

基于 Mask-RCNN 迁移学习的红外图像电力设备检测

刘子全¹, 付 慧², 李玉杰¹, 张国江², 胡成博¹, 张照辉¹

(1. 国网江苏省电力有限公司电力科学研究院, 南京 211103; 2. 国网江苏省电力有限公司, 南京 210000)

摘 要: 红外图像诊断是电力系统故障诊断的重要方式,但目前仍依靠人工辅助框图来实施图像中目标的检测。为提升检测效率,本文借鉴并改进在目标分割任务中表现优异的 Mask-RCNN 方法,利用图像自动语义分割识别红外图像中的一个或多个电力设备,并提取设备轮廓。为了缓解标注样本相对不足的问题,研究 Mask-RCNN 的迁移学习机制,设计并实现了训练数据重要性采样、参数迁移映射等方法,使改进后的方法适应于红外图像电力设备检测任务。在实际采集数据集上的实验表明,改进后的算法能在仅有少量像素级标注样本的条件下,较好地提取出电力设备的轮廓,并进一步识别出设备类别。所提模型和算法为进一步的设备分区和故障区域检测提供了精确有效的预处理手段。

关键词: 红外图像; 电力设备检测; 卷积神经网络; 图像分割; 迁移学习

中图分类号: TM73 **文献标志码:** A

Electrical Equipment Detection in Infrared Images Based on Transfer Learning of Mask-RCNN

LIU Ziquan¹, FU Hui², LI Yujie¹, ZHANG Guojiang², HU Chengbo¹, ZHANG Zhaohui¹

(1. Electric Power Research Institute of State Grid Jiangsu Power Co. Ltd., Nanjing 211103, China; 2. State Grid Jiangsu Power Co. Ltd., Nanjing 210000, China)

Abstract: Infrared fault image recognition is an important method to diagnose electrical equipment, but the recognition relies on the manually created bounding boxes over objects. In this paper, in order to improve the detection efficiency, automatic semantic segmentation of infrared images is investigated to recognize one or more electrical equipment objects. The proposed method is based on Mask-RCNN which has demonstrated good performance on instance segmentation. Our main contribution is applying transfer learning to Mask-RCNN, where importance sampling and parameter mapping are conducted to alleviate the data-shortage problem on pixel-level annotating. Experimental results on real-world datasets have shown that the improved version of Mask-RCNN is able to extract the shapes of electrical equipment, even with limited data with pixel-level annotations. The proposed algorithm provides an efficient way to the subsequent steps of fault region detection and classification.

Key words: infrared images; electrical equipment detection; convolutional neural network; semantic segmentation; transfer learning

引 言

保持电力设备安全稳定的运行是电力系统运维的首要任务。常年运行的设备可能出现各种故障,因此需要定位、检测和维修。红外检测方式无需直接接触设备,可在不停电的状态下发现故障类型,在系统运维的安全性、可靠性和经济性等方面具有明显优势。当前技术通常基于故障设备的温度会发生变化这个事实来实施红外成像检测,通过故障状态图像和正常状态图像的对比来实施定位和检测。

目前,红外图像中电力设备的检测主要以人工诊断的方式来实施。该方式需要操作人员手工在图像上画框进行分区,以定位电力设备及其重要部件。这种方式效率较低,对操作人员的专业知识及经验要求较高,容易出错。随着人工智能技术的快速发展,用智能算法取代人工操作成为可能,这需要首先完成红外图像中电力设备的检测任务。

考虑到变电设备种类繁多,为了完成设备区域的检测,可首先运用图像分割技术自动将红外图像分区,然后采用目标特征提取和分类方法对电力设备分类和识别,最后利用各自区域的温度数据完成故障诊断^[1]。

因此,本文研究利用深度学习技术对红外图像中的电力设备进行自动语义分割和目标检测,将设备的轮廓提取出来,便于后续的区域定位和识别。

为了实现图像分割、配准、识别等处理的自动化,很多研究者将计算机视觉等机器学习技术用于处理电力设备红外图像,这些工作在提高检测效率、降低运维成本等方面提供了非常有价值的思路。

目标检测是计算机视觉和模式识别的重要研究方向,主要是确定图像中是否有感兴趣的目标存在,并对其进行探测和精确定位。传统的目标检测流程采用滑动窗口进行区域选择,然后采用 SIFT、HOG 等方法进行特征提取,最后采用 SVM、Adaboost 进行类别判断。文献[2]研究改进了变电站设备红外图像的分割方法,首先对预处理后的红外图像做 K-means 聚类,然后采用最大熵法对聚类后的图像进行分割,在完整地分割实验数据集中目标的任务中取得了较好的效果。文献[3]考虑到设备的形状特征,采用多种不变矩统计量来提取和表示这种形状特征,然后采用支持向量机分类器完成图像分类,得到了较高的目标识别率。

传统的目标检测方法面临的主要问题有:特征提取鲁棒性差,不能反映光照变化、背景多样等情况;区域选择没有针对性,时间复杂度高,窗口冗余。此外,传统机器学习方法能处理的目标类别数目通常较少,且性能相对于近年来出现的深度学习技术有较大差距。因此,本文工作建立在最新的深度学习技术的基础上。

深度学习技术在计算机视觉领域的人脸识别、智能驾驶、场景分类等任务中获得了非常广泛且有效的应用。基于深度学习的目标检测模型有 Faster RCNN、Yolo 和 Yolo2、SSD 等框架。近年来,电网中大量的传感以及检测设备产生规模巨大的红外图像,为使用深度学习构建自动分类器提供了数据基础。深度学习的一些典型网络结构,如深度神经网络(Deep neural network, DNN)、卷积神经网络(Convolutional neural network, CNN)等在电力系统自动故障区域检测领域也有了新的应用^[4-5],并取得了较好的效果。

上述工作集中于电力设备目标检测方面,即检测红外图像中有何目标,并用矩形框将目标框出来。在这些工作中,目标检测和识别是两个分离的过程,识别要依赖于目标检测的结果^[6]。因此,在目标检测阶段的误差会积累到后续识别中,带来负面影响。

为了进一步提高目标检测的效果,本文考虑实体分割,即让算法直接标出每个像素所属的类别^[7]。本文工作将深度学习目标分割和识别方面的最新技术 Mask regional convolutional neural network (Mask-RCNN)^[8-9]应用于红外图像的语义分割和目标识别,将二者置于统一的优化框架下,整体提高红外图像中

目标检测和识别的效果。对电气设备进行分割和分类,将为后续的故障检测提供有力的依据。

本文首次将Mask-RCNN用于红外图像电力设备的目标检测,并针对该场景做了小样本条件下的迁移学习,在迁移参数映射、迁移样本设计等方面做了相应的改进,最终在框出电力设备目标的同时完成目标分类。这为后续的故障分区和识别提供了良好的基础。

1 基于Mask-RCNN的目标检测和语义分割

本文拟采用的Mask-RCNN是一个灵活通用的目标分割和识别框架。该网络不仅可对图像中的目标区域进行检测,还可以对每一个目标给出一个高质量的分割结果。Mask-RCNN是在Faster RCNN^[10]基础上的扩展,在目标区域框(Bounding box)分支的并列位置添加一个用于预测目标掩模(Object mask)的新分支,如图1所示。

该网络主要包括两个关键网络:蓝色区域所示的Faster RCNN和红色区域所示的实体分割网络。此外,还有候选区域投票和对齐、施加掩模等步骤。最终,所有网络和步骤涉及的参数都通过一个统一的目标函数来优化。

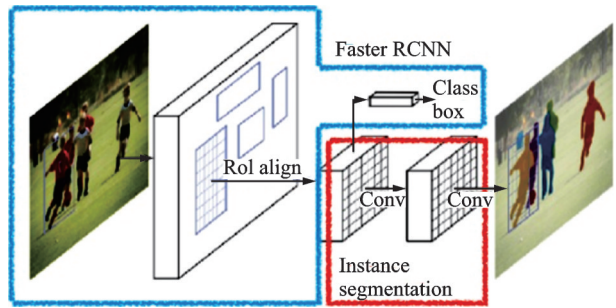


图1 Mask-RCNN 概念图

Fig.1 The conceptual illustration of Mask-RCNN

1.1 Faster RCNN 网络结构

Faster RCNN 将区域生成网络(Region proposal networks, RPN)和Fast RCNN 整合到一个端到端的网络中,提高了目标检测的速度及精确度。

首先向共享的卷积特征模块提供输入图片。一方面提供RPN网络的特征提取信息,另一方面继续向前传播,生成特征图供ROI pooling 模型使用。针对RPN网络的输入特征图,通过滑动窗口方法生成锚点(Anchors)和相应的特性信息,然后生成区域建议和区域得分,并对区域得分采用非极大值抑制方法得到建议区域,给ROI池化层用。将上述两部得到的特征图和区域建议同时输入ROI池化层,提取对应区域的特征。对建议区域的特征信息,输出分类得分以及回归后的Bounding box。

1.2 实体分割网络

实体分割网络采用ResNet50/101+FPN的网络结构,其中FPN为特征金字塔网络(Feature pyramid networks)^[11]。在ResNet50/101网络中,使用了ResNet中Stage 2、Stage 3、Stage 4、Stage 5的特征图^[12]。基于上述4个阶段的特征图,可构建FPN结构。在目标检测里面,低层的特征图信息量比较少,但是特征图比较大,目标位置准确,所以容易识别一些小物体;高层特征图信息量比较丰富,但是目标位置比较粗略。

Mask-RCNN是一种特殊的FPN网络结构,该结构包含一个自底向上的线路、一个自顶向下的线路以及对应层的链接。两个线路中各层生成和上层特征相同大小的特征后与上层特征的对应该元素相加,生成新的特征图。

1.3 候选区域对齐和投票

在卷积特征图上,用较小尺度的窗口(如3*3)执行卷积操作。对特征图中的每个中心点,选取 k 个不同尺度的锚点,然后按照原尺度映射回原图,生成候选区域(Region proposals)。借助于特征低维编码及其卷积操作,分别判断每个锚点属于目标的概率以及它所对应物体的坐标信息。一般可以采用如

下规则:

- (1) 假如某 ROI 与任一目标区域的 IoU 最大,则该锚点判定为有目标;
- (2) 假如某 ROI 与任一目标区域的 $\text{IoU} > 0.5$,则判定为有目标;
- (3) 假如某 ROI 与任一目标区域的 $\text{IoU} < 0.5$,则判定为无目标。

其中 IoU 为预测框和真实框的覆盖率,其值等于两个框所代表区域的交集除以两个区域的并集。

把原图的左上角和右下角的候选区域映射到特征图上的两个对应点,并基于图像的缩放比例进行映射,然后在特征图中把归一化的坐标映射为特征图中的坐标,就生成了特征图中对应的 ROI 区域坐标。

1.4 施加掩模

Mask-RCNN 通过设置掩模(Mask)部分的损失函数,使对于一个属于第 k 个类别的 RoI,仅仅与第 k 个掩模的输出做对比,而不考虑其他的掩模输出。这种损失函数的定义对像素区域不会引入类间竞争,从而允许对每个类别都在该区域生成掩模,有利于得到完整的区域分割。特别是在多个目标有重叠、有些目标只观测到部分区域的情况下,对提取多个完整目标区域更加有利。

Mask-RCNN 能在有效检测图片中感兴趣目标的同时,生成每个目标高质量的分割掩模,从而达到对图像进行语义分割的效果。根据掩模对应的物体类别,同时完成目标的识别任务。

1.5 总体代价函数

在 Mask-RCNN 训练时,对每个采样 RoI 的多任务损失函数定义为

$$L = L_{\text{cls}} + L_{\text{box}} + L_{\text{mask}} \quad (1)$$

式中: L_{cls} 为分类器损失, L_{box} 为边框回归损失, L_{mask} 为掩模分割损失。 L_{cls} 和 L_{box} 分支设置与 Faster RCNN 一致; L_{mask} 分支和其他的分类分支一样,使用全卷积网络输出作为输入,输出为以 Sigmoid 为激活函数的 k 个类别的 L_{mask} 分支。通过上述目标函数设置,避免了类间掩模对同一区域标注的竞争,将分类的任务交给专业的 L_{cls} 分支。而 L_{mask} 对于每一个像素使用二值的 Sigmoid 交叉熵损失来度量。

2 样本缺乏条件下的迁移学习

电力设备红外图像中关于目标的标注多是人工加框标注的数据,无法直接满足 Mask-RCNN 的训练需求。由于现有的目标实例分割要求所有训练实例都必须标记有像素级的分割掩模,使得注释新类别的成本十分昂贵。因此,本文在上述场景下研究样本缺乏条件下的迁移学习,首先根据与电力设备目标的相似性,对预训练数据库 ImageNet 中的图像样本进行重采样;再利用重采样后的数据集对网络参数进行预训练;然后制作少量带有像素级掩模标注的电力设备图像;接下来利用这些图像完成 Mask-RCNN 的调优;最后在完成目标检测的同时实现目标分割。

2.1 基于相似度重采样的迁移样本数据集构建

为了评估来自 ImageNet 源域训练样本与任务所在的电力设备图像目标域样本的相似性,首先需要建立目标相似度的评估函数。综合考虑准确度和复杂度,本文采用度量方式为

$$\text{sim}(I_1, I_2) = e^{-\|f(I_1) - f(I_2)\|_2^2} \quad (2)$$

式中: I_1 和 I_2 分别表示训练样本集和目标样本集中的任意两幅图像; $f(I_1)$ 表示从图像提取的特征矢量,此处用灰度共生矩阵来计算;两幅图像的相似度取为灰度共生矩阵距离的负指数次方。

容易看出,式(2)所得数值介于 0 和 1 之间,可解释为训练样本与目标样本的相似概率。依据此概率分布对训练数据集抽样,可得到抽样后的训练数据集。此数据集中,相似度较高的样本有较大概率被多次抽中;相似度较低的样本出现的频度则显著降低。

2.2 借助 OpenCV 生成少量掩模标注图像

现有的电力设备红外图像数据多为施加了矩形框标注的样本,该类样本适用于 Faster RCNN 网络的有监督学习,从而预测目标矩形框的位置和类别,如图 2(a)所示。然而,Mask-RCNN 还需要预测矩形框中每个像素所属的目标类别。为了实施有监督学习,相应的掩模标注信息也需要提供,可用多边形标注出来,如图 2(b)所示。

生成掩模标注后,图像中每个像素的掩模信息一般用 0 或 1 表示,其中 0 表示该像素不属于特定目标区域,1 表示对应像素属于特定目标区域。该掩模标注可通过 OpenCV 的 API 生成。

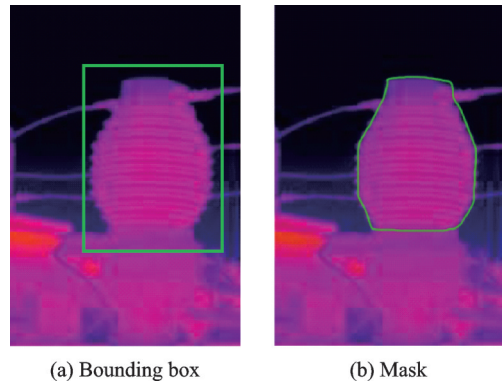


图 2 边框标注和掩模标注的制作

Fig.2 Creating labels of bounding box and mask

2.3 迁移参数映射

Mask-RCNN 非常适合本文的任务,因为它将实例分割问题分解为目标的边界框检测和掩模预测两个子任务。这两个子任务共享网络的底层部分,而底层部分可由 2.1 节生成的样本实施预训练。预训练完成后,边界框网络的顶层参数编码嵌入到每个目标类别,使该类别的视觉信息转移到掩模网络的顶层参数上,如图 3 所示。

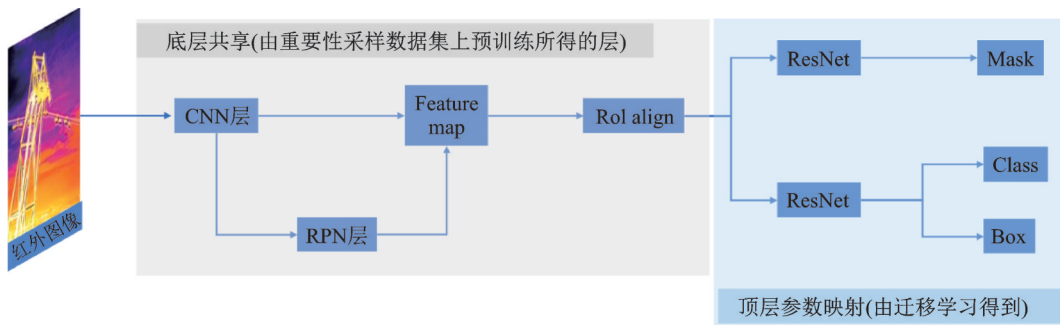


图 3 “底层共享”和“顶层参数映射”的迁移学习示意图

Fig.3 Transfer learning based on “sharing of bottom layers” and “feature mapping of top layers”

由边界框网络到掩模网络的参数编码嵌入可通过一个参数传递函数来实现,其中的未知参数可根据目标边界框和目标掩模的成对数据来训练。由于其参数规模较小,所需数据量较小,从而借助于 2.2 节生成的少量样本,实现由边界框预测到掩模预测的迁移。整个网络训练完成后,在推理时,参数传递函数用于预测每个类别的分割参数,从而使模型能够分割所有目标的类别,包括在训练时没有掩模标注的类别。

3 算法实例和分析

本节首先分析针对电力设备红外图像场景改进后的 Mask-RCNN 算法的效果,然后讨论该算法对后续故障区域检测效果的贡献。

3.1 目标检测和语义分割的效果

本文实验采用了 1.1 节所述的网络结构。预训练在通过 2.1 节所述的相似度重采样后的 ImageNet 数据集上进行,这种相似度实际上是图 3 重要性采样中样本重要程度的一种具体量化,样本构成为 ImageNet 的子集,含有 100 万张图像。如 2.2 节所述,边框到掩模的映射在少量带有掩模标注的数据集上实

施,每个类别用于迁移学习的带掩膜的样本数量为50。测试数据集中每个类别含有100~200个图像样本,共含有3 000张测试图片。

在带有4块Nvidia 1080Ti GPUs的服务器上训练,每个GPU的批处理图像数为2,初始步长取为0.01,动量系数为0.99,经过200 k次迭代可收敛。



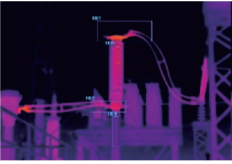
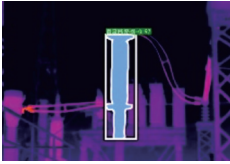
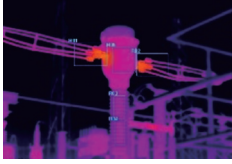

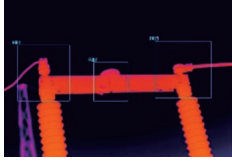
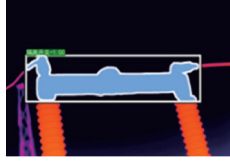


网络推断阶段,采用的“目标候选区域”的数量为300,在这些候选区域上,采用目标框网络分支来预测目标位置。然后,采用Mask网络分支对最高得分的100个检测框进行处理。值得说明的是,此处与训练阶段不同,基于更少、更精确的RoIs,能够加速推断效率,提升预测精度。Mask网络分支对每个RoI预测 K 个掩模区域,但这里只使用第 k 个掩模(其中 k 是分类网络分支预测的类别标签),得到的浮点型掩模输出再重新缩放回RoI的尺寸,并以0.5为阈值进行二值化,最终得到二值分割图像。

掩模预测的平均精确度(Average precision)很难在电力设备红外数据集上进行计算,因为大多数设备只包含了目标边界框的标注,而缺乏像素级掩模的标注。因此本文对算法的评价将采用直观显示的方式,而将平均精确度的评价留作后续工作。Mask-RCNN的设备主体检测效果示例如表1所示。

根据领域知识,可将红外图像中的电力设备按照外形划分为两类大的类别:(1)外形规则的电力设备包含GIS套管、中性点套管、中性点避雷器、套管、断路器、电力电缆、电压互感器、电流互感器、结合滤波器、耦合电容器、避雷器、高压套管等;(2)外形不规则的电力设备包括母线、油枕、油管、电抗器主

表1 设备检测效果

Table 1 Detection results of equipment objects

设备名称	Faster RCNN效果	Mask RCNN效果
直流断路器 (样例1)		
直流断路器 (样例2)		
电流互感器		
隔离开关		
避雷器		

体、电缆、电缆终端、线夹、主体、主变主体、所用变主体。

对于第一大类的电力设备检测算法,采用Mask-RCNN算法实现红外图像中的电力设备检测模块,目前能够识别的电力设备类型有:GIS套管、中性点套管、中性点避雷器、套管、断路器、电力电缆、电压互感器、电流互感器、结合滤波器、耦合电容器、避雷器和高压套管。

这个可视化的结果显示了所提改进版的Mask-RCNN在电力设备红外图像这个缺乏像素级标注的数据集上有较好的表现。该算法具有一定的检测抽象概念的能力,如圆柱、长条等。该算法对孤立的设备能进行合理的分割,共存组成更复杂的物体时,该模型可能会分割失败。该算法在辨识物体的整体和部分时,具有较好的效果,能把整个物体形状预测出来。综合来看,借助于迁移学习技术改进的Mask-RCNN,在训练样本缺乏情况下分割出物体形状,具有很好的潜力。

3.2 所提算法对后续故障检测效果的影响

针对红外图像,在完成拍摄主体检测任务的基础之上,根据复杂程度将设备分区划分为分区规则简单和分区规则复杂两类,分别设计红外分区算法。针对待检测的红外图像的算法运行如下流程:

- (1) 使用基于Mask-RCNN的目标检测算法估计当前图像拍摄主体设备的位置;
- (2) 在设备检测框内,使用姿态估计算法估计当前设备的所需分区的中心点;
- (3) 针对每个分区的中心点,使用回归算法估算当前分区的分区框尺寸。

由上述步骤所得的主要类型设备分区算法效果统计如表2所示。

表2 分类效果对比图

Table 2 Comparison on the classification performance

设备名称	设备占比/%	平均准确率/%
隔离开关	24.80	97.3
避雷器	11.89	96.7
断路器	11.34	94.3
电流互感器	10.80	92.1
电压互感器	7.40	93.5
套管	6.70	96.2
GIS套管	3.80	95.7
电抗	3.40	89.5
主变主体	3.40	85.3
油枕	2.70	98.3
电力电容器	1.90	86.5
熔断器	0.60	84.5

4 结束语

本文介绍了电力行业红外图像诊断的现状以及传统的图像诊断方法,分析了传统方法效率不高的症结所在。提出了一种利用图像自动分区技术辅助红外图像诊断的方法,方法由Mask-RCNN算法构成,经过测试方法可有效提高红外图像诊断的效率,一定程度上摆脱对人工配准和分区的依赖。

目前,将深度学习应用于电力等传统工业领域仍有许多问题需要解决。本文工作对计算机视觉领域的最新技术扬长避短,初步提高了电力设备红外图像的智能诊断水平。此外,不同场景下数据分布之间的差异,深度模型的运行效率等问题仍需要进一步研究。

参考文献:

- [1] 张浩. 图像识别技术在电力设备在线监测中的应用[D]. 北京: 北京交通大学, 2009.
ZHANG Hao. Application of image recognition techniques on the online surveillance of electrical equipment[D]. Beijing: Beijing Jiaotong University, 2009..
- [2] 任新辉. 基于红外技术的变电站设备识别与热故障诊断[D]. 成都: 西南交通大学, 2016.
REN Xinhui. Equipment identification and thermal faults diagnostic of transformer substations based on infrared technology[D]. Chengdu: Southwest Jiaotong University, 2016.
- [3] 张明. 基于红外图像的变电设备分类及在故障诊断中的应用[D]. 合肥: 合肥工业大学, 2012.
ZHANG Ming. Substation equipment classification and its application in fault diagnosis based on infrared image[D]. Hefei: Hefei University of Technology, 2012.
- [4] 印俊. 面向电力设备图片分类大数据应用的深度学习网络研究[D]. 上海: 上海交通大学, 2015.
YIN Jun. Research on deep learning networks for big data application on the classification of images of electronic equipment[D]. Shanghai: Shanghai Jiaotong University, 2015.
- [5] 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): 703003-0703003.
JIA Xin, ZHANG Jinglei, WEN Xianbin. Infrared faults recognition for electrical equipment based on dual supervision signals deep learning[J]. Infrared and Laser Engineering, 2018, 47(7): 703003-0703003.
- [6] 黄新波, 李文君子, 宋桐, 等. 采用遗传算法优化装袋分类回归树组合算法的变压器故障诊断[J]. 高电压技术, 2016, 42(5): 1617-1623.
HUANG Xinbo, LI Wenjunzi, SONG Tong, et al. Application of bagging-CART algorithm optimized by genetic algorithm in transformer fault diagnosis[J]. High Voltage Engineering, 2016, 42(5): 1617-1623.
- [7] HU Ronghang, DOLLAR P, HE Kaiming, et al. Learning to Segment Everything[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Venice, Italy: [s.n.], 2017.
- [8] Matterport Inc. Mask R-CNN for object detection and segmentation[EB/OL]. (2019-04-01) [2021-01-18]. https://github.com/matterport/Mask_RCNN.
- [9] HE Kaiming, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Venice, Italy: [s.n.], 2017.
- [10] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [11] LIN Tsungyi, DOLLAR P, GIRSHICK R, et al. FPN: Feature pyramid networks for object detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017.
- [12] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. ResNet: Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016.

作者简介:



刘子全(1989-), 通信作者, 男, 工程师, 研究方向: 图像识别技术与电力设备状态评价, E-mail: liu_ziquan_sgo@qq.com。



付慧(1979-), 女, 教授级高级工程师, 研究方向: 电力设备状态评价, E-mail: fuhui@js.sgcc.com.cn。



李玉杰(1990-), 男, 工程师, 研究方向: 开关设备故障诊断与状态评价, E-mail: 1275911032@qq.com。



张国江(1975-), 男, 高级工程师, 研究方向: 电力设备状态评价, E-mail: gjzhang@js.sgcc.com.cn。



胡成博(1984-), 男, 高级工程师, 研究方向: 电力设备在线监测与状态评价, E-mail: 15105168989@163.com。



张照辉(1988-), 男, 高级工程师, 研究方向: 电力设备试验与状态诊断, E-mail: 363680788@qq.com。