



## 基于Faster R-CNN的多任务增强裂缝图像检测方法

毛莺池, 唐江红, 王静, 平萍, 王龙宝

引用本文:

毛莺池, 唐江红, 王静, 等. 基于Faster R-CNN的多任务增强裂缝图像检测方法[J]. 智能系统学报, 2021, 16(2): 286–293.

MAO Yingchi, TANG Jianghong, WANG Jing, et al. Multi-task enhanced dam crack image detection based on Faster R-CNN[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(2): 286–293.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201910004>

## 您可能感兴趣的其他文章

### 基于深度学习的空间非合作目标特征检测与识别

Feature detection and recognition of spatial noncooperative objects based on deep learning

智能系统学报. 2020, 15(6): 1154–1162 <https://dx.doi.org/10.11992/tis.202006011>

### 基于注意力融合的图片描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

### 一种基于级联神经网络的飞机检测方法

Cascade convolutional neural networks for airplane detection

智能系统学报. 2020, 15(4): 697–704 <https://dx.doi.org/10.11992/tis.201908028>

### 注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

### 基于改进的Faster R-CNN高压线缆目标检测方法

Object detection of high-voltage cable based on improved Faster R-CNN

智能系统学报. 2019, 14(4): 627–634 <https://dx.doi.org/10.11992/tis.201905026>

### 基于Faster R-CNN的人体行为检测研究

Research on human behavior detection based on Faster R-CNN

智能系统学报. 2018, 13(6): 967–973 <https://dx.doi.org/10.11992/tis.201801025>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201910004

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20200717.1727.008.html>

# 基于 Faster R-CNN 的多任务增强裂缝图像检测方法

毛莺池, 唐江红, 王静, 平萍, 王龙宝

(河海大学计算机与信息学院, 江苏南京 211100)

**摘要:** 针对 Faster R-CNN 算法对多目标、小目标检测精度不高的问题, 本文提出一种基于 Faster R-CNN 的多任务增强裂缝图像检测 (Multitask Enhanced Dam Crack Image Detection Based on Faster R-CNN, ME-Faster R-CNN) 方法。同时提出一种基于 K-means 的多源自适应平衡 TrAdaBoost 的迁移学习方法 (multi-source adaptive balance TrAdaBoost based on K-means, K-MABtrA) 辅助网络训练, 解决样本不足问题。ME-Faster R-CNN 将图片输入 ResNet-50 网络提取特征; 然后将所得特征图输入多任务增强 RPN 模型, 同时改善 RPN 模型的锚盒尺寸和大小以提高检测识别精度, 生成候选区域; 最后将特征图和候选区域发送到检测处理网络。K-MABtrA 方法利用 K-means 聚类删除与目标源差别较大的图像, 再在多元自适应平衡 TrAdaBoost 迁移学习方法下训练模型。实验结果表明: 将 ME-Faster R-CNN 在 K-MABtrA 迁移学习的条件下应用于小数据集大坝裂缝图像集的平均 IoU 为 82.52%, 平均精度 mAP 值为 80.02%, 与相同参数设置下的 Faster R-CNN 检测算法相比, 平均 IoU 和 mAP 值分别提高了 1.06% 和 1.56%。

**关键词:** 裂缝图像检测; Faster R-CNN; 多任务检测; 小目标检测; 迁移学习; 大坝安全; 区域建议网络; 小样本中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2021)02-0286-08

中文引用格式: 毛莺池, 唐江红, 王静, 等. 基于 Faster R-CNN 的多任务增强裂缝图像检测方法 [J]. 智能系统学报, 2021, 16(2): 286-293.

英文引用格式: MAO Yingchi, TANG Jianghong, WANG Jing, et al. Multi-task enhanced dam crack image detection based on Faster R-CNN[J]. CAAI transactions on intelligent systems, 2021, 16(2): 286-293.

## Multi-task enhanced dam crack image detection based on Faster R-CNN

MAO Yingchi, TANG Jianghong, WANG Jing, PING Ping, WANG Longbao

(College of Computer and Information, Hohai University, Nanjing 211100, China)

**Abstract:** To improve the accuracy of the detection of multiple small targets using the faster R-CNN model, we propose a multi-task enhanced dam-crack image detection method based on faster R-CNN (ME-Faster R-CNN). In addition, to solve the problem of insufficient dam-crack samples, we propose a transfer learning method, multi-source adaptive balance TrAdaBoost based on K-means (K-MABtrA), to assist with network training. In the ME-Faster R-CNN, the ResNet-50 network is adopted to extract features from original images, obtain the feature map, and input a multi-task enhanced region-proposal-network module to generate candidate regions by adopting the appropriate size and dimensions of the anchor box. Lastly, the features map and candidate regions are processed to detect dam cracks. The K-MABtrA method first uses K-means clustering to delete unsuitable images. Then, models are trained using the multi-source adaptive balance TrAdaBoost method. Our experimental results show that the proposed ME Faster R-CNN with the K-MABtrA method can obtain an 82.52% average intersection over union (IoU) and 80.02% mean average precision (mAP). Compared with Faster R-CNN detection method using the same parameters, the average IoU and mAP values was increased by 1.06% and 1.56%, respectively.

**Keywords:** crack image detection; Faster R-CNN; Multi-task detection; small targets detection; transfer learning; dam safety; RPN; small sample

收稿日期: 2019-09-15. 网络出版日期: 2020-07-18.

基金项目: 国家重点研发课题 (2018YFC0407105); 国家自然科学基金重点项目 (61832005); 国网新源科技项目 (SGTYHT/19-JS-217); 华能集团重点研发课题 (HNKJ19-H12).

通信作者: 唐江红. E-mail: [15195897810@163.com](mailto:15195897810@163.com).

我国是世界上拥有水库大坝最多的国家<sup>[1]</sup>, 但随着时间的推移和坝龄的增长, 大坝表面和内部发生形变, 出险几率增加, 威胁人民生命财产安全。裂缝是大坝的主要危害之一。

近年来,图像处理、模式识别和深度学习等技术的发展,为大坝裂缝图像检测提供技术支持。但由于大坝环境复杂等一系列因素的限制,导致裂缝图像收集和标记成本过高,因此难以获得大坝裂缝图像检测的分类模型。迁移学习主要是针对规模不大,样本数量有限的特定领域数据集使用机器学习容易产生过拟合而导致无法训练与学习的问题,通过利用具有一定相似性的领域中已训练好的较好优秀模型和样本构建满足任务需求的模型,从而实现小数据集下构建良好模型的效果。

Faster R-CNN<sup>[2]</sup>是目前基于区域卷积神经网络系列的目标检测算法中综合性能最好的方法之一,但其对多目标、小目标情况检测精度不高。本文提出了一种基于Faster R-CNN的多任务增强裂缝图像检测方法,以适应大坝在不同光照环境、不同长度裂缝情况下的检测。同时提出了一种基于K-means多源自适应平衡TrAdaBoost迁移学习方法辅助网络训练,解决样本不足问题。

## 1 相关工作

根据卷积神经网络的使用方式,将基于CNN目标检测算法<sup>[3-4]</sup>分为两大类:基于区域建议的深度学习目标检测算法和基于回归思想的深度学习目标检测算法。前者的主流算法有:R-CNN算法<sup>[5]</sup>、Fast R-CNN算法<sup>[6]</sup>和Faster R-CNN算法。R-CNN首次将神经网络应用在目标检测算法上,在Pascal VOC 2012的数据集上将平均精度mAP提升了30%。Fast R-CNN将候选框识别分类和位置回归合成到一个网络中,不再对网络进行分步训练,提高了训练速度。Faster R-CNN与Fast R-CNN最大的区别就是提出了区域建议网络(region proposal networks, RPN)网络,极大地提升了检测框的生成速度。基于回归思想的深度学习目标检测的主流算法有:SSD算法<sup>[7]</sup>和YOLO V2算法<sup>[8]</sup>。SSD算法和YOLO算法均没有区域建议过程,极大地提高了检测速度,但识别精度和位置回归精度不足。

从20世纪90年代起迁移学习开始逐渐进入机器学习领域,受到研究者的关注。常用的迁移学习方法有AdaBoost<sup>[9]</sup>和TrAdaBoost<sup>[10]</sup>算法等。AdaBoost算法基本思想:当一个训练样本被错误分类时候,对此样本增加样本权重,再次训练时该样本分错的概率就会大大降低。TrAdaBoost算法是由AdaBoost算法演变而来的,该算法通过降低误分类的源域训练数据权重,增加误分类的目标域训练数据权重,使得分类面朝正确的方向移动并训练出强分类模型。Al-Stouhi等<sup>[11]</sup>总结TrAdaBoost算法存在的问题,在此基础上进

行改进,提出一种动态TrAdaboost(dynamic TrAdaboost, DtrA)方法,DtrA方法能够在迭代过程中动态调整样本权重;郭勇<sup>[12]</sup>在DtrA方法基础上进一步改进,提出一种自适应TrAdaBoost(adaptive TrAdaBoost, AtrA)方法,AtrA方法能够反映出源领域训练数据集与目标领域训练数据集之间是否具有相似性关系。

## 2 ME-Faster R-CNN 与 K-MABtrA 方法

本文提出了一种基于Faster R-CNN的多任务增强裂缝图像检测的网络模型,以适应大坝在不同光照环境、不同长度裂缝情况下的检测。同时,提出了一种基于K-means多源自适应平衡TrAdaBoost迁移学习方法解决样本不足问题。本文采用基于K-MABtrA迁移学习方法训练ME-Faster R-CNN网络模型,该方法通过已准备的多源裂缝图像数据集对卷积神经网络的参数进行预训练,然后使用预训练得到的网络权重作为初始权值,迁移到目标数据集上进行微调,得到适用于大坝裂缝检测的模型。

### 2.1 ME-Faster R-CNN 模型

ME-Faster R-CNN在Faster R-CNN模型基础上进行改进,改进之处如图1所示。其中,特征提取部分:选取轻量级的ResNet-50作为卷积神经网络;特征融合部分以及候选区域生成部分:改进使用多任务增强RPN模型,改善锚盒尺寸大小提高Faster R-CNN搜索能力,提高检测识别精度;检测处理部分:特征图和选择区域建议经过感兴趣区域(ROI)池、全连接(FC)层分别发送给边界回归器和SVM分类器得到分类与回归结果。

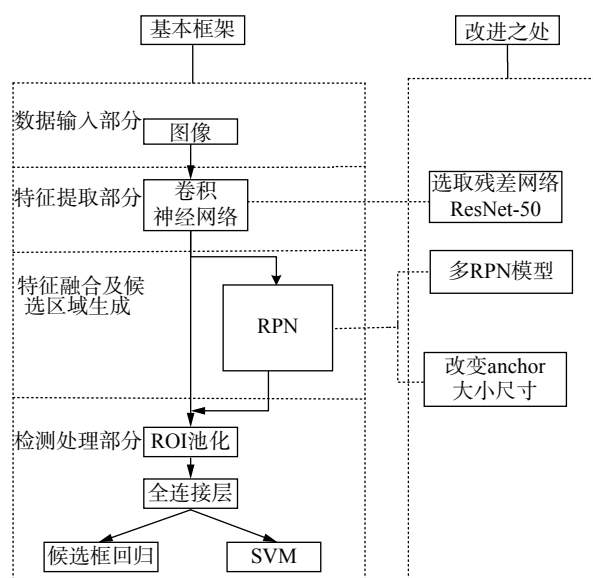


图1 ME-Faster R-CNN 模型改进之处

Fig. 1 Improvements of ME-Faster R-CNN

ME-Faster R-CNN 检测流程主要分为 3 个部分, 分别是特征提取、特征融合以及候选区域生成、检测处理。

1) 特征提取: 本文选用 ResNet-50 深度残差网络<sup>[13]</sup>作为大坝裂缝图像特征提取器, 通过 5 级 ResNet-50 将图片转换成特征图。

2) 特征融合以及候选区域生成: 将所得特征图输入多任务增强 RPN 模型, 并改善 RPN 模型的锚盒尺寸和大小以提高检测识别精度, 最后生成候选框。具体方法如下:

①多任务增强 RPN 方法: 最初 Faster R-

CNN 模型结构中只有一个 RPN, RPN 使用最后一个卷积层获得特征图<sup>[2]</sup>。称之为原始 RPN, 其结构如图 2(a) 所示, 输入图像大小为  $224 \times 224$ , 原始 RPN 在网络中感受野要远远大于  $224 \times 224$ , 仅能获得少量典型裂缝特征。然而, 图像中裂缝存在不同大小和比例。如果检测到裂缝大小对于检测区域太大, 则检测区域周围多余裂缝形状可能会被视为噪音。如果检测到裂缝大小对于检测区域太小, RPN 将无法生成 ROI。因此, 原始 RPN 的功能不足以检测不同大小和比例的完整裂缝对象。

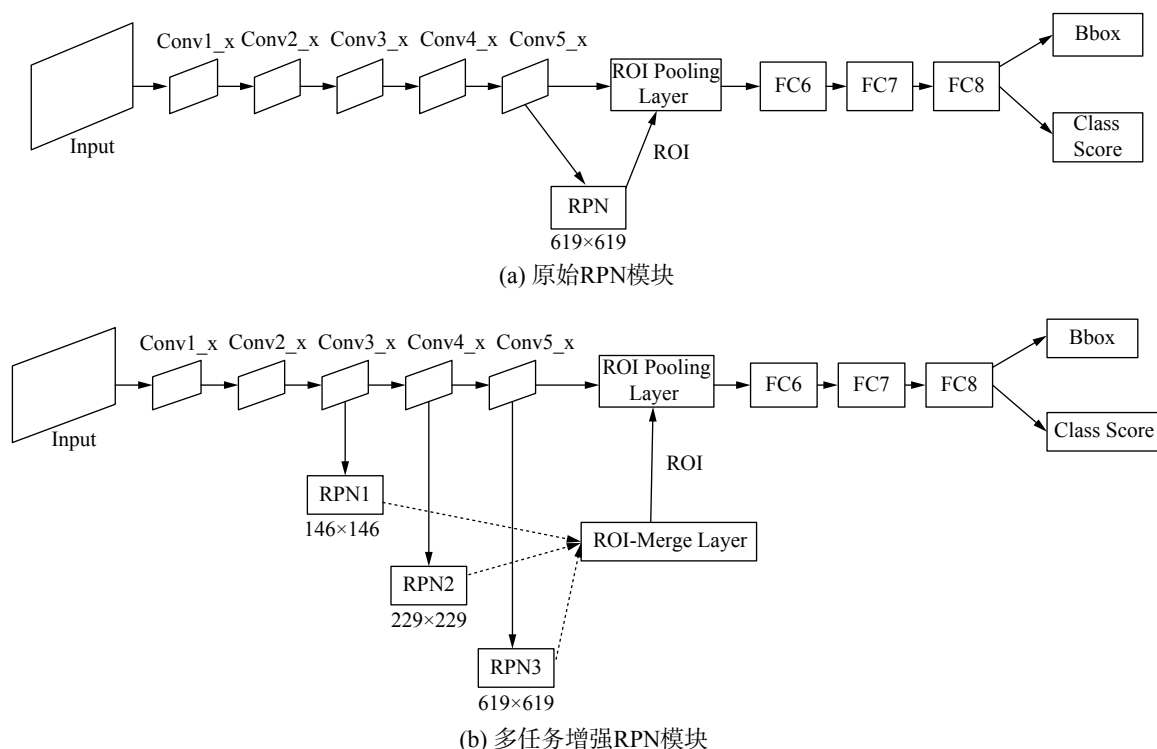


图 2 原始 RPN 模型与多任务增强 RPN 模型

Fig. 2 Primitive RPN module & multi-task enhanced RPN module

针对以上问题, ME-Faster R-CNN 方法提出一种多任务增强 RPN 方法, 其结构如图 2(b) 所示。该方法在 ResNet-50 的基础上引入多个 RPN 来产生 ROI, 提取不同大小特征图。具体是在 ResNet-50 的第 3 卷积层 Conv3\_x 后加入一个 RPN 模块, 其感受野大小为  $146 \times 146$ , 用来检测较小目标; 同时在 ResNet-50 的第 4 卷积层 Conv4\_x 后加入一个 RPN 模块, 其感受野大小为  $229 \times 229$ , 用来检测较大目标; 在 ResNet-50 的第 5 卷积层 Conv5\_x 之后利用多 RPN 任务可以输出图像总体信息。

由于每个 RPN 输出独立的 ROI 数组, 为聚集和选择有效的区域, 多任务增强 RPN 方法提供 ROI-Merge Layer 用于接受独立 ROI 数组, ROI-

Merge Layer 仅输出一个数组。为了避免重复的 ROI 和低的 ROI 裂缝似然分数, 本文使用非极大值抑制方法, 不同卷积层后 RPN 输出的候选区域中, 在对应位置两 ROI 的交并比大于 0.7 的 ROI 为同一 ROI。具体方法为, 3 个卷积层后 RPN 输出的候选区域均带有建议得分, 该分数对应的是目标的可能性, 在对应位置选取分数最高的一个 ROI 区域, 另外两层对应位置的 ROI 与所选 ROI 交并比 IoU 若大于 0.7, 则认为是同一 ROI, ROI-Merge Layer 的输出数组中对应位置仅输出该得分最高的数组。在使用非极大值抑制方法之后, 选择前 100 个值较高的 ROI。因此, ROI-Merge Layer 只需要调整超参数即可控制 ROI 的



数量。

②改善 RPN 模型的锚盒尺寸和大小: Faster R-CNN 模型经过卷积层提取特征图, 然后特征图输入 RPN 区域进行特征融合以及生成候选区域, 此时特征图上每个像素点映射不同比例宽度锚点, 每个锚点放置若干个不同大小的锚, 为解决不同尺度间隔的锚盒搜索能力不平衡的问题, 本文设计了一种新型锚盒, 新型锚盒尺度为:  $50 \times 50$ 、 $200 \times 200$ 、 $350 \times 350$  和  $500 \times 500$ , 其中,  $50 \times 50$  和  $200 \times 200$  适用于较小的裂缝检测,  $350 \times 350$  和  $500 \times 500$  适用于较大的裂缝检测。这 4 种尺度每种尺寸按 1:1, 1:2, 2:1 的长宽比例缩放, 共 12 种尺度作为 RPN 需要评估的候选框, 在预测时候选框的顺序是固定的。RPN 的目标就是对原图中的每个锚点对应的 12 个框, 预测其是否是一个存在目标的框。框与真值框的  $\text{IoU} > 0.7$  就认为这个框是一个候选框, 反之, 则不是。

## 2.2 K-MABtrA 方法

针对大坝裂缝图像较少, 训练样本分布不平衡, 以及 TrAdaBoost 算法在训练过程中易削弱辅助数据集作用的问题, 本节提出一种基于 K-means 的多源自适应平衡 TrAdaBoost 的迁移学习方法 K-MABtrA, 使用迁移学习充分利用多个领域的大量相关训练集, 动态调整样本权值, 训练出网络的强分类器, 提高大坝裂缝检测的准确率。如图 3 所示为基于 K-means 的多源自适应平衡 TrAdaBoost 迁移学习的过程, 主要分为两个阶段: K-means 图像聚类方法<sup>[14]</sup>和多源自适应平衡 TrAdaBoost 迁移学习。

1) K-means 图像聚类方法: 通过 K-means 图像聚类方法, 利用欧式距离将图像进行聚类排序。将聚类距离远的图片从裂缝图像库中删除, 有利于后续分类器的训练, 提高训练效率。K 代表聚类质心数目, means 表示簇内数据的均值。K-means 图像聚类方法具体步骤如下:

①首先将图像库中图像  $X_i (i = 1, 2, \dots, n)$  进行灰度化, 依次存储到一维矩阵  $D_x$  中;

②接着以 10 像素长度, 3 像素移动步长依次进行分块存储, 记录每小块的首位置, 得到  $n$  个像素块数据集, 从中任意选择 30 个图像小块的灰度均值作为初始聚类中心;

③根据每个图像矩阵小块的灰度均值, 利用欧几里得距离, 如式 (1) 所示, 计算这些对象与 30 个图像样本聚类中心的距离; 并根据最小距离

重新对相应图像小块灰度均值进行划分, 将每个图像矩阵小块赋给最相近的类;

$$\text{dis}(x_i, y_j) = \left[ \sum_{k=1}^d (x_{ik} - x_{jk})^2 \right]^{\frac{1}{2}}, i, j = 1, 2, \dots, n \quad (1)$$

式中:  $\text{dis}(x_i, y_j)$  为数据对象  $x_i$  和  $y_j$  之间的距离。该值越大, 说明  $x_i$  和  $y_j$  越相似; 反之  $x_i$  和  $y_j$  差距越大。

④重新计算每个有变化的图像小块像素灰度均值的质心;

⑤重复上述步骤 3)、4) 直至各个数据类的集合中心不再发生变化为止。

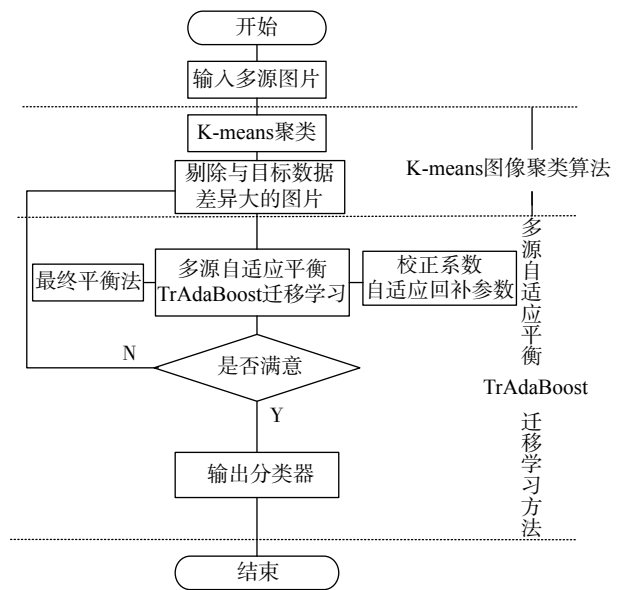


图 3 K-MABtrA 方法流程图

Fig. 3 Flow chart of K-MABtrA method

2) 多源自适应平衡 TrAdaBoost 迁移学习: 利用不同领域裂缝图像和真实大坝裂缝图像——组合进行训练, 生成基分类器; 在 TrAdaBoost 基础上引入校正系数<sup>[12]</sup>, 避免由于迭代次数的增加, 导致源领域权重下降过快, 与目标源领域权重之间差距过大的问题; 在校正系数中引入自适应回补参数<sup>[13]</sup>, 反映源领域训练数据集与目标领域训练数据集之间是否具有相似性关系, 提高方法检测性能; 最后, 使用最终平衡权重法, 使最终得到的目标源数据集与各领域裂缝数据集重要度一致。

①增加校正系数更新源领域样本的权值

迁移学习在训练过程在, 各领域辅助训练集随着迭代次数增加得到的权值不断减小以至于与目标数据集不相关, 无法起到辅助目标数据集学习的作用。为了更好地利用各领域辅助训练集和目标数据集训练, 在 TrAdaBoost 基础上增加校正系数更新源领域样本的权值。当迭代次数  $m$  不

断增大,各个领域辅助训练集都能被正确回归,当  $m$  次迭代结束后,各个辅助领域样本权值之和为

$$S_n = \sum_{i=1}^{n_a} w_{a_i}^{m+1} = \sum_{i=1}^{n_a} w_{a_i}^m = n_a w_a^m \quad (2)$$

式中:  $n_a$  为辅助训练集  $a$  中样本个数;  $w_a^m$  为  $a$  中各训练样本权重。

目标数据集  $b$  中预测样本正确的样本权值不变,  $n_b$  为目标数据集  $b$  中样本个数,  $w_b^m$  为  $b$  中训练样本权重,  $\varepsilon_b^m$  为弱分类器在  $b$  上的错误率,正确样本的权值之和为

$$S_{b_1} = n_b w_b^m (1 - \varepsilon_b^m) \quad (3)$$

目标数据集  $b$  中预测错误样本需要更新  $\varphi_m$ ,  $\varphi_m = \frac{\varepsilon_b^m}{1 - \varepsilon_b^m}$ , 则  $b$  中错误样本的权值之和  $S_{b_2}$  为

$$S_{b_2} = n_b w_b^m \varepsilon_b^m \varphi_m^{-1} = n_b w_b^m (1 - \varepsilon_b^m) \quad (4)$$

所有目标域样本权值之和,即正确样本和错误样本权值之和:

$$S_b = S_{b_1} + S_{b_2} = 2n_b w_b^m (1 - \varepsilon_b^m) \quad (5)$$

当  $m+1$  次迭代的辅助数据集样本权值分布为

$$w_a^{m+1} = \frac{w_a^m}{S_a + S_b} = \frac{w_a^m}{n_a w_a^m + 2n_b w_b^m (1 - \varepsilon_b^m)} \quad (6)$$

当迭代次数足够大时,各领域辅助训练集都能被正确回归,迭代结束后,  $w_a^{m+1} = w_a^m$ , 联系式 (6) 可得:

$$S_a + S_b = n_a w_a^m + 2n_b w_b^m (1 - \varepsilon_b^m) = 1 \quad (7)$$

设辅助数据集样本增加校正系数为  $C^m$ , 其权值变为

$$w_a^{m+1} = C^m w_a^m / (C^m S_a + S_b) = C^m w_a^m / [C^m n_a w_a^m + 2n_b w_b^m (1 - \varepsilon_b^m)] \quad (8)$$

由于此时辅助数据集样本权值稳定不变,即  $w_a^{m+1} = w_a^m$ , 根据关系式 (7) 和 (8) 可得校正系数为

$$C^m = 2(1 - \varepsilon_b^m) \quad (9)$$

从式 (9) 可以看出,校正系数  $C^m$  与弱分类器在目标数据集  $b$  上的错误率  $\varepsilon_b^m$  负相关,辅助数据集样本权值增加,对下一次迭代训练弱分类器的影响增加;反之对下一次迭代训练弱分类器的影响减小。因此,在 TrAdaBoost 算法基础上加入校正系数  $C^m$  能够同时保持目标数据集和辅助数据集样本权值得到收敛。

## ② 引入自适应回补参数

然而,即使  $\varepsilon_b$  较低时,弱分类器对源领域训练集的分类效果也会存在差异,这种差异同样可以反映出源领域训练集与目标领域训练集之间的相关性。为了反映这种相似性关系,在校正系数中引入自适应回补参数,自适应回补参数为基分类器在辅助数据集和目标数据集上的分类正确率

之和,即

$$B^m = (1 - \varepsilon_b^m) + (1 - \varepsilon_a^m) \quad (10)$$

第  $m+1$  次迭代后辅助领域数据样本权值:

$$w_a^{m+1} = \frac{B^m w_a^m \alpha^{|f(x_i) - y_i|}}{S_a + S_b} = \frac{B^m w_a^m \alpha^{|f(x_i) - y_i|}}{B^m n_a w_a^m + 2n_b w_b^m (1 - \varepsilon_b^m)} = \frac{B^m w_a^m \alpha^{|f(x_i) - y_i|}}{B^m w_a^m \alpha^{|f(x_i) - y_i|}} \quad (11)$$

## ③ 最终平衡权重法

最后,循环达到设定迭代次数  $M$ ,得到强学习器。但在迭代后,目标数据集与源数据集的权重已经严重偏离,所以,造成最终分类器也过于偏向目标小数据集的问题。针对上述问题,在最终分类器生成方式中引入最终平衡权重法。最终平衡权重法的基本概念是:在迭代过程中,源数据权重不断下降,目标数据权重不断增加,迭代结束后,源数据权重与目标数据权重之间差距较大,但在最终分类器生成形式上,将目标数据集最终权重重置为最后一次迭代中各领域辅助训练集权重的平均值,使最终得到的目标源数据集与各领域辅助训练集要度一致,提高算法的检测准确率。

## 3 实验与结果

### 3.1 数据集介绍

目前公开的大坝裂缝图像数据库较少,为了实现对大坝裂缝图像检测与识别,从大坝日常监测过程以及 Google 图像搜索引擎中收集并整理已标记好的裂缝图像组建成数据库。该数据库包含大坝、公路、混凝土墙壁和桥梁 4 个领域裂缝图像,其中大坝裂缝图像 635 张,其他领域裂缝图片每个领域各 2500 张,总计 8135 张裂缝图片。

### 3.2 实验结果与分析

本次实验根据选取数据集的特点,选取 mAP<sup>[15]</sup> (mean average precision) 和检测评价函数交并比<sup>[16]</sup> (intersection over union, IoU) 作为目标检测算法的评价指标。mAP 作为目标检测中用于衡量识别精度指标;IoU 表示感兴趣区域和标定区域的重叠率。

本实验主要从以下 4 个方面对基于 Faster R-CNN 参数迁移的裂缝图像检测训练方法的优劣进行对比分析:

#### 1) 视觉对比分析

在裂缝检测过程中,对裂缝图像提取感兴趣区域,并对感兴趣区域进行裂缝特征提取,通过训练好的分类器进行裂缝检测识别后,每个感兴趣区域边框都会得到一个分数,即置信度。随机选取 3 组实验结果进行视觉对比分析如图 4 所

示。以图4(a)为例,其中,原图中央有一条长度和开合度明显的裂缝痕迹,其下方有一条短且开合度不明显的裂缝痕迹。Faster R-CNN能够检测出长度和开合度明显的裂缝,在其下方的裂缝并未准确检测出来。而ME-Faster R-CNN模型不仅提高IoU重叠度,更为准确地检测出长裂缝,同时能够准确检测出下方小裂缝痕迹,做到不误检也不漏检。实验结果表明,在相同的实验条件下,ME-Faster R-CNN方法不仅能提高检测精度,而且在应对目标小、多目标情况时,能获得很好的检测效果。

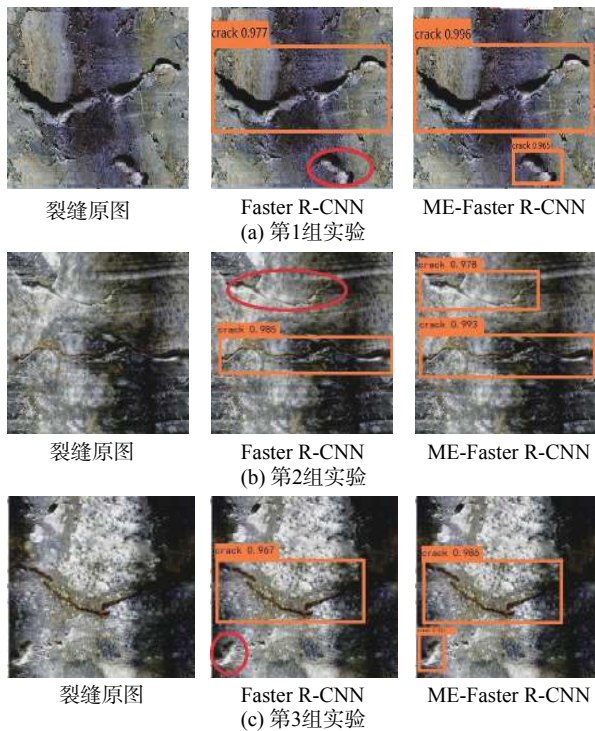


图4 视觉对比分析

Fig. 4 Visual contrast analysis

## 2) 与不同基准网络模型之间对比分析

该部分采用ZF网络<sup>[17]</sup>、VGG-16<sup>[18]</sup>网络、Res-

Net-50和ResNet-101网络作为特征提取基准网络,Faster R-CNN作为目标检测模型进行实验。

由表1可得,在同样数据集的训练测试下,ZF-Net可以达到66.51%的mAP值,VGG-16网络可以达到71.90%的mAP值,而ResNet网络的mAP值可以超过78%,提高了6个百分点,网络的检测准确度得到提高,由表1可以看出ResNet-50比ResNet-101的检测精度略低,但是ResNet-50的参数量为ResNet-101参数量的1/2,能够有效减少网络权重数量,加速模型训练<sup>[14]</sup>,综合训练速度与检测精度本文选取ResNet-50为基准模型。

表1 不同基准网络模型的准确度

Table 1 Accuracy of different baseline network models

#	基准网络模型	mAP/%
1	ZF-Net	66.51
2	VGG-16	71.90
3	ResNet-50	78.46
4	ResNet-101	80.00

## 3) 与不同目标检测算法对比分析

该部分以ResNet-50网络为基准网络,以SSD算法、YOLO V2算法、Faster R-CNN算法和ME-Faster R-CNN算法作目标检测模型进行实验。

表2给出不同目标检测算法获得的平均IoU、召回率、准确度以及平均精度。其中,ME-Faster R-CNN算法的平均IoU是最高的,表明ME-Faster R-CNN算法在裂缝位置检测的准确性方面更优异一些,且其mAP值也是最大的,达到80.02%,表明ME-Faster R-CNN检测模型的综合性能很好。

表2 不同目标检测算法对比分析

Table 2 Comparison of different target detection algorithms

目标检测算法	平均IoU/%	召回率/%	准确度/%	mAP/%
SSD算法	81.11	74.25	77.63	75.94
YOLO V2算法	73.96	75.39	79.07	77.23
Faster R-CNN算法	81.46	77.06	79.86	78.46
ME-Faster R-CNN算法	82.52	76.53	83.51	80.02

此外,注意到裂缝尺寸大小对准确度也会存在一定影响。因此,将采集到的真实大坝裂缝图像根据其尺寸大小分为3组。第1组包含100个

样本,其尺寸大小在[0,50]范围内,第2组包含100个样本,其尺寸大小在[50,200]范围内,第3组包含100个样本,其尺寸大小超过200像素。



各目标检测算法在不同尺寸裂缝图像的准确度如图 5 所示。

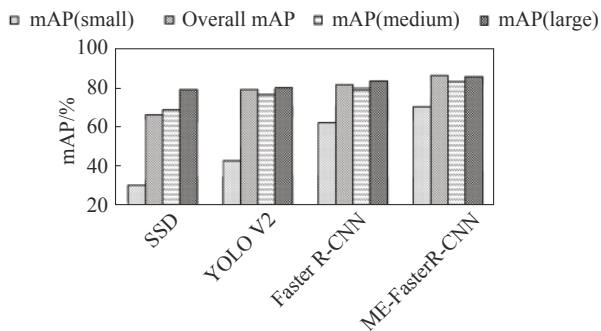


图 5 不同尺寸裂缝图像的准确度

Fig. 5 Accuracy of crack images with different sizes

从图 5 可以看出, Faster R-CNN 系列模型检测的准确度要整体优于 SSD 算法和 YOLO V2 算法, 所有的检测算法在较大裂缝图像上都能表现得最好, 而在小裂缝图像的检测上, 准确度却不是很高。Faster R-CNN 算法和 ME-Faster R-CNN 算法在较大裂缝图像检测性能上实力相当, 而在小裂缝图像的检测上, ME-Faster R-CNN 算法要更优于 Faster R-CNN 算法。综上所述, ME-Faster R-CNN 算法在保持一定准确度的基础上, 在面对小目标检测难度较大的情况, 也能获得很好的效果。

#### 4) 迁移学习对比实验

该部分以 ResNet-50 作为基准网络, ME-Faster R-CNN 作为目标检测模型, 目标数据集样本占源训练集样本的比例为  $r$ ,  $r$  取 2%、5% 和 10%, 分别用 K-MABtrA 方法、ATrA 方法、DTrA 方法、TrAdaBoost 迁移学习方法进行分类器训练。

由表 3 可得出, 同一方法, 不同比例  $r$  下训练得到的分类器, 在一定的范围内随着比例  $r$  不断增加, 分类器的各评价指标都有所提升, 说明在一定范围内目标源数据占总数据比例越大, 分类器的各评价指标越高, 检测效果越好。在目标源数据所占比例  $r$  相同情况下: ATrA 和 DTrA 方法各评价指标均高于 TrAdaBoost, 证明了引入校正系数和自适应回补参数的有效性。本文提出的 K-MABtrA 方法各指标均高于 ATrA 方法。说明 K-MABtrA 方法引入最终平衡权重法, 使最终得到的目标源数据集与各领域裂缝数据集重要度一致, 提高算法的检测准确率。综上所述, K-MABtrA 方法能够更多地利用其他领域的共享信息, 得到更好的迁移学习效果, 训练出强分类器, 高效地完成大坝裂缝图片的检测任务。

表 3 不同迁移学习方法对比分析  
Table 3 Comparison of different transfer learning methods

迁移学习方法	$r$	召回率	准确率	mAP
TrAdaBoost	2	54.73	55.19	54.96
	5	64.94	66.35	65.65
	10	66.18	71.29	68.74
DTrA	2	60.31	62.37	61.34
	5	65.93	70.13	68.12
	10	71.26	75.41	73.34
ATrA	2	64.95	71.64	68.12
	5	71.13	76.81	73.97
	10	74.28	81.17	77.73
K-MABtrA	2	70.09	78.01	74.05
	5	72.81	79.73	76.27
	10	76.53	83.51	80.02

## 4 结束语

经实验验证, 本文提出的 ME-Faster R-CNN 方法在多目标、小目标检测准确性方面更优异; 且本文提出的迁移学习方法更有效地解决了样本不足的问题。本文所提出方法的局限性在于 ME-Faster R-CNN 相比于 Faster R-CNN 只在特定检测任务中准确度较高, 比如本文的应用场景: 大坝裂缝检测, 或类似的检测任务。而对于目标大小相似、亮度相同的目标检测其结果与 Faster R-CNN 所差无几。

## 参考文献:

- [1] 苏南. 我国 200 米级高坝密集, 安全风险不可轻视 [EB/OL]. (2017-11-09). [https://www.thepaper.cn/newsDetail\\_forward\\_1858088](https://www.thepaper.cn/newsDetail_forward_1858088).
- [2] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge, USA, 2015: 91–99.
- [3] HABER E, RUTHOTTO L, HOLTHAM E, et al. Learning across scales—a multiscale method for convolution neural networks[C]//Proceedings of the 23rd AAAI Conference on Artificial Intelligence. 2017.
- [4] GERBER D, MEIER S, KELLERMANN W. Efficient target activity detection based on recurrent neural networks[C]//Proceedings of 2017 Hands-free Speech



- Communications and Microphone Arrays. San Francisco, USA, 2017: 46–50.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA, 2014: 580–587.
- [6] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440–1448.
- [7] LIU WEI, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands, 2016: 21–37.
- [8] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 6517–6525.
- [9] KANG H H, LEE S W, YOU S H, et al. Novel vehicle detection system based on stacked DoG kernel and AdaBoost[J]. *PLoS one*, 2018, 13(3): e0193733.
- [10] DAI Wenyuan, YANG Qiang, XUE Guirong, et al. Boosting for transfer learning[C]//Proceedings of the 24th International Conference on Machine Learning. New York, USA, 2007: 193–200.
- [11] AL-STOUHI S, REDDY C K. Adaptive boosting for transfer learning using dynamic updates[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Berlin, Germany, 2011: 60–75.
- [12] 郭勇. 基于单源及多源的迁移学习方法研究[D]. 西安: 西安电子科技大学, 2013.
- GUO Yong. Research of transfer learning based on single-source and multi-source[D]. Xi'an: Xidian University, 2013.
- [13] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 770–778.
- [14] WICAKSONO Y A, RIZALDY A, FAHRIAH S, et al. Improve image segmentation based on closed form matting using K-means clustering[C]//Proceedings of 2017 International Seminar on Application for Technology of Information and Communication. Semarang, Indonesia, 2018: 26–30.
- [15] AKCAY S, KUNDEGORSKI M E, WILLCOCKS C G, et al. Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery[J]. *IEEE transactions on information forensics and security*, 2018, 13(9): 2203–2215.
- [16] RAHMAN M A, WANG Yang. Optimizing intersection-over-union in deep neural networks for image segmentation[C]//Proceedings of 12th International Symposium on Advances in Visual Computing. Las Vegas, USA, 2016: 234–244.
- [17] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//Proceedings of 13th European Conference on Computer Vision. Zurich, Switzerland, 2014: 818–833.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of 3rd International Conference on Learning Representations. San Diego, USA, 2015.

#### 作者简介:



毛莺池, 教授, 博士, 博士生导师, 主要研究方向为云计算和边缘计算、分布式技术和物联网数据分析。曾获大禹水利科学技术奖一等奖; 华能集团科技进步奖二等奖; 江苏省科学技术奖三等奖; 2018年度江苏省计算机学会优秀科技工作者。发表学术论文50余篇。



唐江红, 硕士研究生, 主要研究方向为图像处理。



王静, 硕士研究生, 主要研究方向为图像处理。