

DOI: 10.19650/j.cnki.cjsi.J2107571

输电杆塔弱纹理部件的可迁移式检测

吴华¹, 梁方正¹, 刘草¹, 白晓静¹, 吕敏²

(1. 华北电力大学 控制与计算机工程学院 电力智能机器人工程技术研究中心 北京 102206;
2. 中国信息通信研究院 北京 100083)

摘要:输电杆塔关键弱纹理部件的通用检测方法依赖大量样本的标注和学习。在无相关部件样本训练情况下,本文提出一种可迁移的输电杆塔弱纹理部件检测方法。本文方法结合了孪生神经网络和互相关卷积用于融合样本块与待搜索区域的特征,其中通过样本部件掩膜裁剪以有效滤除背景噪声,最后在尺度、位置、交并比3个方面提出了对应的分数修正策略以提高检测精度。实验结果表明,本文提出的掩膜裁剪和分数修正策略有效提高了弱纹理部件的检出精度,按照 AP₅₀ 的标准,本文检测方法的平均准确率达到了 98.0%,能有效避免由于观测视角、尺度以及光照带来的干扰。

关键词:孪生神经网络;掩膜裁剪;弱纹理部件检测;输电杆塔

中图分类号: TH74 TP183 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Transferable detection for low texture components of transmission tower

Wu Hua¹, Liang Fangzheng¹, Liu Cao¹, Bai Xiaojing¹, Lyu Min²

(1. AMaze industrial robot research center, School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 2. China Academy of Information and Communications Technology, Beijing 100083, China)

Abstract: For key low texture components of the transmission tower, most detection methods rely on the labeling and training of a large number of sample images, while these samples are usually scarce. Proposes a transferable detection method-PowerNet for the key low textured components. It combines Siamese network and cross-correlation convolution to fuse the features of sample block and search region, and effectively filters out the background of the component sample image using mask cropping. Finally, scoring candidate region is proposed where the scale, location, and IoU are adopted to improve the accuracy of detection. Experiments show that mask cropping and scoring candidate region play an important role in accurate detection. According to AP50 criterion, the average accuracy rate of our detection method reaches 98%, which can detect various components of the transmission tower concerning the variance of view angle, scale, or illumination.

Keywords: Siamese network; mask cutting; low texture components detection; transmission tower

0 引言

电力发输变配用环节的安全就是国家安全。精准高效地识别巡检影像关键缺陷是实现电力输送安全保障的必要方法。近几年,随着我国无人机技术的规模化行业应用,输电线路的巡检数据采集速度和精度快速上升,海量巡检影像的分析任务超出了作业班组人员能力和精力的边界。为此,学者们提出了不少面向依赖前期训练大量样本的识别算法。Apostolopoulos 等^[1]修改了 virtual

geometry group (VGG) 网络,提出了 multilevel VGG19 (MVGG19)用于全自动的工业检测,该网络允许提取更多的局部和全局特征,而多余的特征则通过级联融合。Pan 等^[2]在工业检测锯材图像中的椭圆形缺陷的应用场景下,在区域候选网络(region proposal network, RPN)的基础上添加了感兴趣区域,合并成了 GPN (Gaussian proposal network),从而提高了检测的精度。Nguyen 等^[3]利用4个中等大小的数据集训练多阶段的检测和分类模型,并基于单发多盒检测器和深度残差网络来检测电力系统元件的常见故障。陈志强等^[4]总结了诸多在大多

据的驱动下的神经网络模型,在设备故障预测和健康管理中的应用。王永利等^[5]设计的卷积神经网络模型,经过训练可以有效应对印刷电路板的缺陷检测问题。鲍光海等^[6]利用残差网络结构和可分离卷积,有效改进了卷积神经网络,并在汽车高度调节器的缺陷检测中成功应用。

针对杆塔上关键连接弱纹理部件的准确检测是实现大规模杆塔结构语义理解的重要基础。然而依赖大量标注训练过程得到的模型的通用性不足,工程实际应用难度较大。采用可迁移式检测的方法能够较好地解决无训练少标注的工业检测需求。然而由于检测对象是杆塔上的弱纹刚性目标,通常采用的模板/特征匹配的方法不能够很好地匹配目标。然后本文后续实验表明,由于模板与待检测影像内目标的特征相关性不明显,无法正确地检测出刚性目标。

针对同一物体的特征匹配,已有很多知名算法。常见的图像匹配方法以角点检测与匹配为主,从 Harris 角点检测算子^[7]到加速分割测试特征算子^[8](features from accelerated segment test, FAST),以及对这类角点检测算子的改进方法。Lowe^[9]提出了尺度不变特征变换(scale-invariant feature transform, SIFT),将研究者的思维从角点检测中解放出来,它在空间尺度中寻找极值点,并提取出其位置、尺度、旋转不变量。Bay 等^[10]在吸取了 SIFT 的灵感后,提出了一种加速稳健特征(speeded up robust features, SURF),采用了哈尔特征以及积分图像的概念,这极大加快了程序的运行时间。但是这些经典的算法大多以采用底层的低级局部特征为基础,没有上升到类特征。Revaud 等^[11]提出了一种稠密匹配算法,该算法提取底层特征,在自下而上建立起大范围的匹配来模拟类间匹配,再通过层层对应来找到底层的碎片级对应方法。

针对图像检测的算法,随着计算机算力的不断提高和算法的改进,卷积神经网络^[12]逐渐取代传统机器学习算法,成为了目标检测的主流方法。经过 Fast R-CNN 等^[13]算法的积淀,Ren 等^[14]提出了 Faster R-CNN 算法,该算法将特征抽取、候选区域提取、边界框回归、分类任务都整合在了一个完整的网络里,使综合性能有了大幅的提高,在检测速度方面有了显著地进步。YOLO (you only look once) 系列^[15]通过修改特征提取主干、采用新模块(如空间金字塔池化)等方式,不断提高目标检测模型的精度和速度。但以上的神经网络只能检测已训练的类别,若做迁移学习则需要对应线路的标注好的数据集,在更换线路后,准确率达不到工业的需求,因此不适合应用到本文的问题中。Fan 等^[16]提出 FSOD (few-shot object detection) 来解决小样本的目标检测问题,文中以孪生神经网络基础,提出了一种新的注意力机制网络,它在 RPN

模块和检测器上学习支持集与的查询集之间的一般匹配关系。Wang 等^[17]通过设计一个两阶段的 fine-tune 策略,准确率远超之前的元学习方法。

1 问题描述

在实际无人机电力巡检过程中,同一条巡检线路上有诸多同类杆塔,且无人机拍摄的相对位置和角度近似相同。因此,利用同类杆塔相同位置这个重要的先验信息,以检测后续杆塔的目标部件,是实现杆塔弱纹理部件缺陷分析的有效途径。

然而,由于 GPS 定位精度、定位数据链路中的干扰、局部环境气象条件变化以及机载磁罗盘受电磁场的干扰等原因,相同位置的巡检图像会出现视角、尺度、背景等变化。正是这种情况使弱纹理金具的检测难度大大增加。针对该挑战性问题,本文提出面向输电线路弱纹理部件检测的神经网络算法——PowerNet,旨在解决电路巡检同类杆塔相同位置部件的检测问题,并对比研究了如 SIFT、SURF、CANet^[18]、DeepMatching 等相关匹配算法。

测试图像为实际无人机真实巡检时在同一线路不同杆塔相同位置拍摄的影像结果,测试结果如图 1 所示。

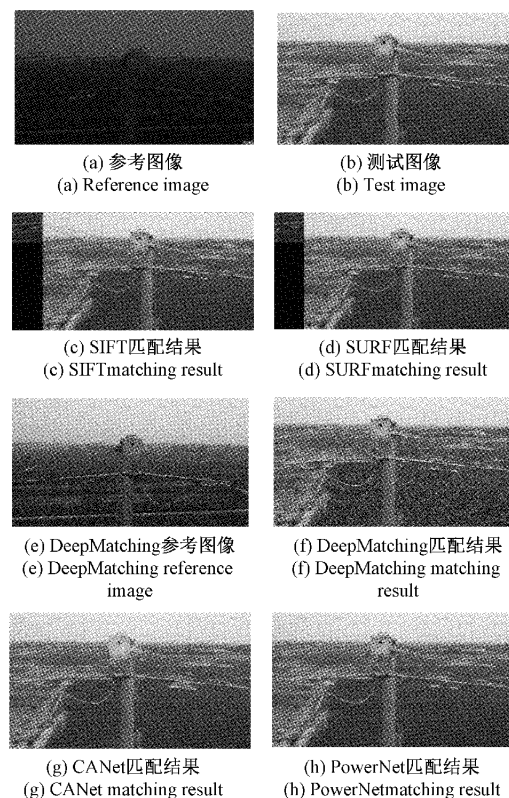


图 1 多种算法匹配结果

Fig. 1 Evaluation results of several algorithms

图1(a)、(b)是参考图像与测试图像,目的是能检测出图1(a)中的左侧防震锤。图1(c)、(d) SIFT和SURF算法均使用参考图像的前景部分与测试图像进行匹配, SIFT和SURF均未正确匹配到前景部分,由实验结果可见匹配效果不佳。图1(e)、(f) DeepMatching算法整体表现不佳,对于视角和背景变化较大的前景物体无法匹配。图1(g) CANet算法错误地匹配到了金属板和右侧防震锤,容易匹配到颜色相近的错误目标。图1(h) PowerNet利用了合理的分数修正策略,精密的掩膜裁剪以及孪生神经网络,在测试中取得了较好的实验效果。

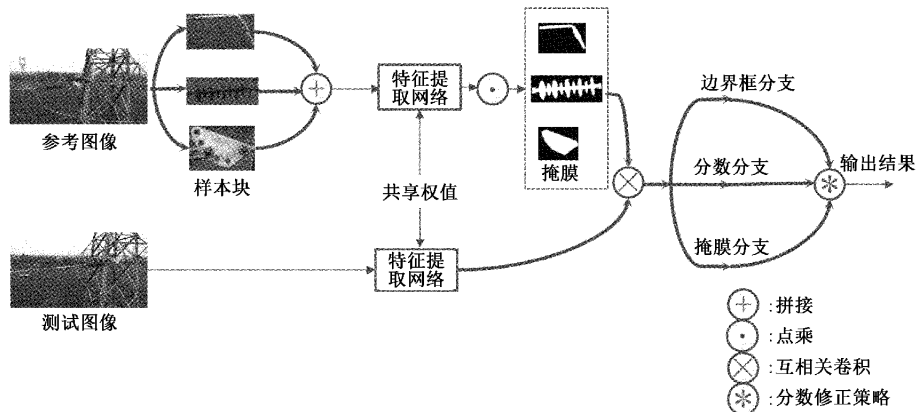


图2 PowerNet神经网络结构

Fig. 2 PowerNet structure

2.1 网络输入与特征提取

电力巡检图像里可能有多处需要检查的弱纹理部件,包括绝缘子、金具、防震锤等。为此,提出的PowerNet网络可检测多种部件。对于一张巡检图像允许选定多个样本块,经过图像预处理后再通过拼接融合为统一的四维张量。在检测相同位置的部件时,以样本块在样本图像中的位置作为先验信息,从而缩小待搜索区域的范围,提高算法效率。

PowerNet选用了ResNet-50作为特征提取的主干网络。ResNet使用了残差学习的方法解决了神经网络加深而退化的问题。除此之外,PowerNet采用了共享权值的孪生神经网络,孪生神经网络能有效度量相似度。样本块和待搜索区域通过同一ResNet网络来提取特征,并以互相关卷积作为特征融合的方法。

2.2 掩膜剪裁与互相关特征融合

利用掩膜剪裁特征能够有效滤除背景噪声,从而提高检测的精度。掩膜是由0和1组成的一个二进制图像,在图像分割数据集标注时,二进制掩膜将帮助算法区分前景和背景。掩膜块与样本块一一对应,由于样本块在特征提取阶段经过了降采样,所以需要利用双线性插

2 PowerNet神经网络检测算法

本文提出的神经网络PowerNet结构如图2所示。PowerNet采用了多类目标的网络输入,允许在一张巡检图像内检测多种弱纹理部件,使用ResNet-50^[19]提取特征,并利用二进制的掩膜来滤除样本图像中的背景部分。除此之外,PowerNet使用孪生神经网络^[20]和互相关卷积来融合特征^[21],汇总各分支结果,然后通过有效的分数修正策略选取最佳的区域,根据最佳区域输出掩膜的最小包围框来确定边界框的位置。

值将掩膜块降采样为与特征长宽相同尺寸,再利用张量计算的传播性与三维特征点乘,从而将背景区域的特征置0。

本文做了使用掩膜的消融实验如图3所示。实验A:以图3(a)中的绝缘子上方的金属板作为样本,通过对比图3(c)、(e)可以发现,图3(c)中错误匹配成与样本金属板尺度相近的绝缘子,这是由于金属板样本在提取样本块时经过了尺度放大,将绝缘子顶部也纳入样本块中,再加上待搜索图中的绝缘子顶部的长宽和尺度合适所导致的;为了解决在尺度放大时收入其他前景目标而导致匹配错误,从而使用了掩膜对特征剪裁,能够有效避免这种情况,如图3(e)所示。实验B:以图3(b)中金属板为样本,对比图3(d)、(f)可以发现,图3(d)中由于在尺度放大时收入了其他前景目标导致分割了多余的部分,从而导致最小包围框大于真实框,在使用掩膜剪裁后,能够有效增大交并比(intersection over union, IoU),如图3(f)所示。

本文孪生神经网络中度量相似度的方法采用的是互相关融合特征。具体如式(1)所示: F_{corr} 代表融合特征, Σ 在式(1)里代表拼接, c 代表通道数, F_{search}^k 和 $F_{template}^k$ 代表第 k 个通道的测试图像特征和样本图像特征, $mask$ 代

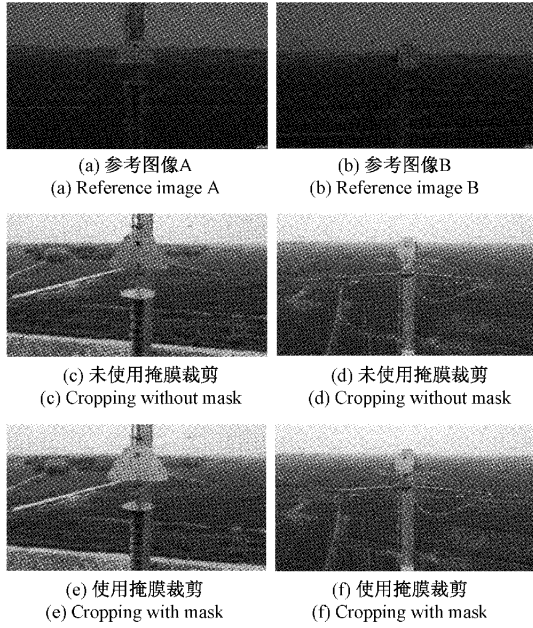


图3 掩膜使用消融实验

Fig. 3 Ablation experiments for mask cropping

表样本图像掩膜, Σ 代表特征拼接, $B \otimes A$ 代表 B 以 A 为卷积核进行卷积, \cdot 代表矩阵点乘。

$$F_{corr} = \sum_{k=1}^c (F_{search}^k \otimes (F_{template}^k \cdot mask)) \quad (1)$$

输出的融合特征 F_{corr} 中, 每一个通道方向的向量都代表待搜索图像的一个候选区域的融合特征。

2.3 分数修正策略

每一个候选区域都由边界框分支产生多个尺度不同的边界框以搜索前景目标, 由分数分支输出每个边界框的前景分数。则分数最高的区域作为最佳区域, 将对应 mask 分支输出的最小包围框作为输出结果。

由于相近视角任务的特殊性, 本算法设计了3种分数修正策略提高检测的精度, 分别是: 尺度修正、位置修正以及交并比修正。首先, 尺度修正利用了每个边界框的长宽比和面积与样本框的区别, 如式(2)所示, S_0^t 代表分数分支产生的第 t 个边界框的初始前景分数, r^t 代表第 t 个边界框的长宽比与样本框的长宽比的比值(取大于1的值), c^t 代表第 t 个边界框的面积与样本框面积的比值(取大于1的值), α 代表尺度修正的修正系数。因此, 当 r^t 和 c^t 越大, 更新后的分数 S_1^t 的值越小。其次, 位置修正由每个边界框位置与样本框初始位置的差异决定的。如式(3)所示, $f_{Hamming}(x)$ 代表 Hamming 函数, x 距离中心越远, $f_{Hamming}(x)$ 的函数值越低。 l^t 代表第 t 个边界框与样本框初始区域的相对位置, β 代表位置修正的修正系数。因此, 边界框与样本框相差距离越远, 更新后的分数 S_2^t 越低。

$$S_1^t = S_0^t \times \exp((1 - r^t \times c^t) \times \alpha) \quad (2)$$

$$S_2^t = S_1^t \times (1 - \beta) + f_{Hamming}(l^t) \times \beta \quad (3)$$

$$S_{final}^t = S_2^t \times (1 - \gamma) + IoU^t \times \gamma \quad (4)$$

最后对各候选区域的分数进行交并比修正, 如式(4)所示; S_{final}^t 代表第 t 个候选区域的最终分数, IoU^t 代表第 t 个候选区域产生的掩膜与样本掩膜的交并比, γ 代表交并比修正的修正系数。当 IoU^t 越大时, 经修正后的最终分数 S_{final}^t 也越大。

本文针对分数修正策略做了消融实验如图4所示。实验A: 参考目标为图4(a)中右侧靠左的防震锤, 对比图4(c)、(e)的实验结果, 未使用分数修正策略的结果图4(c)错误的匹配到了右侧靠右的防震锤, 经过分数修正策略后, 成功匹配到了正确的目标, 如图4(e)所示。实验B: 参考目标为图4(b)中间的线夹, 在未使用分数修正策略的结果图4(d)中, 错误地匹配到了线夹的左侧部分, 利用分数修正策略, 正确匹配到了目标线夹, 如图4(f)所示。

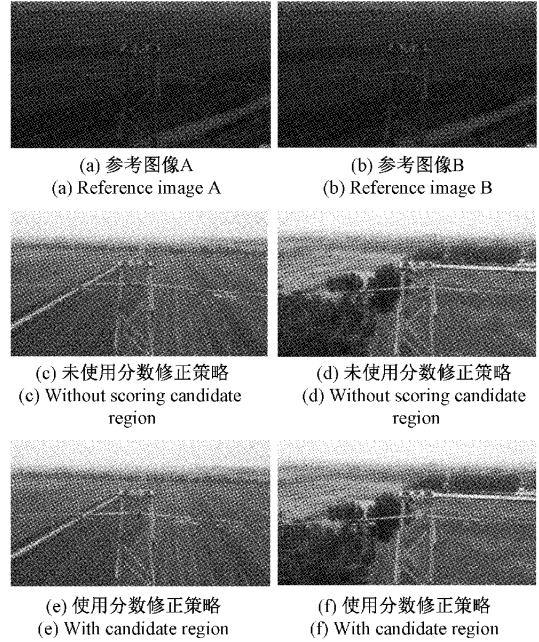


图4 分数修正策略消融实验

Fig. 4 Ablation experiments for scoring candidate region

3 实验结果

PowerNet 神经网络算法针对同类杆塔相同位置的弱纹理部件的检测。实验中在测试集里找到与参考图像标注的对应部件, 并实验中做了掩膜剪裁和分数修正策略的消融实验, 以验证其优化效果。测试集来源于多条不同线路的同类杆塔上多种目标, 具体分类为绝缘子、三角连板、线夹、防震锤、金属板、组件。实验

的挑战性在于背景、尺度、位姿等都有较大的区别,识别起来有较大的难度,具有相当的实验价值。评价指标为掩膜的最小包围框与样本框的交并比,评价标准包括 AP_{50} 和 AP_{75} 。

PowerNet 中 ResNet-50 选用 ImageNet-1k 的预训练网络,除此之外使用了 COCO^[22] 训练整个模型,未使用电力巡检相关数据集训练。实验使用 NVIDIA Tesla V100 的 GPU 进行推理,在一幅图单目标且利用位置先验信息的情况下,模型平均推理速度为 0.055 s。

在图 5、图 6 中,折线图的横坐标代表不同类型的检测部件,其中总计代表整体的统计结果,纵坐标分别为 AP_{50} 和 AP_{75} 的统计结果,折线颜色由浅至深分别代表未优化、仅使用掩膜剪裁、仅使用分数修正、以及两种优化策略都使用的检测结果。由折线图可以直观地分析出,掩膜剪裁和分数修正可以有效提高检测的准确性,在使用两种优化策略的情况下,以 AP_{50} 标准提高了 7.9%,以 AP_{75} 的标准提高了 13.9%,在纹理不明显的三角连板的检测方面尤为明显,具体的实验数据统计结果详如表 1、2 所示。由图 7 实验结果所示,PowerNet 神经网络可以有效应对电力巡检中同类杆塔同位置的弱纹理部件的检测问题,且可以自由定义部件的组合方式,从而整体检测,对于细小的紧固件也有良好的检测效果。

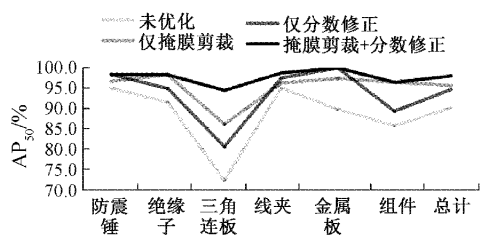


图 5 AP_{50} 标准实验数据折线图

Fig. 5 Line chart of AP_{50} experimental results

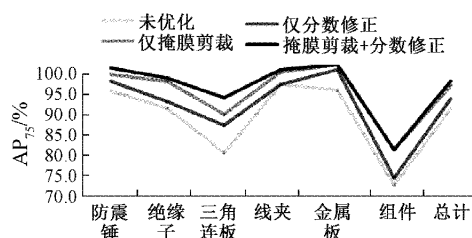


图 6 AP_{75} 标准实验数据折线图

Fig. 6 Line chart of AP_{75} experimental results



(a) 参考图像 (b) 测试图像
(a) Reference image (b) Test image

图 7 典型检测结果

Fig. 7 Typical detection results

表 1 AP_{50} 标准实验数据
Table 1 AP_{50} experimental results

AP_{50}	防震锤	绝缘子	三角连板	线夹	金属板	组件	总计
未优化	95.0	91.5	72.2	95.0	89.7	85.7	90.1
仅掩膜剪裁	96.7	98.3	86.1	96.3	97.4	96.4	95.7
仅分数修正	98.3	94.9	80.6	97.5	100	89.3	94.7
掩膜剪裁+分数修正	98.3	98.3	94.4	98.8	100	96.4	98.0

表2 AP₇₅ 标准实验数据
Table 2 AP₇₅ experimental results

AP ₇₅	防震锤	绝缘子	三角连板	线夹	金属板	组件	总计
未优化	86.7	78.0	55.6	90.0	87.2	39.3	77.8
仅掩膜剪裁	95.0	91.5	75.0	96.3	100	57.1	89.4
仅分数修正	91.7	81.4	69.4	90.0	97.4	42.9	82.8
掩膜剪裁+分数修正	98.3	93.2	83.3	97.5	100	57.1	91.7

4 结 论

本文针对输电杆塔弱纹理部件的检测问题设计了一套以深度神经网络为基础的方法——PowerNet。在网络里创新了多种关键环节,并在实验中证明了其有效性。

首先,多目标的输入设计能够有效提高检测的效率,更加符合实际应用中的需要。其次,像素级的掩膜剪裁能够有效的避免背景带来的干扰,能够有效减少误检率以及检测多余区域的情况。最后,有效融合了多种成熟的网络结构来提高神经网络的能力,本文利用的3种分数修正策略:尺度修正、位置修正以及交并比修正,能够充分利用参考目标的信息,从而有效提高检测的准确率。

本文使用现场巡检图像进行实验验证了本方法能够有效避免由于观测视角、尺度以及位姿带来的干扰,具有良好的工程应用价值。

参考文献

- [1] APOSTOLOPOULOS I D, TZANI M. Industrial object, machine part and defect recognition towards fully automated industrial monitoring employing deep learning. The case of multilevel VGG19 [J]. Arxiv Preprint, 2020, Arxiv:2011.11305.
- [2] PAN S, FAN S, WONG S W K, et al. Ellipse detection and localization with applications to knots in sawn lumber images [C]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020: 3892-3901.
- [3] NGUYEN V N, JENSSEN R, ROVERSO D. Intelligent monitoring and inspection of power line components powered by UAVs and deep learning [J]. IEEE Power and Energy Technology Systems Journal, 2019:11-21.
- [4] 陈志强,陈旭东, JOS E, 等. 深度学习在设备故障预测与健康中的应用 [J]. 仪器仪表学报, 2019, 40(9):206-226.
- [5] 王永利,曹江涛,姬晓飞. 基于卷积神经网络的PCB缺陷检测与识别算法 [J]. 电子测量与仪器学报, 2019, 33(8):78-84.
- [6] 鲍光海,林善银,徐林森. 基于改进型卷积网络的汽车高度调节器缺陷检测方法 [J]. 仪器仪表学报, 2020, 41(2):157-165.
- [7] HARRIS C G, STEPHENS M. A combined corner and edge detector [C]. Alvey Vision Conference, 1988, 15(50):10-5244.
- [8] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection [C]. European Conference on Computer Vision, 2006:430-443.
- [9] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [10] BAY H, TUYTELAARS T, VAN GOOL L. Surf: Speeded up robust features [C]. European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2006:404-417.
- [11] REVAUD J, WEINZAEPFEL P, HARCHAOUI Z, et al. Deepmatching: Hierarchical deformable dense matching [J]. International Journal of Computer Vision, 2016, 120(3):300-323.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural

- networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [13] GIRSHICK R. FAST R-CNN [C]. Proceedings of the IEEE International Conference on computer Vision, 2015: 1440-1448.
- [14] REN S, HE K, 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, 2016, 39(6):1137-1149.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. Arxiv Preprint, 2020, Arxiv: 2004.10934.
- [16] FAN Q, ZHUO W, TANG C K, et al. Few-shot object detection with attention-RPN and multi-relation detector[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 4013-4022.
- [17] WANG X, HUANG T E, DARRELL T, et al. Frustratingly simple few-shot object detection[J]. Arxiv Preprint, 2020, Arxiv:2003.06957.
- [18] ZHANG C, LIN G, LIU F, et al. Canet: Class-agnostic segmentation networks with iterative refinement and attentive few-shot learning[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 5217-5226.
- [19] 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, 2016: 770-778.
- [20] KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition [C]. ICML Deep Learning Workshop, 2015, 2.
- [21] WANG Q, ZHANG L, BERTINETTO L, et al. Fast online object tracking and segmentation: A unifying approach[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 1328-1338.
- [22] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C]. European

Conference on Computer Vision, 2014: 740-755.

作者简介



吴华,2010年于北京航空航天大学获得博士学位,现为华北电力大学副教授,主要研究方向为高可信智能识别、自主巡检、视点规划及SLAM等。

E-mail: wuhua@ncepu.edu.cn

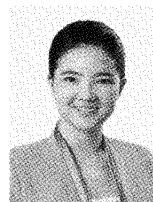
Wu Hua received his Ph. D. degree from Beihang University in 2010. He is currently an associate professor at North China Electric Power University. His main research interests include high-reliability intelligent recognition, edge reasoning, autonomous inspection, viewpoint planning, SLAM, etc.



白晓静(通信作者),2017年于华北电力大学获得博士学位,现为华北电力大学讲师。主要研究方向为:电力设备缺陷智能检测、图像处理、图像识别等。

E-Mail: baixiaojing@ncepu.edu.cn

Bai Xiaojing (Corresponding author) received her Ph. D. degree from North China Electric Power University in 2017. She is currently a lecturer at North China Electric Power University. Her main research interests include intelligent detection of defects in electric power equipment, image processing, image recognition, etc.



吕敏,2010年于北京理工大学获得学士学位,2013年于华北电力大学获得硕士学位,现为中国信息通信研究院工程师,主要研究方向为5G、人工智能算法、ICT行业研究。

E-mail: lvmin@caict.ac.cn

Lyu Min received her B. Sc. degree from Beijing Institute of Technology in 2010, and received her M. Sc. degree from North China Electric Power University in 2013. She is currently an engineer at China ICT Research Institute. Her main research interests include 5G, artificial intelligence, and ICT industry research.