

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2023)11-3485-12

论文引用格式: Liang H G, Zhao H X, Liu L H, Yue P and Zheng Z Y. 2023. Combining cross-layer feature fusion with cascade detectors for anti-vibration hammer defects detection. Journal of Image and Graphics, 28(11):3485-3496(梁华刚, 赵慧霞, 刘丽华, 岳鹏, 郑振宇. 2023. 结合跨层特征融合与级联检测器的防震锤缺陷检测. 中国图象图形学报, 28(11):3485-3496)[DOI:10.11834/jig.220789]

结合跨层特征融合与级联检测器的防震锤缺陷检测

梁华刚, 赵慧霞*, 刘丽华, 岳鹏, 郑振宇

长安大学电子与控制工程学院, 西安 710064

摘要: 目的 防震锤可以减少输电线路的周期性震动以降低线路的疲劳损害, 定期对防震锤进行巡检非常必要。针对目前无人机巡检输电线路所得航拍图像背景复杂, 而种类较多、形状各异以及特性不一的防震锤在航拍图像中占据像素面积很小, 导致防震锤检测过程中出现的检测精度低、无法确定缺陷类型等问题, 提出了一种结合跨层特征融合和级联检测器的防震锤缺陷检测方法。方法 本文使用无人机对防震锤部件巡检的航拍图像进行数据扩充建立防震锤缺陷检测数据集, 并划分了4种缺陷类型, 为研究提供了数据基础。首先, 以VGG16(Visual Geometry Group 16-layer network)为基础对1、3、5层特征进行特征融合得到特征图, 平衡了语义信息和空间特征; 其次, 使用3个级联检测器对目标进行分类定位, 减小了交并比(intersection over union, IoU)阈值对网络性能的影响; 最后, 将非极大值抑制法替换为Soft-NMS(soft non-maximum suppression)算法, 去除边界框保留了最佳结果。结果 在自建数据集上验证网络模型对4种防震锤缺陷类型的检测效果, 与现有基于深度学习的其他6种先进算法相比, 本文算法的平均准确率比Fast R-CNN(fast region-based convolutional network)、Faster R-CNN、YOLOv4 (you only look once version 4)分别提高了13.5%、3.4%、5.8%, 比SSD300(single shot MultiBox detector 300)、YOLOv3、RetinaNet分别提高了9.5%、8.5%、8%。与Faster R-CNN相比, 本文方法的误检率降低了5.61%, 漏检率降低了3.01%。结论 本文提出的防震锤缺陷检测方法对不同背景、不同光照、不同角度、不同尺度、不同种类和不同缺陷种类的防震锤均有较好的检测结果, 不但可以更好地提取防震锤的特征, 而且还能提高分类和位置预测网络的定位精度, 从而有效提高了防震锤缺陷检测算法的准确率, 在满足防震锤巡检工作实际检测要求的同时还具有较好的鲁棒性和有效性。

关键词: 防震锤缺陷; 深度学习; 小尺度目标检测; 跨层特征融合; 级联检测器

Combining cross-layer feature fusion with cascade detectors for anti-vibration hammer defects detection

Liang Huagang, Zhao Huixia*, Liu Lihua, Yue Peng, Zheng Zhenyu

School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China

Abstract: Objective Anti-vibration hammers can reduce the periodic vibration of transmission lines to minimize line fatigue damage. Therefore, a regular inspection of these hammers is necessary. Previous studies on anti-vibration hammers for transmission lines have mostly focused on the identification or fault classification of anti-vibration hammers with only one case of deficiency. However, the detection methods for such small-scale targets are limited, and their leakage and false

收稿日期: 2022-08-22; 修回日期: 2023-01-20; 预印本日期: 2023-01-27

* 通信作者: 赵慧霞 18809299832@163.com

基金项目: 国家自然科学基金项目(61973243); 陕西省重点研发计划项目(2020ZDLGY09-03)

Supported by: National Natural Science Foundation of China (61973243); Shaanxi Provincial Key R&D Program, China (2020ZDLGY09-03)

detection rates remain very high, which cannot guarantee the safe operation of transmission lines. Therefore, in view of the complex background of the aerial images obtained from the current unmanned aerial vehicle (UAV) inspection of transmission lines, the different types, shapes, and characteristics of anti-vibration hammers occupying a small pixel area in these aerial images can lead to problems, such as inability to determine the types of defects and the low detection accuracy in the anti-vibration hammer detection process. In this paper, we propose an anti-shock hammer defect detection method based on cross-layer feature fusion and cascade detector. **Method** Given the lack of a public anti-vibration hammer dataset at home and abroad for targeted research, this paper takes the aerial images taken via the UAV inspection of anti-vibration hammer components as the original data and then expands the dataset by geometric and contrast transformation to ensure the equalization of the number of different types of anti-vibration hammers in the dataset and to establish an anti-vibration hammer defect detection dataset. The anti-vibration hammer defects are then refined and classified into four categories, namely, normal, corroded, broken, and collision, to serve as the data basis for the algorithm research designed in this paper. The anti-shock hammer defect detection algorithm is then designed. The proposed anti-shock hammer defect detection method is mainly divided into two major parts, namely, the feature extraction network and classification location prediction network. The feature extraction network extracts accurate anti-seismic hammer features by fusing the Visual Geometry Group 16-layer network (VGG16). The main idea is to insert a convolutional kernel of size 1×1 into the last convolutional layer of the first, third, and fifth layers, after which the last convolutional layer of the first layer is connected to the maximum pooling layer and fused with the third layer using a deconvolution operation after the fifth layer. These two layers are then fused to form the final feature map, which balances the semantic information and spatial features. To reduce the impact of the intersection over union (IoU) threshold on network performance, in the localization and classification network part, the classification and location prediction network proposed in this paper uses three cascade detectors to gradually increase the IoU threshold and improve the quality of samples and the training effect of the network. The non-maximal suppression method is then replaced by the soft non-maximum suppression (Soft-NMS) algorithm to remove the bounding box. **Result** The main contributions of this paper are as follows: 1) the dataset part expands the data using the aerial images taken via the UAV inspection of anti-vibration hammer components to establish the anti-vibration hammer defect detection dataset. To ensure the validity of this dataset, part of those samples that cannot be easily separated even through the naked eye are reasonably removed; 2) the network model is based on VGG16, feature fusion is performed on 1, 3, and 5 layers of features to effectively obtain more features, and three cascade detectors are used to classify the target to reduce the impact of the IoU threshold on network performance and improve the detection capability of the algorithm. The average accuracy of the network model was improved by 13.5%, 3.4%, and 5.8% over the fast region-based convolutional network (Fast R-CNN), Faster R-CNN, and you only look once version 4 (YOLOv4), respectively, and by 9.5%, 8.5%, and 8% over single shot MultiBox detector 300 (SSD300), YOLOv3, and RetinaNet, respectively, when compared with six other advanced algorithms based on deep learning. These results may be explained by the incorporation of multiple layers of features, which enable the model to obtain more information about low- and high-level image features. Moreover, the uses of cascade detectors reduce the impact of the IoU thresholds on network performance and eventually improve the detection accuracy of small-scale targets. Compared with Faster R-CNN, the false detection rate of the proposed method is reduced by 5.61%, whereas the missed detection rate is reduced by 3.01%. Therefore, the proposed method improves its accuracy while effectively reducing its false detection and missed detection rates and works particularly well in practical applications. **Conclusion** The proposed anti-shock hammer defect detection method obtains good detection results for different backgrounds, illuminations, angles, scales, and types of anti-shock hammers. Experimental results show that this method not only efficiently extracts the characteristics of anti-shock hammers but also improves the localization accuracy of the network, thus effectively improving the accuracy of the algorithm and satisfying the actual detection requirements of anti-shock hammer inspection work while showing improved robustness and effectiveness.

Key words: shockproof hammer defects; deep learning; small-scale object detection; cross-layer feature fusion; cascade detector

0 引言

输电线路大部分位于高海拔地区,环境恶劣(赵振兵等,2021)。采用传统的人工巡检方式不仅无法保证巡检人员的安全,而且巡检人员工作质量、水平不一,很难保证巡检质量(吕明等,2012)。无人机灵活性较强,能极大提高巡检效率以及巡检质量(陈昕哲,2021),同时降低巡检风险,广泛应用于电力巡检领域中(陈昕哲,2001)。防震锤作为输电线路的典型部件,对于减小输电线路受风产生的震动,从而延长输电线路的工作寿命有着非常重要的作用(Liang等,2020)。然而受防震锤本身结构和工作环境的复杂性以及地震载荷等因素影响,使其很容易出现缺陷,一旦防震锤产生缺陷,线路也将发生问题。准确定位防震锤,及时发现缺陷并对其进行维护,可大大延长输电线路的使用时长(Simonvan和Zisserman,2015)。但防震锤在无人机航拍图像中所占比例较小,且其种类多样(许保瑜等,2021)。防震锤位置与导线平行,特征不明显,与塔杆距离较近,会有部分特征相似导致误检。航拍图像背景复杂、角度多样,均给防震锤的缺陷检测工作带来了许多挑战。

防震锤的缺陷检测方法主要分为传统的图像处理方法和深度学习方法。基于传统的图像处理方法大多利用部件的形状、纹理等特征将所需要的部件正确分割出来,再进行部件识别以及故障检测,如针对基于链码跟踪与角点检测的识别方法(Pan和Xiao,2009)、利用计算机视觉导航技术识别出防震锤(陈晓娟等,2010)、将类 Haar 特征与级联 Ada-Boost 算法结合识别单一防震锤(金立军等,2012)、根据分割出的防震锤的 RGB 颜色直方图判断防震锤是否锈蚀(宋伟等,2016)。然而传统的图像处理方法中,大部分方法的准确率决定性因素是对图像正确分割,很难满足当前复杂背景下的输电线路的部件缺陷检测需求。随着科学技术的发展,深度学习方法开始在输电线路的部件缺陷检测中得以应用。通过 Faster R-CNN (faster region-based convolutional network)(Ren 等,2017)完成绝缘子的缺陷检测,该方法同时提高了绝缘子检测的精确度和实时性,但是不能很好地解决遮挡问题(Liu 等,2018)。针对移动端防震锤、绝缘子和悬垂线夹的检测算法

MSFF-KCD (multi-scale feature fusion in key component detection)(杨昱等,2020),提高了3种部件的检测精度与速度。将自我注意机制与ECA-Net (efficient channel attention for deep convolutional neural network)(Wang 等,2020)网络相结合,大大降低了原有YOLOv4(you only look once version 4)算法的复杂度,提高了检测速度。将迁移学习与YOLOv3结合并使用图像增广技术解决了防震锤和线夹样本少的问题(张永翔等,2020)。焦润童等人(2021)对防震锤进行分割并利用Faster R-CNN(Huang等,2015)对防震锤进行识别。充分利用上下文信息(贾立业等,2021),结合DeepLabV3+语义分割(Chen等,2018)网络,实现防震锤的识别和缺陷检测。Faster R-CNN引入双注意力机制来检测航拍输电线路螺栓缺陷,有效避免了螺栓缺陷检测中的误判漏判问题,为输电线路的其他缺陷检测任务奠定了良好的基础(戚银城等,2021)。将动态非极大值抑制(non-maximum suppression, NMS)方法应用到YOLOv4目标检测,能够检测出重叠的防震锤,检测精度有所提高(贾雁飞等,2022)。

深度学习解决了很多输电线路部件检测方面的难题。但是现有的针对输电线路防震锤的研究大都只针对防震锤的识别或防震锤的故障分类只有缺损一种情况,而不同的缺陷对输电线路安全的影响程度也不同,需要巡检部门有不同的应对策略,所以需要防震锤的缺陷细化。其次航拍图像背景复杂多变,防震锤在市面上的类型较多、形状各异、特性不一,且防震锤属于小尺度目标,在航拍图像中占据的像素面积很小,检测难度更高。现有方法针对输电线路中大尺度目标已经有了较高的准确率,但是针对防震锤这类小尺度目标的检测方法较少,而且漏检率和误检率仍然很高,无法保障输电线路的安全运行。

防震锤缺陷检测的关键技术在于防震锤特征的提取方法,有效的防震锤特征提取工作将大大提高防震锤检测的精度。因此,本文主要针对特征提取和分类位置预测进行改进,提出了一种结合跨层特征融合与级联检测器的防震锤缺陷检测方法。主要贡献有:1)利用无人机对防震锤部件巡检的航拍图像进行数据扩充,建立防震锤缺陷检测数据集,同时为保证数据集的有效性,合理去除一部分通过肉眼也难以分别的样本。2)网络模型以VGG16(Visual

Geometry Group 16-layer network)为基础,对第1、3、5层特征进行特征融合有效地获取更多的特征,并使用3个级联检测器对目标分类减小交并比(intersection over union, IoU)阈值对网络性能的影响,提高算法的检测能力。3)本文防震锤缺陷检测任务包括不同光照、不同角度、不同背景、不同尺度和不同防震锤种类等多个场景,并通过大量的实验论证了本文方法的有效性和准确性。

1 防震锤缺陷网络模型

1.1 网络结构

基于跨层特征融合和级联检测器的防震锤缺陷检测方法整体框架如图1所示。主要分为特征提取

网络和分类位置预测网络两大部分。算法主要内容包括:

1)通过融合VGG16提取更准确的防震锤特征。主要思想是分别在第1、3、5层的最后一个卷积层中插入一个卷积核,大小为 1×1 ,之后对第1层的最后一个卷积层接入最大池化层,在第5层后使用反卷积操作,分别与第3层融合。最终将这两层特征融合形成最终的特征图。

2)IoU 阈值用于定义网络的正样本和负样本,训练检测器的阈值决定了检测的质量。为了降低IoU 阈值对网络性能的影响,本文的分类和位置预测网络使用3个级联检测器,逐渐增加IoU 阈值,提高样本的质量和网络的训练效果,并使用Soft-NMS(Bodla等,2017)代替NMS去除边界框保留最佳结果。

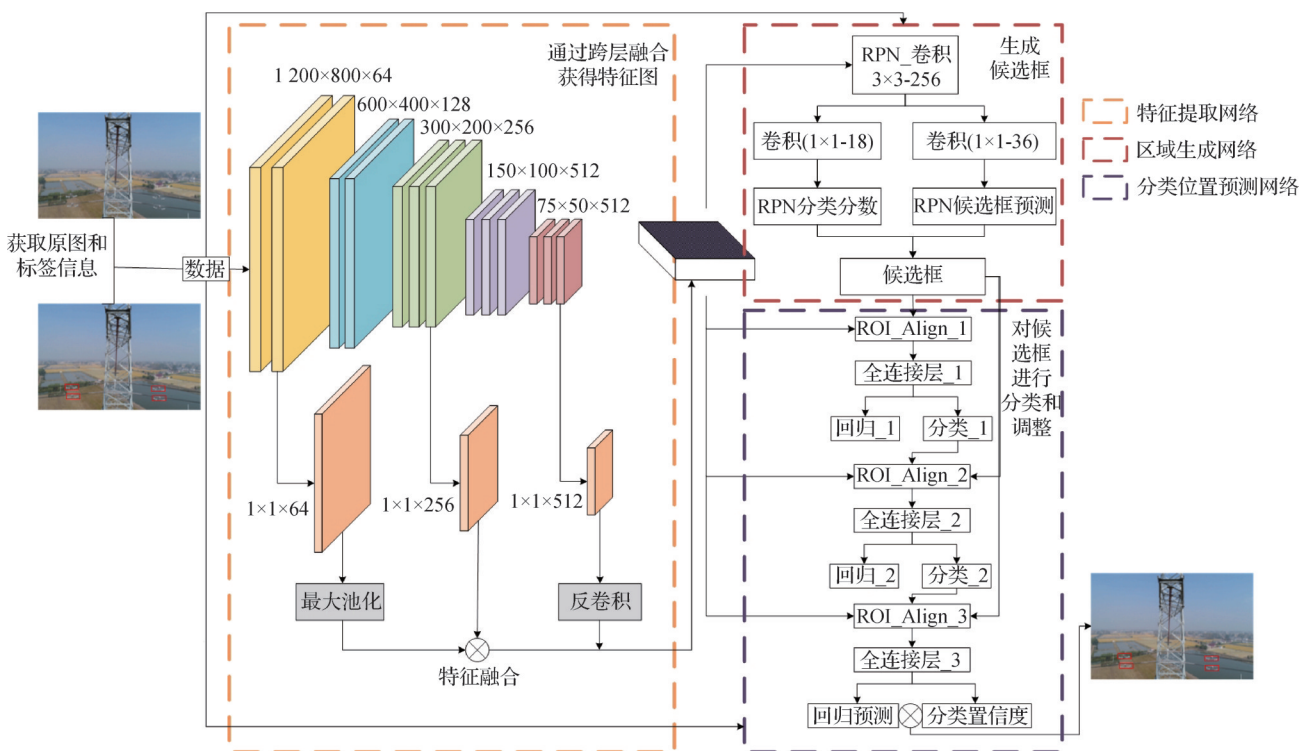


图1 防震锤缺陷检测算法整体框架

Fig. 1 Overall framework of vibration hammer defect detection algorithm

1.2 特征提取网络

作为目标检测中的主干网络,特征提取网络利用卷积神经网络(convolutional neural network, CNN)对图像的特征信息进行提取,并将获取到的特征图输入进分类位置预测网络,网络预测待检测目标的类别和位置,特征图的质量严重影响目标分类和定位的准确度。研究发现,深层的网络可以输出更复

杂的特征,从而具有更强的特征表达的能力,但随着网络层数的加深,信息越来越抽象,会使图像原有信息丢失,也会出现梯度消失的问题。传统的特征提取网络为了获得更大的感受野使用大的下采样倍数,但这只在图像分类时有利,而在对目标进行定位时则表现很差。

在特征提取网络中,通常采用的网络结构包括

VGG16 (Simonvan 和 Zisserman, 2015)、ResNet-50 (residual network-50) (He 等, 2016)、ResNet-101 (He 等, 2016)、Darknet-53 (Redmon 和 Farhadi, 2018)。由于小尺度目标在图像中覆盖面积较小,可提取到特征信息较少,在使用深度 CNN 提取小尺度目标特征时,小尺度目标的特征信息会部分丢失,从而缺乏了很多目标的外观信息,难以将目标从复杂的背景或相似物体中区分出来。航拍图像中的防震锤在图像中占比很小,小尺寸目标的定义为目标的长宽与原图的长宽比均小于 0.1,而航拍图像中防震锤与图像的长宽比范围经计算分别为 0.055~0.095 和 0.033~0.075,且航拍图像背景复杂,使用深度神经网络提取防震锤的特征图难度很大。所以本文的特征提取网络以 VGG16 为基础进行改进。随着特征提取网络层数加深,获得的特征具有更多的语义信息,但分辨率会降低。本文专注于小尺度目标的防震锤,对

VGG16 进行改进,改进后的结构如图 2 所示。

卷积神经网络的高层特征包含较多的语义信息有助于分类,但是如果将第 5-3 卷积层的输出作为最终特征图,很容易忽略目标之间的细节性差异。如果可以将第 3-3 卷积层的输出作为最终特征图,由于该层拥有更高的空间分辨率,所以会包括更加丰富的空间特性和纹理信息,此时检测精度较高。第 5-3 卷积层中具有语义信息会使得分类效果更好。

可以选择融合这两层特征,在具有较高空间分辨率的同时获取语义信息。如图 2 所示,使用 1×1 卷积、最大池化以及 ReLU 激活函数融合第 1-2 卷积层与第 3-3 卷积层特征,通过 1×1 卷积、反卷积操作以及 ReLU 激活函数将第 5-3 卷积层与第 3-3 卷积层特征合并。融合后的特征通过 concat 层连接在一起获取最终特征图。本方法通过融合这些特征可获得更准确的防震锤特征。

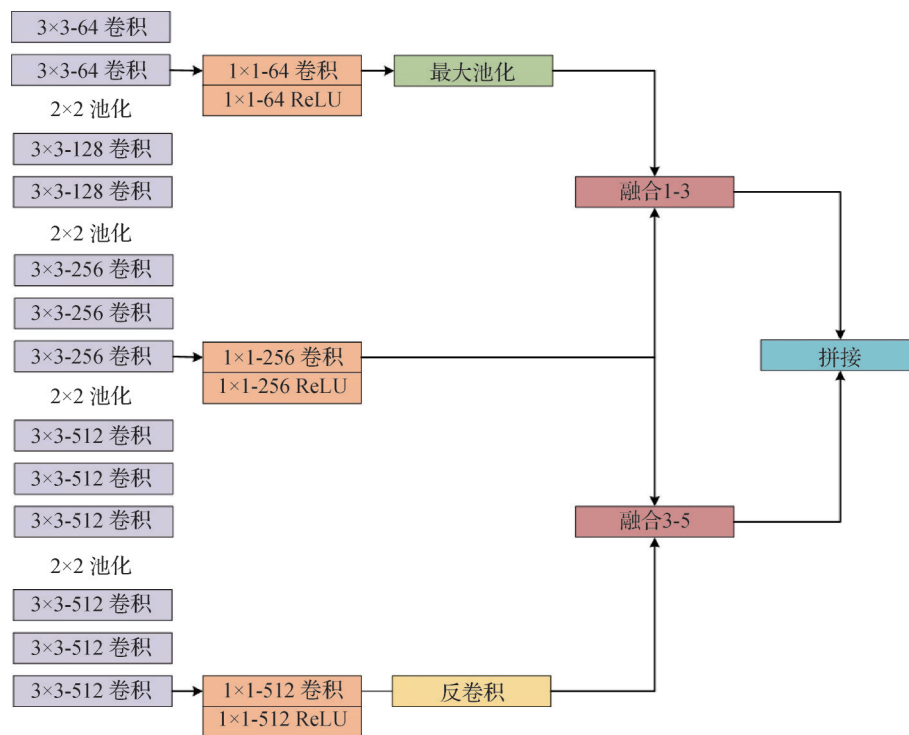


图2 跨层融合的VGG网络结构图

Fig. 2 Structure diagram of VGG network with cross-layer fusion

1.3 分类和位置预测网络

1.3.1 区域生成网络

Faster R-CNN 在进行候选框提取时的区域生成网络采用区域候选网络 (region proposal network, RPN)。RPN 使用非极大值抑制法 (NMS) 来选择回归损失后最大得分 S_i 的目标框,这可以获得更高的

置信度。但是 NMS 的阈值需要人为设定,但是如果阈值设置过大或过小都会导致模型精度降低。本文使用 Soft-NMS 方法代替 NMS 算法。在该方法中,相邻框检测方法根据重叠部分的重叠程度设定了衰减函数,重叠越大得分 S_i 越小,但当 $S_i=0$ 时,会导致相邻检测目标被漏检,因此采用式 (1) 保证 S_i 取值不为

0。Soft-NMS计算为

$$S_i = \begin{cases} S_i & IoU(M, b_i) > N_i \\ S_i(1 - IoU(M, b_i)) & IoU(M, b_i) \leq N_i \end{cases} \quad (1)$$

式中, S_i 表示预测框的分数, M 表示得分最高的框, b_i 表示当前框, N_i 表示NMS阈值。 $IoU(M, b_i)$ 表示得分最高的预测框与当前框的交并比。

1.3.2 分类和位置预测

在Faster R-CNN中, IoU阈值通常用于定义正样本和负样本。训练网络的阈值定义了检测的质量。阈值过低将导致检测质量低下,但随着阈值的增加,检测性能往往会下降。原因包括:1)正样本数量的减少很容易导致过拟合;2)当RPN输出的候选框阈值与训练时设置的阈值相差太远时,会出现不匹配问题。为避免这些因素影响检测性能,本文使用级

联检测器,提高网络的定位精度。分类和位置预测网络结构图如图3所示。该结构通过一系列回归器实现,具体为

$$f(x, b) = f_T \circ f_{T-1} \circ \dots \circ f_1(x, b) \quad (2)$$

式中,“ \circ ”表示点乘, T 表示级联的总数, b 代表候选框 BBox, x 代表图像批次。其中,级联中的每个回归因子 f_r 都经过了边界位置优化,表明本次回归使用该阶段对应的候选框分布 $\{b'\}$,并非初次回归时的 $\{b\}$,这种方式可逐步改进假设。级联回归本质上为一个重采样过程,每一级的输入都为上一级的输出,可以改变不同阶段要处理的假设分布。并且多元回归器 $\{f_T, f_{T-1}, \dots, f_1\}$ 针对不同阶段重采样的分布进行了优化,定位更加精确。本文选择使用3级级联检测器,每一级的阈值依次设置为0.5, 0.6, 0.7。

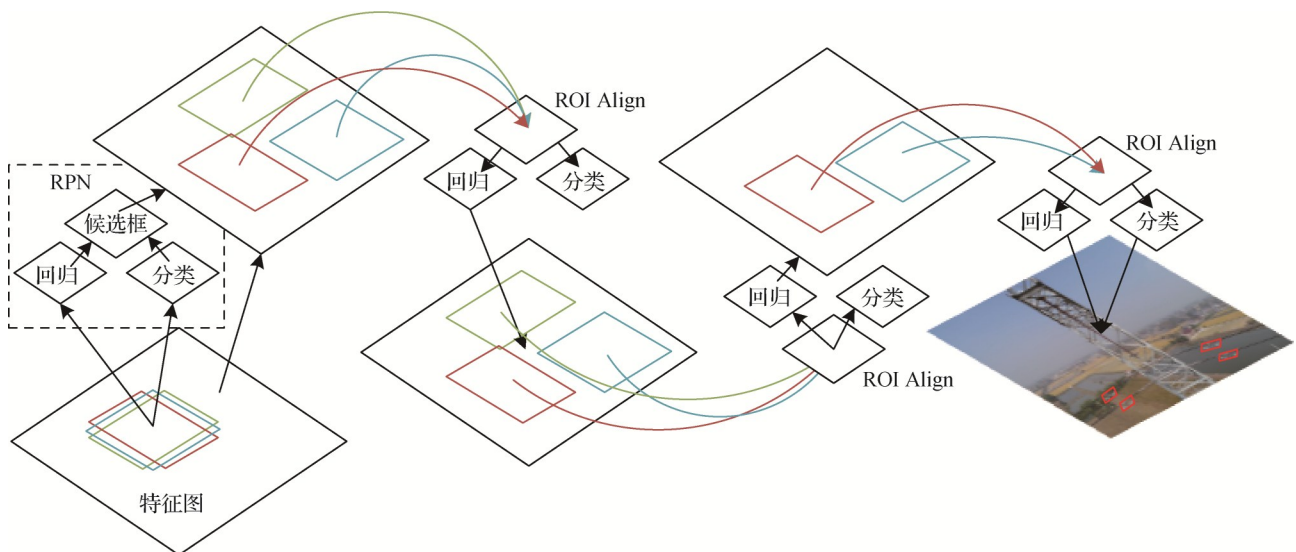


图3 分类和位置预测网络结构

Fig. 3 Classification and location prediction network structure

2 实验及结果分析

2.1 实验数据集

目前国内外没有防震锤缺陷检测相关的公开数据集,为进行防震锤的缺陷检测,本文建立了防震锤缺陷检测数据集,并在该数据集上完成防震锤缺陷检测研究工作。本文数据集的原始图像为无人机对阳泉市的输电线路部件进行航拍巡检时获得。为了保证网络的泛化性以及鲁棒性,本数据集符合实际,包含较多复杂的背景,如树林、天空、田地和住宅等。根据目前防震锤常见缺陷种类并分析航拍采集的原

始图像中防震锤的外观特征和姿态,将数据集将防震锤分为正常、腐蚀、破损和相撞4类,如图4所示。

对数据集扩充后,数据集共有图像4 200幅,样本组成情况如表1所示。随机将图像分为训练集和测试集,训练集与测试集的比例为8:2。

2.2 实验设置

实验在Linux中进行,采用Caffe(convolutional architecture for fast feature embedding)深度学习框架搭建本文网络模型。所使用计算机处理器为Intel(R) Core(TM) i7-10700 CPU @2.90 GHz,显卡为NVIDIA GeForce RTX 3070。

本文网络训练过程如下:

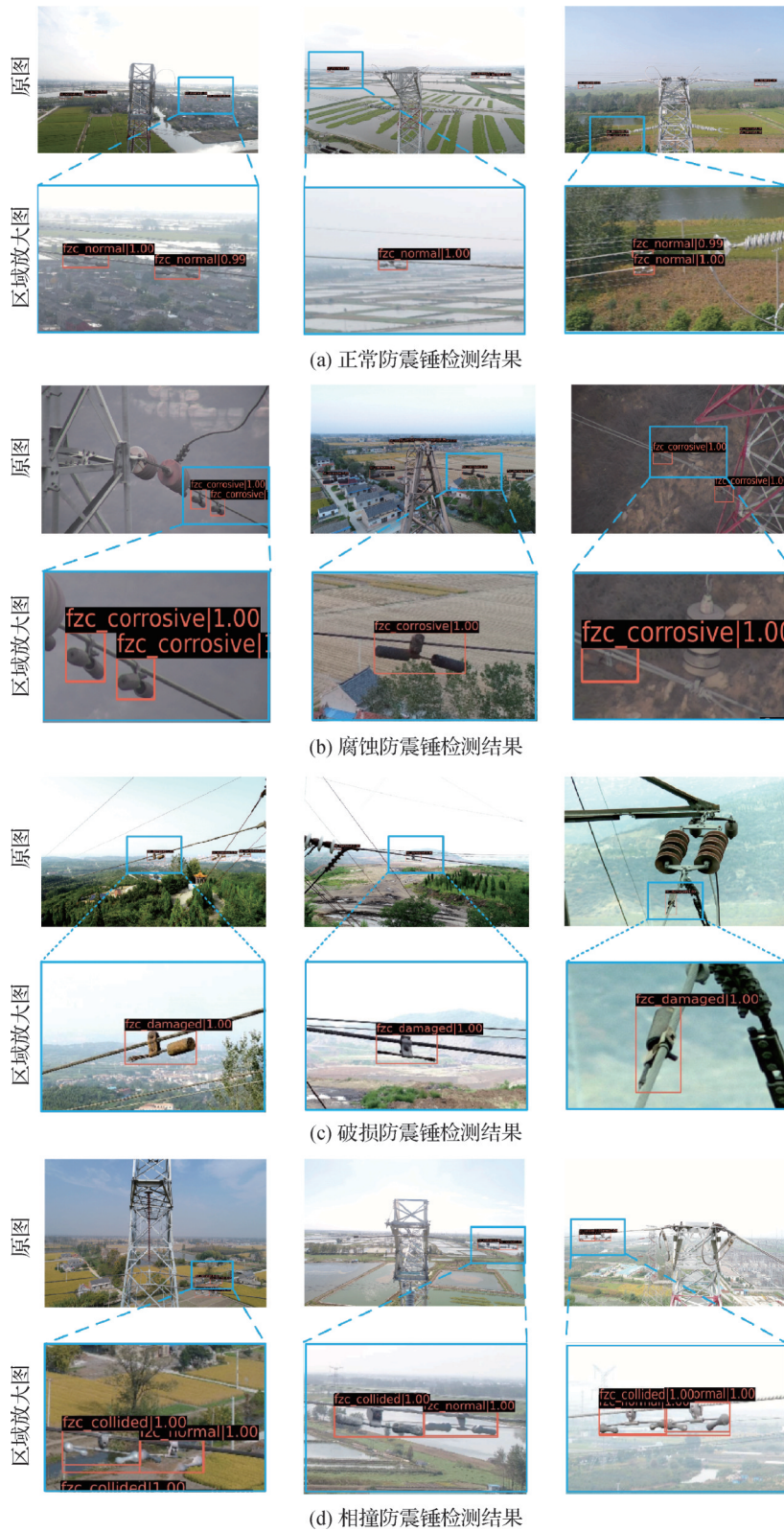


图4 防震锤缺陷检测数据集示例

Fig. 4 Examples of vibration hammer defect detection dataset ((a) normal anti-vibration hammer test results; (b) corrosion anti-shock hammer test results; (c) broken anti-vibration hammer test results; (d) collision anti-shock hammer test results)

- 1) 预训练跨层融合的VGG16网络;
- 2) 用步骤1)的模型初始化RPN,并训练RPN;
- 3) 用步骤1)的模型初始化分类和位置预测网络,然后使用经过训练的RPN计算候选区域,并提

表1 数据集样本组成情况

Table 1 Sample composition of the dataset

类别	标签	数量/幅
正常	fzc_normal	1 000
腐蚀	fzc_corrosive	1 200
损坏	fzc_damaged	1 000
相撞	fzc_collided	1 000

交给分类和位置预测网络;

4)微调步骤2)中的RPN网络生成共享特征层;

5)使用步骤4)生成的候选区域,微调用于目标检测的分类和位置预测网络;

6)输出一个在步骤4)和步骤5)中训练的统一网络作为最终模型。

在步骤1)—3)中,分别训练RPN和分类位置预测网络。通过步骤4)和步骤5)对两个网络进行微调,且这两个网络共同使用特征提取网络。最后,本文将两个独立的网络合并成一个统一的网络。对于RPN和分类位置预测网络,本文在训练和微调步骤中,设置初始学习率为0.005,采用step作为学习率下降方式。本文设置迭代次数为80k,在每次迭代中,每个图像的感兴趣区域(region of interest, ROI)数量设置为64个。本文参数中动量等于0.9,权重衰减系数等于0.0001。

2.3 防震锤缺陷检测结果

实验在本文建立的防震锤缺陷检测数据集上进行,将本文模型结果与其他基于深度学习的先进算法对比,并分析本文算法的改进策略以证明其有效性。最后进行错误检测分析,分析算法的误检率和漏检率。

2.3.1 实验结果与分析

首先分析本文模型训练时损失值和准确率随迭代次数变化情况,如图5和图6所示。本文设计的算法模型在训练过程中损失值和准确率可以基本达到稳定,在训练起始阶段损失值较大,准确率较低。随着迭代次数(iter)的增加和学习率的不断调整,损失值骤降至大约1.0,然后开始缓慢下降,在0.1左右整个曲线趋于平滑,与此同时准确率急剧上升,最终模型训练时准确率可以稳定在99.5%左右。因此整个训练过程中的损失值持续下降,准确率持续上升,最终曲线基本稳定,说明模型在训练过程中可以

较好地学习到目标的特征。

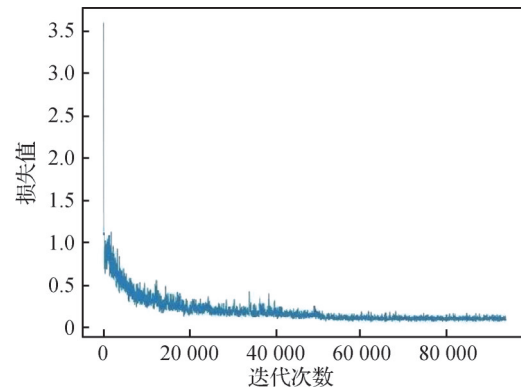


图5 模型训练时损失变化情况

Fig. 5 Variation of loss during model training

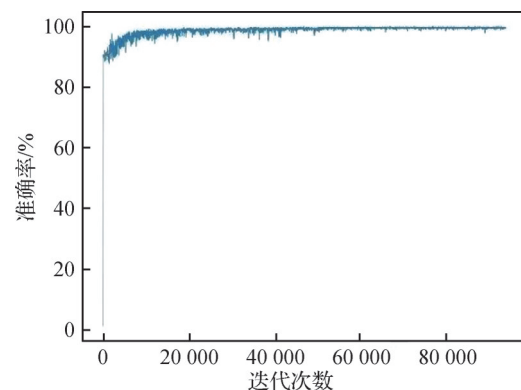


图6 模型训练时准确率变化情况

Fig. 6 Variation of accuracy during model training

因此,整个模型训练时设置的网络参数是本文方法的网络结构中较合适的参数值。为了验证本文模型的检测效果,对常用的目标检测算法在本文建立的防震锤缺陷检测数据集上的检测结果进行分析。

表2是不同方法的网络结构参数量,表3是不同方法的测试结果。检测结果为每类防震锤的平均精度(average precision, AP)以及平均精度均值(mean average precision, mAP),由表3可以得出,本文方法的mAP可以达到95.3%,是对比算法中精度最高的,并且Faster R-CNN的检测精度高于SSD(single shot MultiBox detector)、Retinanet、YOLO系列,原因是这些检测算法为提高检测速度放弃了部分检测精度,但是本文数据集为小尺度目标,更需要关注目标的检测精度。Faster R-CNN的mAP为91.9%,本文方法的精度与Faster R-CNN相比提高了3.4%,每一类的AP也有相应提高。

本文方法融合了多层特征,使模型能够得到更

表2 不同方法的网络结构参数量

Table 2 The amount of network structure parameters of different methods

方法	参数量/个
SSD300	23 878 495
YOLOv3	40 378 960
RetinaNet	170 691 895
YOLOv4	57 671 300
Fast R-CNN	69 765 442
Faster R-CNN	61 834 673
本文	79 108 519

表3 不同方法在防震锤缺陷检测结果对比

Table 3 Comparison of anti-shock hammer defect detection results of different methods

方法	主干网络	检测准确率				mAP
		正常	腐蚀	破损	相撞	
SSD300	VGG16	85.8	85.3	84.5	87.8	85.8
YOLOv3	DarkNet-53	86.4	84.8	85.9	89.0	86.5
RetinaNet	ResNet101	88.3	85.7	85.1	89.9	87.3
YOLOv4	CSPDarknet-53	89.6	88.4	88.5	91.7	89.5
Fast R-CNN	VGG16	82.6	80.4	79.5	84.5	82.0
Faster R-CNN	VGG16	92.2	90.0	90.0	94.3	91.9
本文	VGG16-Cross	95.6	94.0	94.3	97.2	95.3

注:加粗字体表示各列最优结果。

多低层和高层图像特征信息,并且还使用级联检测器降低了IoU阈值对网络性能的影响,最终显著提高了小尺度目标的检测精度。

本文模型对防震锤缺陷检测数据集中4个缺陷类型的防震锤测试的效果图如图4所示。从图中可以看出,在不同背景、不同光照、不同天气下,不同角度、不同尺度和不同种类的防震锤均表现出很好的检测效果,所以本文提出的防震锤缺陷检测算法具有较强的鲁棒性。

2.3.2 特征融合策略对比分析

特征图的选取是保证后续网络对目标的分类和位置预测精度的关键。在训练时采用不同的特征融合策略,以研究融合不同层的特征得到的特征图对算法检测精度的影响。融合策略包括:1)融合第1层和第3层;2)融合第3层和第5层;3)融合第1、2、

3层;4)融合第3、4、5层;5)本文特征融合策略,即融合第1、3、5层。不同特征融合策略在防震锤缺陷检测数据集上的精度如表4所示。

本节所有网络的训练使用相同配置和参数。根据表4,融合第1、3、5层特征的网络检测的效果最好,对应的mAP达到95.3%,而其他融合3层特征的策略效果较差的原因是相邻层之间的特征具有很强的相关性,无论是低层还是高层融合的效果都不好。这也表明更广泛的由粗到细的卷积层特征融合的网络检测效果最好。本文提出的融合策略能够保证较高的小尺寸目标的检测精度,在考虑高层特征具有较多的语义信息的同时还考虑到低层特征中丰富的空间特征和纹理信息。

表4 特征融合策略对比

Table 4 Comparison of feature fusion strategies

层	召回率	mAP
1+3	94.2	91.5
3+5	95.0	93.7
1+2+3	94.7	92.8
3+4+5	95.4	94.1
1+3+5	96.1	95.3

注:加粗字体表示各列最优结果。

2.3.3 消融实验

通过消融实验可分析跨层特征融合、Soft-NMS和级联检测器对本文改进后的Faster R-CNN模型的影响。具体结果如表5所示;Improve代表加入3种改进的其中一种,例如Improve1代表加入第1种改进融合特征;Combine代表加入3种改进的其中两种,例如Combine23代表加入第2种改进和第3种改进,即组合Soft-NMS和级联检测器。以Faster R-CNN为基础进行相应的算法改进,表5中对比了单独加入3种改进以及对3种改进两两组合加入Faster R-CNN中时模型的精度,网络训练过程中其余参数和结构相同。从表5中可以看出,每种方法都在一定程度上提高了模型的精度,跨层特征融合、Soft-NMS和级联检测器分别将模型精度提高了1.6%、0.2%和2.0%,其中跨层特征融合和级联检测器都极大地提高了模型的精度。在对3种改进进行两两组合时模型的精度以及网络性能均有更大提

升。实验结果表明,本文提出的针对Faster R-CNN的改进策略对于提高输电线路中防震锤缺陷的检测效率具有重要意义。

表5 消融实验结果

Table 5 Results of ablation experiments

方法	融合特征	Soft-NMS	级联检测器	mAP/%
Faster R-CNN	-	-	-	91.9
Improve1	√	-	-	93.5
Improve2	-	√	-	92.1
Improve3	-	-	√	93.9
Combine12	√	√	-	93.8
Combine23	-	√	√	94.2
Combine13	√	-	√	95.0
本文	√	√	√	95.3

注:加粗字体表示最优结果,“√”表示采用,“-”表示未采用。

2.3.4 错误检测分析

本文设计的防震锤缺陷检测算法得到了95.3%的平均精度,可以认为符合实际防震锤缺陷检测的要求,但是对防震锤缺陷仍无法完全检测,经分析主要原因因为误检和漏检。接下来针对本文检测结果中的误检和漏检情况及原因进行分析。

对数据集中误检情况进行分析,得出误检原因主要是将塔架或复杂背景中部分干扰物误认为防震锤,误检示意图如图7所示,展示情况为将塔架部分特征误认为破损防震锤。这些数据中,塔架或复杂背景中部分特征与防震锤的特征过于相似,导致防震锤与这部分特征发生混淆。由于防震锤安装位置距离塔架较近,数据中无法避免复杂背景和塔架的出现。

对数据集中漏检情况进行分析,得出漏检原因主要是防震锤本身颜色等特征与图像背景颜色极为相似,同时由于拍摄角度问题防震锤会被线路遮挡

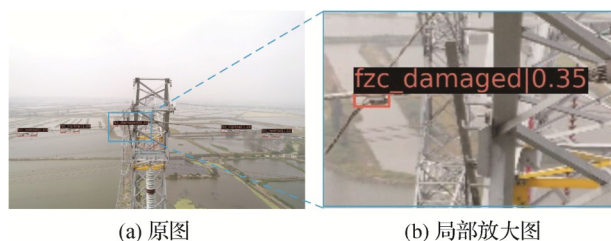
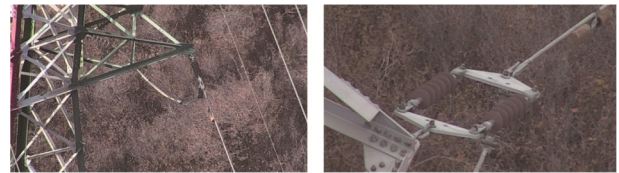


图7 误检情况示例

Fig. 7 Example of anti-vibration hammer misdetection situation ((a) original drawing; (b) fractionated gain)

主要特征,两种原因叠加导致漏检情况的出现,如图8所示。并且此类数据较少,网络无法较好地学习到该情况下的防震锤特征。在实际应用中,由于无法避免该类背景的出现,可通过在巡检过程中寻找较好的拍摄角度避免漏检。



(a) 拍摄角度不佳 (b) 目标与背景颜色相似

图8 漏检情况示例

Fig. 8 Example of anti-vibration hammer leak detection ((a) poor shooting angle; (b) targets are similar in color to the background)

对这两种情况出现的原因进行分析发现,检测结果受数据集的影响较大,关于防震锤缺陷检测的数据较少,并且数据集中困难样本更少。

本文对比了Faster R-CNN和本文方法的漏检率和误检率,结果如表6所示,虽然本文无法避免漏检率和误检率的出现,但是本文方法在提高精度的同时有效降低了误检率和漏检率,在实际应用中效果更好。

表6 不同方法下错误检测对比

Table 6 Comparison of error detection under different methods

方法	mAP	误检率	漏检率
Faster R-CNN	91.9	9.08	4.33
本文	95.3	3.47	1.32

注:加粗字体表示最优结果。

3 结论

针对防震锤这类小尺度目标检测精度低、防震锤缺陷缺乏细化检测等问题,本文提出基于跨层特征融合和级联检测器的防震锤缺陷检测方法。主要工作如下:

1)利用无人机对输电线路进行航拍,采集包含防震锤的航拍图像,经过筛选和数据扩充等处理后建立防震锤数据集并细化了防震锤缺陷种类。

2)网络模型采用VGG16为主干网络,通过对其

1、3、5层进行跨层融合,将底层特征与高层特征结合,得到更加准确地防震锤特征图。

3)对分类和位置预测网络部分引入3个级联检测器,逐渐增加IoU阈值提高样本的质量和网络的训练效果。

最后在自建数据集上进行本文网络模型与其他6种基于深度学习的先进算法的对比实验,实验结果表明本文算法的检测效果最佳,其检测精度与其他6种算法相比均有不同程度的提高;消融实验结果表明本文提出的跨层特征融合、Soft-NMS和级联检测器分别将模型精度提高了1.6%、0.2%和2.0%,而且本文方法较Faster-RCNN的误检率和漏检率均有降低。因此上述实验共同验证了本文方法的鲁棒性和有效性。但是本文算法主要关注小尺度目标防震锤的缺陷检测精度,并未关注算法实时性,下一步将围绕着该算法的实时性展开深入研究。

参考文献(References)

- Bodla N, Singh B, Chellappa R and Davis L S. 2017. Improving object detection with one line of code [EB/OL]. [2022-08-22]. <https://arxiv.org/pdf/1704.04503.pdf>
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40 (4): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Chen X J, Wu Y S and Zhao L. 2010. Identification of OPGW vibration damper based on random Hough transformation. *Heilongjiang Electric Power*, 32(1): 1-2, 5 (陈晓娟, 吴英石, 赵亮. 2010. 基于随机Hough变换的OPGW防震锤识别. *黑龙江电力*, 32(1): 1-2, 5) [DOI: 10.13625/j.cnki.hljep.2010.01.021]
- Chen X Z. 2021. Application and prospect of transmission line UAV inspection. *China Power Enterprise Management*, (36): #86 (陈昕哲. 2021. 输电线路无人机巡检应用及展望. *中国电力企业管理*, 36): #86) [DOI: 10.3969/j.issn.1007-3361.2021.36.040]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Huang L L, Xie R and Xu Y. 2015. Invasion detection on transmission lines using saliency computation//Proceedings of 2015 IEEE International Symposium on Signal Processing and Information Technology. Abu Dhabi, United Arab Emirates: IEEE: 320-325 [DOI: 10.1109/ISSPIT.2015.7394352]
- Jia L Y, Han J and Yu H F. 2021. Shock hammer recognition based on semantic segmentation and spatial context information. *Computer Measurement and Control*, 29(1): 200-205 (贾立业, 韩军, 余鸿飞. 2021. 基于语义分割和空间上下文信息的防震锤识别. *计算机测量与控制*, 29(1): 200-205) [DOI: 10.16526/j.cnki.11-4762/tp.2021.01.041]
- Jia Y F, Chen G D, Yang M, Xing L Y, Zhao L Q and Li S Y. 2022. Damper detection of high voltage transmission line based on deep learning. *Machine Tool and Hydraulics*, 50(13): 21-25 (贾雁飞, 陈广大, 杨森, 邢砾云, 赵立权, 李帅洋. 2022. 基于深度学习的高压输电线路防振锤检测. *机床与液压*, 50(13): 21-25 [DOI: 10.3969/j.issn.1001-3881.2022.13.004]
- Jiao R T, Ni H X and Wang Z Y. 2021. Identification of Transmission line Damper based on Faster RCNN Algorithm. *Journal of Changchun Institute of Technology (Natural Science Edition)*, 22 (1): 38-43 (焦润童, 倪虹霞, 王智昱. 2021. 基于Faster RCNN算法的输电线路防震锤识别研究. *长春工程学院学报(自然科学版)*, 22 (1): 38-43) [DOI: 10.3969/j.issn.1009-8984.2021.01.010]
- Jin L J, Yan S J and Liu Y. 2012. Vibration damper recognition based on Haar-Like features and cascade AdaBoost classifier. *Journal of System Simulation*, 24(9): 1806-1809 (金立军, 闫书佳, 刘源. 2012. 基于类Haar特征与级联AdaBoost算法的防震锤识别. *系统仿真学报*, 24(9): 1806-1809) [DOI: 10.16182/j.cnki.joss.2012.09.022]
- Liang H G, Zuo C and Wei W M. 2020. Detection and evaluation method of transmission line defects based on deep learning. *IEEE Access*, 8: 38448-38458 [DOI: 10.1109/ACCESS.2020.297498]
- Liu X Y, Jiang H, Chen J, Chen J J, Zhuang S B and Miao X R. 2018. Insulator detection in aerial images based on faster regions with convolutional neural network//Proceedings of the 14th IEEE International Conference on Control and Automation (ICCA). Anchorage, USA: IEEE: 1082-1086 [DOI: 10.1109/ICCA.2018.8444172]
- Lyu M, Sheng G H, Zhang W D and Jiang X C. 2012. A tower and line tracking algorithm for power transmission line inspection based on unmanned aerial vehicles. *Automation of Electric Power Systems*, 36(9): 92-97 (吕明, 盛戈峰, 张卫东, 江秀臣. 2012. 无人飞行器巡检输电线路的杆塔和导线跟踪算法. *电力系统自动化*, 36(9): 92-97)
- Pan L Z and Xiao X H. 2009. Image recognition for on-line vibration monitoring system of transmission line//Proceedings of the 9th International Conference on Electronic Measurement and Instruments. Beijing, China: IEEE: 1078-1081 [DOI: 10.1109/ICEMI.2009.5274379]
- Qi Y C, Wu X L, Zhao Z B, Shi B Q and Nie L Q. 2021. Bolt defect detection for aerial transmission lines using faster R-CNN with an embedded dual attention mechanism. *Journal of Image and Graphics*, 26(11): 2594-2604 (戚银城, 武学良, 赵振兵, 史博强, 聂礼强. 2021. 嵌入双注意力机制的Faster R-CNN航拍输电线路螺栓缺陷检测. *中国图象图形学报*, 26(11): 2594-2604)

- [DOI: 10.11834/jig.200793]
- Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement [EB/OL]. [2022-08-22]. <https://arxiv.org/pdf/1804.02767.pdf>
- Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]
- Simonvan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2022-08-22]. <https://arxiv.org/pdf/1409.1556.pdf>
- Song W, Zuo D, Deng B F, Zhang H B, Xue K W and Hu H. 2016. Corrosion defect detection of earthquake hammer for high voltage transmission line. *Chinese Journal of Scientific Instrument*, 37(S1): 113-117 (宋伟, 左丹, 邓邦飞, 张海兵, 薛凯文, 胡泓. 2016. 高压输电线防震锤锈蚀缺陷检测. *仪器仪表学报*, 37(S1): 113-117) [DOI: 10.19650/j.cnki.cjsi.2016.s1.019]
- Wang Q L, Wu B G, Li P F, Li P H, Zuo W M and Hu Q H. 2020. ECA-Net: efficient channel attention for deep convolutional neural network//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 11531-11539 [DOI: 10.1109/CVPR42600.2020.01155]
- Xu B Y, Chen Q N, Wang S W, Shen F and Wu Y B. 2021. Implementation and research of automatic line patrol based on 5G UAV in power transmission line. *Yunnan Electric Power*, 49(6): 8-12 (许保瑜, 陈庆宁, 王胜伟, 沈凤, 吴圆波. 2021. 基于5G无人机在电力输电线路自动巡线的实现与研究. *云南电力技术*, 49(6): 8-12) [DOI: 10.3969/j.issn.1006-7345.2021.06.002]
- Yang G, Sun C W, Wang D W, Jin T, Xu C Y, Lu Z B and Zhang X Z. 2020. Comparative study of transmission line component detection models based on UAV front end and SSD algorithm. *Journal of Taiyuan University of Technology*, 51(2): 212-219 (杨罡, 孙昌雯, 王大伟, 晋涛, 徐澄宇, 卢志博, 张兴忠. 2020. 基于无人机前端和SSD算法的输电线路部件检测模型对比研究. *太原理工大学学报*, 51(2): 212-219) [DOI: 10.16355/j.cnki.issn1007-9432tyut.2020.02.009]
- Zhang Y X, Wu G P, Liu Z Y, Yang S and Xu W Z. 2020. Transfer learning of transmission line damper and clamp detection based on YOLOv3 network. *Journal of Computer Applications*, 40(S2): 188-194 (张永翔, 吴功平, 刘中云, 杨松, 徐玮泽. 2020. 基于YOLOv3网络的输电线路防震锤和线夹检测迁移学习. *计算机应用*, 40(S2): 188-194) [DOI: 10.11772/j.issn.1001-9081.2020020276]
- Zhao Z B, Jiang Z G, Li Y X, Qi Y C, Zhai Y J, Zhao W Q and Zhang K. 2021. Overview of visual defect detection of transmission line components. *Journal of Image and Graphics*, 26(11): 2545-2560 (赵振兵, 蒋志钢, 李延旭, 戚银城, 翟永杰, 赵文清, 张珂. 2021. 输电线路部件视觉缺陷检测综述. *中国图象图形学报*, 26(11): 2545-2560) [DOI: 10.11834/jig.200689]

作者简介

梁华刚,男,副教授,主要研究方向为机器视觉、数字图像处理、模式识别和智能交通。E-mail: hgliang@chd.edu.cn

赵慧霞,通信作者,女,硕士研究生,主要研究方向为模式识别和激光SLAM。E-mail: 18809299832@163.com

刘丽华,女,硕士研究生,主要研究方向为深度学习和模式识别。E-mail: liulihua0925@163.com

岳鹏,男,硕士研究生,主要研究方向为人工智能和机械臂深度学习。E-mail: 2020232057@chd.edu.cn

郑振宇,男,硕士研究生,主要研究方向为图像处理和模式识别。E-mail: 2021232009@chd.edu.cn