB 通道格式的图像, 并按各波段亮度峰值进行归一化操作, 符合计算机视觉领域广泛使用的卷积神经网络学习所需输入图像数据格式.

由于城镇、机场和养殖塘等 3 类地物样本数量有限, 较难满足深度神经网络对训练的数据量的需求, 且场景与作为待检测异常的雷达样本有所冲突, 因此本文参考数据集工作^[25], 选取河流、港口、湖泊、农田、林地和荒漠等 6 类可将雷达视为异常场

景的场景类别进行模型训练和测试,并使用 90°、180°和 270°旋转,上下、左右和上下左右翻转进一步增广样本数据至原始数据量的 7 倍,并剔除了部分传感器导致成像脏污的样本数据. 实验中,使用计算机生成的军事目标可见光仿真图像和网络收集的卫星拍摄图像作为构建异常样本的对象,包括雷达、飞机、车辆和坦克等类别. 如图 7 所示. 实验训练集与测试集划分比例 8:2,共记训练图像 415 张,测试图像 103 张,其中异常测试图像 41 张.



图 7 用于构造异常样本的异常图像样本展示
Fig. 7 Sample display of anomaly image for anomaly sample construction

3.2 实验方法

遵循单分类检测的传统过程,即在正常数据上进行训练,得到特征提取器后,使用正常数据特征构建到单分类器,参考文献 [26],使用 ResNet-18 作为骨干网络,连接一个 3 层 MLP 投影层来学习表示.

训练模型的输入图像尺寸为 100 × 100,使用动量 SGD 优化模型参数,学习率为 0.03,动量参数为 0.9, batch size 大小为 64,训练 500 epochs. 使用余弦学习率衰减计划和 L2 权重正则化,系数为 0.000 03. 在进行异常数据生成时,对异常样本应用随机旋转和颜色抖动进行数据增强,以增强特征表示的鲁棒性.

对 ResNet 模型采用 3 种训练策略:

- 1) 直接在训练集上训练得到权重;
- 2) 通过微调在大规模图像分类数据集 ImageNet^[27] 预训练的权重,在前 100 epochs 同时训练 ResNet 和 MLP,在 100 ~ 400 epochs 冻结 ResNet,训练 MLP;
- 3) 冻结 ImageNet 预训练权重,只训练 MLP.

训练完成后,在测试集上进行验证,以检验自监督训练的有效性.

3.3 实验结果

由于异常检测本质上是一个单分类问题,传统机器学习中两类问题中的指标可以用来衡量异常检测模型的检测效果. 然而,由于数据集的正常样本远多于异常样本,准确率和召回率均不适合用于作为异常检测的评价指标. 但可以绘制接受者操作

特征曲线 (receiver operating characteristic curve, ROC) 来比较异常检测模型的优劣.

表 2 混淆矩阵
Tab. 2 Confusion matrix

		真实值		总数
		正	负	
预测输出	正	TP	FP	P'
	负	FN	TN	N'
总数		P	N	

如表 2 定义的混淆矩阵所示,每一行表示数据的真正归属类别;每一列中的数值表示真实数据被预测的类别. 真阳性率 (true positive rate, TPR) 和假阳性率 (false positive rate, FPR) 为

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{TN + FP} \quad (7)$$

在接受者改变分类阈值时,以假阳性率 FPR 为横坐标,真阳性率 TPR 为纵坐标绘制的曲线即为 ROC 曲线. ROC 曲线到 x 轴的面积为曲线下面积 (area under curve, AUC) .

实验设置异常评分分类器为 GDE,实验结果及与其他深度学习异常检测方法效果如表 3 所示. 其中本文和 CutPaste 方法可以引入 ImageNet 预训练知识,大幅提升检测准确率,在表 3 中用 P 表示. 本文所提出的异常检测算法可以提供优异的异常检测效果.

表 3 异常检测实验结果
Tab. 3 Experimental results of anomaly detection

异常检测方法	AUC
GANomaly ^[9]	0.656
Deep-SVDD ^[13]	0.455 2
CutPaste ^[15]	0.525 2
CutPaste + P	0.779 7
本文方法	0.872 9
本文方法 + P	0.999 6

异常定位相关结果在图 8 中展示,图 8 中前 4 张可见近红外波段图像分辨率为 256 pixel × 256 pixel,后 6 张短波红外波段分辨率为 128 pixel × 128 pixel,可以看出自监督训练能帮助够定位图像中的异常信息,且有助于发现人类肉眼难以发现的异常.

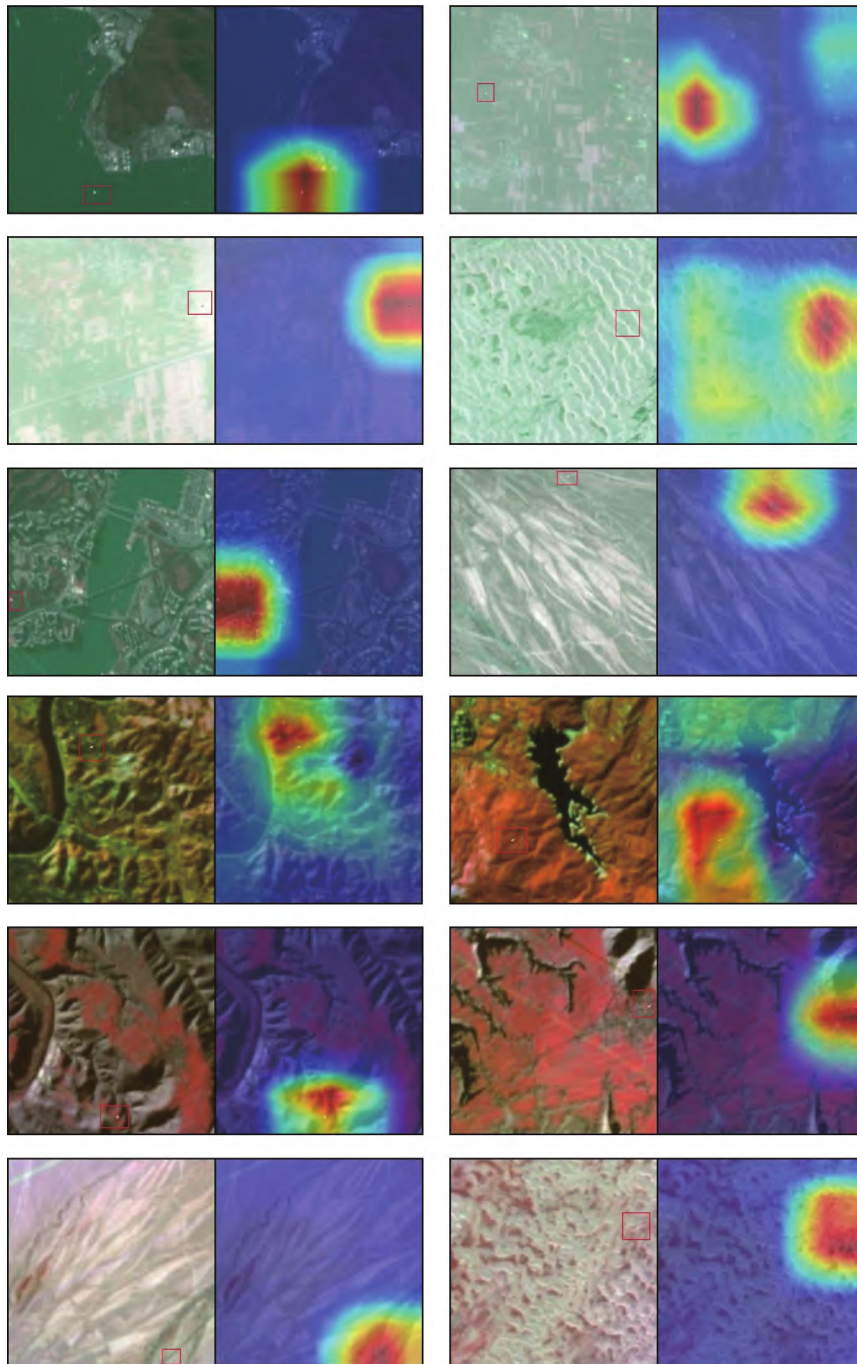


图8 异常定位结果样本展示

Fig. 8 Sample display of anomaly localization

4 结 论

本文在 CutPaste 基础上提出了一种适用于遥感图像的异常检测和定位方法。通过对正常数据和模拟生成的异常数据进行自监督分类,鼓励模型发现异常样本的特性,有效地训练了编码器的表征能力,使建模得到正常样本特征与异常样本产生较大差

距。在仿真数据集上展示出了优秀的图像标签级异常检测性能。直接训练的模型即可得到 AUC 0.872 9 的异常检测模型,引入 ImageNet 预训练知识后对异常检测的 AUC 可达 0.999 6,并通过 Grad-CAM 展示了优秀的异常定位性能。

参 考 文 献

- [1] 李盛阳,刘志文,刘康,等. 航天高光谱遥感应应用研究

- 进展 [J]. 红外与激光工程 2019, 48(3): 9-23.
LI S Y, LIU Z W, LIU K, et al. Advances in application of space hyperspectral remote sensing [J]. Infrared and Laser Engineering, 2019, 48(3): 9-23.
- [2] 璩泽旭, 方火能, 肖化超, 等. 一种基于深度学习的光学遥感影像在轨目标检测方法 [J]. 空间控制技术与应用, 2022, 48(5): 105-115.
QU Z X, FANG H N, XIAO H C, et al. An on-orbit object detection method based on deep learning for optical remotesensing image [J]. Aerospace Control and Application, 2022, 48(5): 105-115.
- [3] 张鹏程, 武文波, 李强, 等. 面向星载边缘计算的遥感目标检测算法轻量化优化研究 [J]. 空间控制技术与应用, 2022, 48(5): 86-94.
ZHANG P C, WU W B, LI Q, et al. Lightweight of remote sensing object detection algorithm for spaceborne edge computing [J]. Aerospace Control and Application, 2022, 48(5): 86-94.
- [4] CHOI J, KIM C. Unsupervised detection of surface defects: a two-step approach [C]//The 19th IEEE International Conference on Image Processing. New York: IEEE, 2012.
- [5] 吕承侃, 沈飞, 张正涛, 等. 图像异常检测研究现状综述 [J]. 自动化学报, 2022, 48(6): 1402-1428.
LYU C K, SHEN F, ZHANG Z T, et al. Review of image anomaly detection [J]. Acta Automatica Sinica, 2022, 48(6): 1402-1428.
- [6] REED I S, YU X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1990, 38(10): 1760-1770.
- [7] VERACINI T, MATTEOLI S, DIANI M, et al. Fully unsupervised learning of Gaussian mixtures for anomaly detection in hyperspectral imagery [C]//2009 9th International Conference on Intelligent Systems Design and Applications. New York: IEEE, 2009.
- [8] HASELMANN M, GRUBER D P, TABATABAI P. Anomaly detection using deep learning based image completion [C]//The 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). New York: IEEE, 2018.
- [9] AKCAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Ganomaly: semi-supervised anomaly detection via adversarial training [C]//Computer Vision-ACCV 2018: 14th Asian Conference on Computer Vision. Perth: Springer International Publishing, 2019.
- [10] YANG H, CHEN Y, SONG K, et al. Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects [J]. IEEE Transactions on Automation Science and Engineering, 2019, 16(3): 1450-1467.
- [11] GONG D, LIU L, LE Y, et al. Memorizing normality to detect anomaly: memory-augmented deep autoencoder for unsupervised anomaly detection [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2019.
- [12] JIANG T, LI Y, XIE W, et al. Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(7): 4666-4679.
- [13] RUFF L, VANDERMEULEN R, GOERNITZ N, et al. Deep one-class classification [C]//International Conference on Machine Learning. Stockholm: PMLR, 2018.
- [14] DEFARD T, SETKOV A, LOESCH A, et al. Padim: a patch distribution modeling framework for anomaly detection and localization [C]//Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event. Cham: Springer International Publishing, 2021.
- [15] LI C L, SOHN K, YOON J, et al. Cutpaste: self-supervised learning for anomaly detection and localization [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2021.
- [16] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization [C]//Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017.
- [17] GIDARIS S, SINGH P, KOMODAKIS N. Unsupervised representation learning by predicting image rotations [J]. ArXiv Preprint ArXiv:1803.07728, 2018.
- [18] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations [C]//International Conference on Machine Learning. Stockholm: 2020.
- [19] YE M, ZHANG X, YUEN P C, et al. Unsupervised embedding learning via invariant and spreading instance feature [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019.
- [20] DEVRIES T, TAYLOR G W. Improved regularization of convolutional neural networks with cutout [J]. ArXiv Preprint ArXiv:1708.04552, 2017.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016.
- [22] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint ArXiv:1409.1556, 2014.
- [23] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Im-

- genet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [24] RIPPEL O, MERTENS P, MERHOF D. Modeling the distribution of normal data in pre-trained deep features for anomaly detection [C]//2020 25th International Conference on Pattern Recognition (ICPR). New York: IEEE, 2021.
- [25] 刘康, 周壮, 李盛阳, 等. 天宫一号高光谱遥感场景分类数据集及应用 [J]. *遥感学报*, 2020, 24(9): 1077-1087.
- LIU K, ZHOU Z, LI S Y, et al. Scene classification dataset using the Tiangong-1 hyperspectral remote sensing imagery and its applications [J]. *Journal of Remote Sensing*, 2020, 24(9): 1077-1087.
- [26] SOHN K, LI C L, YOON J, et al. Learning and evaluating representations for deep one-class classification [J]. *ArXiv Preprint ArXiv: 2011. 02578*, 2020.
- [27] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge [J]. *International Journal of Computer Vision*, 2015, 115: 211-252.

作者简介: 曹哲骁(2000—), 男, 博士研究生, 研究方向为人工智能、计算机视觉; 傅瑶(1983—), 女, 副研究员, 研究方向为光学、遥感成像处理; 王丽(1989—), 女, 工程师, 研究方向为图像检测、跟踪算法等; 苏盈(1987—), 女, 高级工程师, 研究方向为光电系统工程技术、伺服控制技术等; 郭云翔(1993—), 男, 工程师, 研究方向为光电系统工程技术、嵌入式等; 王田(1987—), 男, 副教授, 研究方向为模式识别与智能系统。

An Anomaly Detection Method for Remote Sensing Image Based on Deep Learning Network

CAO Zhexiao^{1,2}, FU Yao^{3*}, WANG Li⁴, SU Ying⁴, GUO Yunxiang⁴, WANG Tian^{1,2,5}

1. Beihang University, Beijing 100191, China
2. State Key Laboratory of Complex & Critical Software Environment, Beijing 100191, China
3. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
4. Wuhan Guide Infrared Co. Ltd, Wuhan 430205, China
5. Zhongguancun National Laboratory, Beijing 100191, China

Abstract: A high-performance anomaly detection model has been constructed to address the problem of sparse anomalous image data in the real world. A two-stage framework anomaly detection model is built using only normal training data and a small amount of synthetic anomaly sample. First, a ResNet-18 encoder model is trained to extract representation by the pretext of classifying normal data and synthetic anomaly data. Then, a single classifier for anomaly images is built through modelling the distribution of normal data representations using Gaussian density estimation. GradCAM is applied to extend the model, enabling the anomaly detection model to locate anomaly regions without labels. Finally, experiments are conducted on a simulated anomaly detection dataset using real-world images, demonstrating that the proposed algorithm can detect anomaly and provide location results in remote sensing images that are even difficult to recognize with human eyes.

Keywords: anomaly detection; remote sensing; deep learning; convolutional neural network

Received: 2023-02-22; Accepted: 2023-06-27

Foundation item: Supported by National Natural Science Foundation of China (61972016 and 62032016) and Beijing Nova Program (20220484106 and 20230484451)

* Corresponding author. E-mail: fuy_cas@foxmail.com