

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2023)11-3497-12

论文引用格式: Yan G W, Zhou X J, Jiao R H and He H. 2023. Defect detection of tower bolts by fusion of priori information and feature constraints. Journal of Image and Graphics, 28(11):3497-3508(阎光伟, 周香君, 焦润海, 何慧. 2023. 融合先验信息和特征约束的杆塔螺栓缺陷检测. 中国图象图形学报, 28(11):3497-3508)[DOI:10.11834/jig.221077]

融合先验信息和特征约束的杆塔螺栓缺陷检测

阎光伟, 周香君, 焦润海*, 何慧

华北电力大学控制与计算机工程学院, 北京 100096

摘要: 目的 螺栓是输电线路中数量最多的紧固件, 一旦出现缺陷就会影响电力系统的稳定运行。针对螺栓缺陷自动检测中存在的类内多样性和类间相似性挑战, 提出了一种融合先验信息和特征约束的 Faster R-CNN (faster regions with convolutional neural network) 模型训练方法。方法 在航拍巡检图像预处理阶段, 设计了基于先验信息的感兴趣区域提取算法, 能够提取被识别目标的上下文区域, 从而减少模型训练阶段的数据量, 帮助模型在训练阶段关注重点区域, 提高其特征提取能力。在模型训练阶段, 首先通过费舍尔损失约束 Faster R-CNN 模型的输出特征生成, 使样本特征具有较小的类内距离和较大的类间间隔; 然后采用 K 近邻算法处理样本特征得到 K 近邻概率, 将其作为难易样本的指示以引导模型后续更加关注难样本。结果 在真实航拍巡检图像构建的螺栓数据集上进行测试, 与基线模型相比, 本文模型使螺栓识别的平均精度均值 (mean average precision, mAP) 提高了 6.4%, 其中正常螺栓识别的平均精度 (average precision, AP) 提高了 0.9%, 缺陷螺栓识别的平均精度提高了 12%。结论 提出的融合先验信息和特征约束的输电杆塔螺栓缺陷检测方法在缺陷螺栓识别上获得了良好的效果, 为实现输电线路螺栓缺陷的自动检测奠定了良好的基础。

关键词: 电力巡检; 螺栓缺陷检测; 类内多样性; 类间相似性; 先验信息; 特征约束; Faster R-CNN

Defect detection of tower bolts by fusion of priori information and feature constraints

Yan Guangwei, Zhou Xiangjun, Jiao Runhai*, He Hui

School of Control and Computer Engineering, North China Electric Power University, Beijing 100096, China

Abstract: **Objective** Bolts, as fasteners, play a role in connecting and fixing key components in power towers. Adverse environments, mechanical vibrations, and material aging may cause the bolt cotter pins to fall off or come out, thus impacting the normal operations of other parts on the transmission line. Monitoring, testing, and maintenance of bolt conditions play crucial roles in ensuring the uninterrupted operation of a power system. However, the timely and efficient automatic detection of transmission line bolt defects in a power system poses a challenge. Aiming at the challenges of intra-class diversity and inter-class similarity in automatic bolt defect detection, a faster regions with convolutional neural network (Faster R-CNN) model training method based on prior information and feature constraints is proposed in this paper. **Method** In the pre-processing of an aerial inspection image, this paper designs a region of interest extraction algorithm based on prior information that can extract the context region of the identified object so as to reduce the amount of data in the training stage,

收稿日期: 2022-11-21; 修回日期: 2023-02-27; 预印本日期: 2023-03-06

* 通信作者: 焦润海 runhaijiao@ncepu.edu.cn

基金项目: 国家自然科学基金项目 (62073133)

Supported by: National Natural Science Foundation of China (62073133)

help the model focus on the key areas in the training stage, and improve its feature extraction ability. In the model training stage, the output features of Faster R-CNN model are initially constrained by the Fisher constraint to reduce the intra-class distances of the sample features while increasing their inter-class intervals, which improves the differentiation of sample features. Afterward, the K -nearest neighbor algorithm is used to process the sample features to obtain the k -nearest neighbor probability, which is used as an indication of difficult samples to make the model pay attention to difficult samples in the future. **Result** All data used in this paper are real aerial inspection images from a power company in Central China. Bolts on the power tower can be divided into bolts with and without pins. Bolts with pins connect and fix the key components and need to bear a large force. Compared with bolts without pins, those with pins are more prone to defects. Therefore, this paper treats bolts with pins as the research object. According to the state of the cotter pin on the bolt, bolts with pins are divided into normal and defective bolts, of which defective bolts have fallen or loose cotter pins. The dataset contains 28 887 images, with each image having a resolution of $5\,000 \times 3\,000$ pixels. The training set, validation set, and test set are divided according to the ratio of 8:1:1. The proposed model is then tested on this dataset. Compared with the baseline model, the proposed model improves the mean average precision of bolt recognition by 6.4%, increases the average precision of normal bolt identification by 0.9%, and increases the average accuracy of defective bolt identification by 12%. **Conclusion** In this paper, the problems of intra-class diversity and inter-class similarity in bolt defect detection are explored, and a method of transmission tower bolt defect detection based on the fusion of prior information and feature constraints is proposed. This improves the recognition effect of the model on bolt defects and lays a good foundation for the automatic detection of bolt defects in transmission lines.

Key words: power inspection; bolt defect detection; intra-class diversity; inter-class similarity; prior information; feature constraints; faster regions with convolutional neural network (Faster R-CNN)

0 引言

螺栓作为紧固件在电力杆塔中起着连接、固定关键部件的作用(赵振兵等,2020)。恶劣环境(如冰雹、酸雨和大风)、机械振动和材料老化等因素会导致螺栓的开口销脱落或脱出,影响输电线路其他部件的正常运行(曾勇斌等,2020)。为了保证电力系统的不间断运行,螺栓状态的监测、检测和维护发挥着至关重要的作用。过去几十年,电网公司采用了多种检查方法,包括徒步巡逻、直升机和攀爬机器人。无人机巡检因其便捷、安全及成本较低的特点被电力部门广泛采用。但航拍巡检图像数量庞大且价值密度低,依靠审核人员对海量巡检数据进行检查的方式效率低下。因此,基于计算机视觉技术的电力智能巡检开始兴起,并取得了一定的研究进展(赵振兵等,2021)。

输电杆塔上存在带销和不带销螺栓,二者仅靠销钉或销孔这种细微特征区分。当天气环境和无人机拍摄角度不佳时,不同类螺栓的视觉形态相似。无人机在不同角度拍摄时,螺栓的三维结构会产生不同的二维视觉信息,这使得航拍图像中同类螺栓

的视觉表现不一致(Toth 和 Gilpin-Jackson, 2010)。此外,光照和遮挡也使螺栓呈现多种形态。图1展示了螺栓数据集中存在的类内多样性和类间相似性问题,这对螺栓缺陷检测造成很大的干扰。

研究人员一直在研究快速、准确的螺栓缺陷检测方法。早期研究大都基于传统图像处理算法。徐丹等人(2017)和冯敏等人(2018)分别提取螺栓的方向梯度直方图和局部二值模式后使用支持向量机分类。这种基于人工设计特征的缺陷检测方法难以提取到小尺寸螺栓的有效特征,无法真正应用于复杂的电力巡检场景(赵振兵等,2021)。

自从卷积神经网络在图像分类任务中取得突破性成功,许多基于卷积神经网络的目标检测方法被提出,如两阶段 R-CNN (regions with convolutional neural network) 系列(Girshick 等, 2014; Girshick, 2015; Ren 等, 2017)和一阶段 SSD (single shot multi-box detector)(Liu 等, 2016)以及 YOLO (you only look once) 系列(Redmon 等, 2016; Redmon 和 Farhadi, 2017, 2018; Bochkovskiy 等, 2020; Jocher 等, 2021)。这些方法在通用目标检测数据集上表现良好,因此研究人员开始使用上述算法识别电力巡检数据集中的螺栓缺陷。Jiao 等人(2021)在 Faster R-CNN 的基

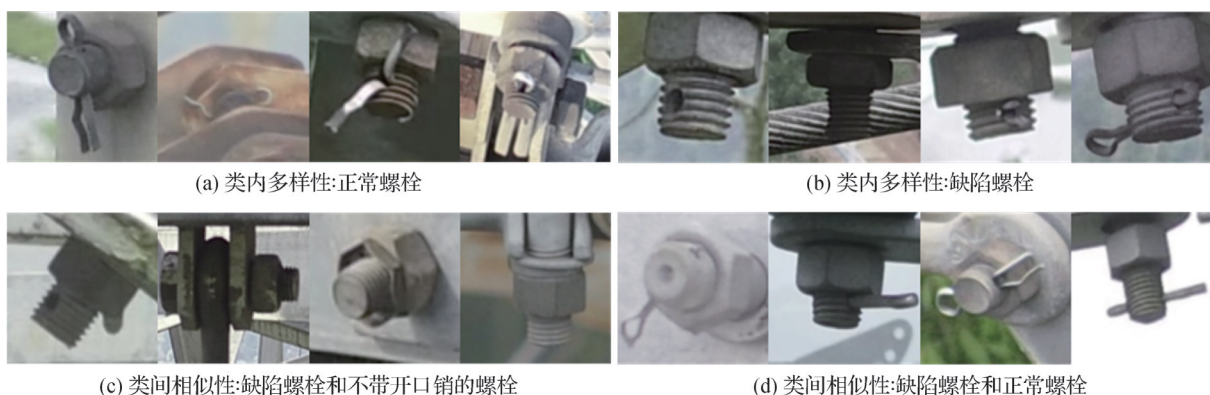


图1 类内多样性和类间相似性

Fig. 1 Intra-class diversity and inter-class similarity ((a) diversity within class of normal bolts; (b) diversity within class of defective bolts; (c) similarity between classes of defective bolts and bolts without pins; (d) similarity between classes of defective bolts and normal bolts)

础上引入残差单元和多尺度特征金字塔以弥补小尺寸螺栓因下采样信息丢失严重的问题。戚银城等人(2021)提出一种双注意力机制方法,分别对不同尺度和不同位置的视觉特征进行分析和增强,增大螺栓与背景的特征差异程度。李雪峰等人(2021)针对带销螺栓设计了 PinFPN (pin feature pyramid network) 网络,加强对低层语义信息和位置信息的提取能力。赵文清和徐敏夫(2022b)以 SSD 模型为基础,加入隔级交叉特征金字塔结构,增强特征图的视觉信息和语义信息。赵文清等人(2022a)在 YOLOv5s (YOLOv5 small) 的基础上嵌入 CA (coordinate attention) 注意力模块,并新增一条“自顶向下”特征信息传递路径,增强对螺栓特征信息的学习能力。

以上方法重点针对螺栓物理体积小、特征语义信息不强以及背景复杂等问题进行算法设计,对螺栓存在的类内多样性和类间相似性问题关注较少。Zhao 等人(2020)通过对螺栓视觉形状聚类来构造关系树,以减少类内数据差异过大而形成的数据矛盾对神经网络学习的误导。但关系树对数据质量和标注质量要求较高,也未考虑类间相似性对模型的影响。

同类螺栓间视觉形态的差异性和不同类螺栓间视觉形态的相似性会加大模型的分类难度,从而影响目标的识别精度。针对这个问题,本文设计了一种基于密度空间聚类 (density-based spatial clustering of applications with noise, DBSCAN) (Ester 等, 1996) 的数据预处理方法,能够将不同类螺栓安装在不同位置的先验信息引入模型,以此来辅助模型分

类。在模型训练阶段,通过为 Faster R-CNN 模型引入并学习费舍尔判别层 (Fisher discriminant layer, FD), 实现视觉特征的类内相聚和类间相离; 然后利用 K 近邻学习区分简单和困难样本,并依据近邻概率调整难易样本对分类损失的贡献,以引导检测模型重点关注难例。

1 研究方法

图2展示了基于 Faster R-CNN 模型构建的螺栓缺陷检测模型的整体训练流程。在数据预处理阶段,利用 DBSCAN 算法对巡检图像进行预裁剪来获取感兴趣区域。预裁剪可以减少模型后续处理的数据量,获得螺栓所在的结构区域信息和抑制复杂背景对模型的干扰。本文对 Faster R-CNN 模型结构进行微调,在第2个全连接层 (full connection layer, FC2) 后添加一个新的全连接层 FD 层。在训练阶段,首先通过对 FD 层的输出特征施加费舍尔约束训练该层,使其输出特征类内距离小、类间间隔大; 然后使用 K 近邻算法对 FD 层的样本特征进行处理,获得 K 近邻概率,并根据 K 近邻概率设计调节因子 w 缩放样本的交叉熵分类损失,引导检测模型重点关注难例。

1.1 聚类裁剪

航拍巡检图像尺寸大,螺栓尺寸小。受计算资源的限制,检测模型训练时会将航拍图像调整到较小的尺寸。这导致原本就不多的螺栓像素信息丢失,使螺栓缺陷检测模型难以提取到有效特征进行

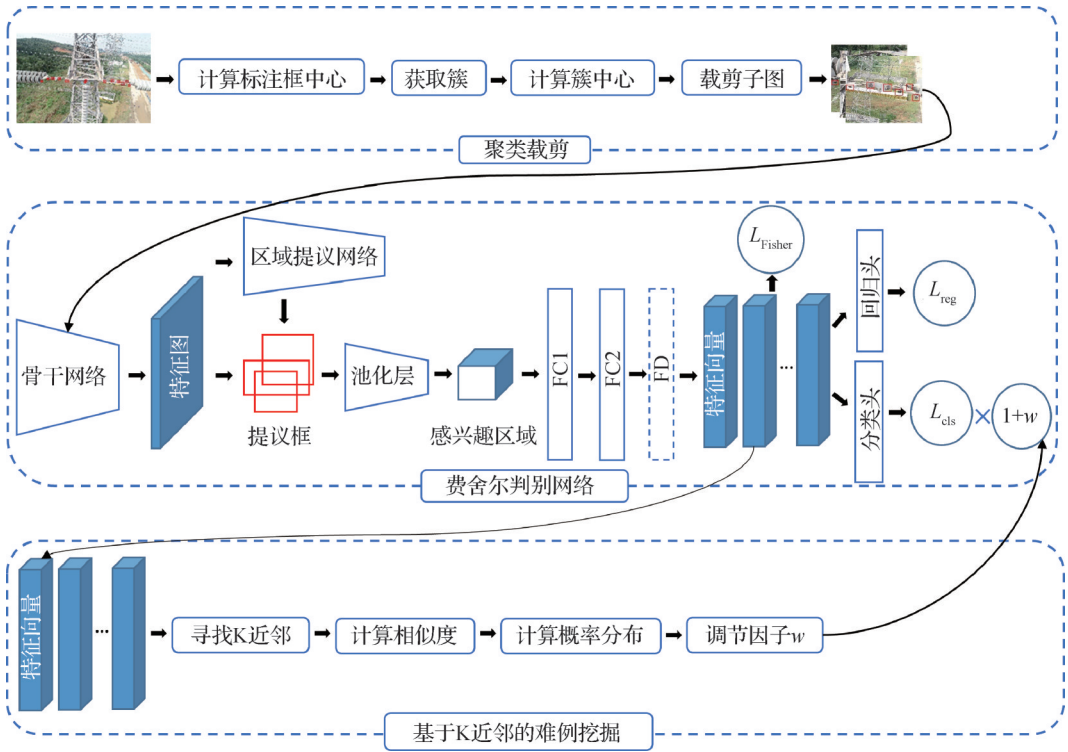


图2 模型训练流程
Fig. 2 Process of model training

学习和训练。
为了解决这个问题,目前常用的解决方案是将原始巡检图像均匀裁剪成多个子图,然后使用这些子图训练模型。均匀裁剪过程如图3所示,在原始图像上从左到右、从上到下依次裁剪固定大小的子图。虽然这缓解了分辨率的挑战,但这种无差别的预裁剪方法不仅产生大量子图,造成计算资源的浪

费,而且破坏了螺栓的区域结构信息。
在电力杆塔上,带销和不带销螺栓被安装在不同的位置。带销螺栓通常安装于三角挂板、绝缘子串连接处。为了将这一先验信息引入检测模型,利用基于密度的空间聚类算法准确提取螺栓出现的上下文区域,然后模型在此区域上做进一步识别。这种预裁剪方法不仅可以减少模型后续处理的数据

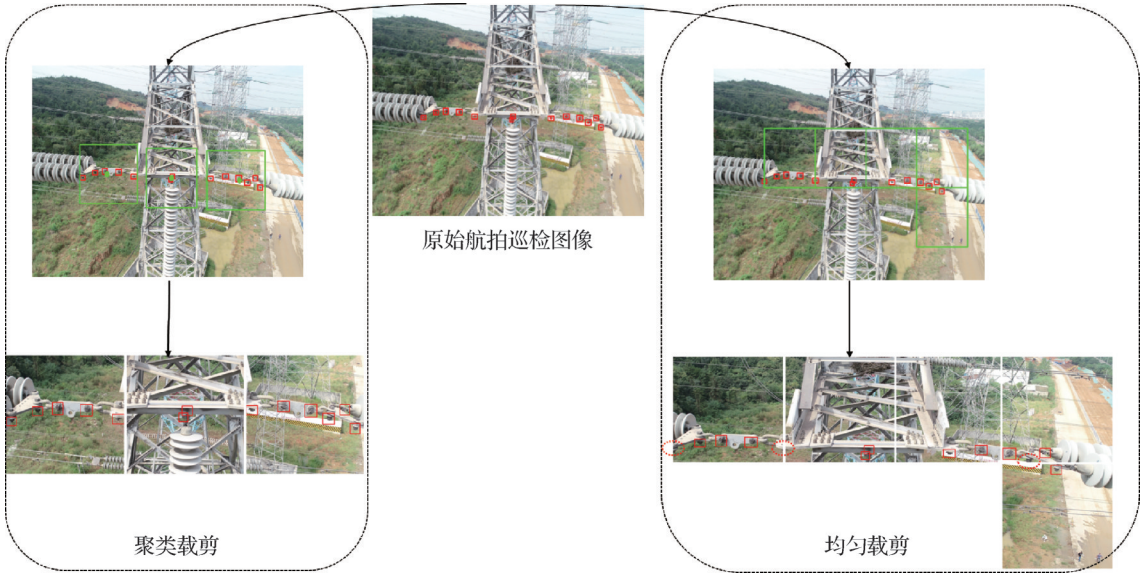


图3 聚类裁剪和均匀裁剪
Fig. 3 Cluster cropping and uniform cropping

量,而且能够帮助模型隐式建模螺栓安装的位置信息,关注重点区域,提高模型的分类能力。

图3左边流程展示了基于聚类算法生成子图边界的过程。首先, $G = \{g_i\}_{i=1}^N$ 表示巡检图像中的目标注释框集合,即图3原始航拍巡检图像中的红色框。 N 表示一幅航拍巡检图像中的注释框总数, $g_i = \{x_i^1, y_i^1, x_i^2, y_i^2\}$ 表示第 i 个注释框, (x_i^1, y_i^1) 和 (x_i^2, y_i^2) 分别表示目标注释框的左上角和右下角坐标。基于 G 可以计算得到这些注释框的中心点坐标集合 $P = \{p_i\}_{i=1}^N$, 即图3原始图像中的红色点。 $p_i = \{x_i, y_i\}$ 标识一个注释框的中心点。然后,对中心点坐标集合 P 使用 DBSCAN 算法进行簇划分,获得簇集合 $C = \{G_i\}_{i=1}^M$ 。 M 表示划分后得到的簇数目,它与每幅图像注释框的分布情况有关,并不是固定的。 G_i 是簇 i 中包含的目标注释框集合。然后,依据每个簇中的注释框中心点坐标计算簇中心集合 $CP = \{cp_i\}_{i=1}^M$, 即图3聚类裁剪流程中的绿色点。 $cp_i = (cx_i, cy_i)$ 表示簇 i 的中心点坐标, cx_i 表示中心点横坐标, cy_i 表示中心点纵坐标。最后,定义子图边界集合为 $B = \{b_i\}_{i=1}^M$, 即图3聚类裁剪流程中的绿色框。第 i 个子图边界 $b_i = \{cx_i, cy_i, w_i, h_i\}$, 它把簇 i 的中心点 (cx_i, cy_i) 做为自己本身的中心点,宽 w_i 和高 h_i 定义为

$$w_i = \max\left(1\ 024, \left(\max\{x_2 | g \in G_i\} - \min\{x_1 | g \in G_i\}\right)\right) \quad (1)$$

$$h_i = \max\left(1\ 024, \left(\max\{y_2 | g \in G_i\} - \min\{y_1 | g \in G_i\}\right)\right) \quad (2)$$

1.2 改进的费舍尔损失

由于拍摄角度、光照强度及遮挡等原因导致航拍图像中同一类别的螺栓视觉形态多样,类内差异大;而各类别螺栓之间仅靠销或销孔这种细微特征区分,类间视觉差异不明显。为了减少类内多样性和类间相似性对样本分类的影响,如图2所示,本文在 Faster R-CNN 模型的 FC2 层后添加一个全连接层,称为费舍尔判别层。通过对其输出特征施加费舍尔约束来学习一个合适的空间变换,使其输出特征类内分散小,类间分隔大。

$\chi_{FD} = \{\chi_{FD}^1, \chi_{FD}^2, \dots, \chi_{FD}^c\}$ 表示 FD 层输出的样本特征,其中, c 表示样本所属类别(包括背景类), χ_{FD}^i

表示第 i 类样本特征集合。除了最小化交叉熵分类损失 ℓ_{cls} 和回归损失 ℓ_{reg} , 费舍尔判别网络还要优化对 χ_{FD} 施加的费舍尔损失 ℓ_{Fisher} , 表示为

$$\ell = \ell_{cls} + \ell_{Fisher} + \ell_{reg} \quad (3)$$

借鉴费舍尔判别分析的思想,费舍尔损失 ℓ_{Fisher} 可以通过最小化 χ_{FD} 的类内散度 S_w 和最大化 χ_{FD} 的类间散度 S_b 实现。 S_w 和 S_b 定义为

$$S_w = \sum_{i=1}^C \sum_{\chi \in \chi_{FD}^i} (\chi - m_i)(\chi - m_i)^T \quad (4)$$

$$S_b = \sum_{i=1}^C n_i (m_i - m)(m_i - m)^T \quad (5)$$

式中, C 表示样本类别数, n_i 表示第 i 类目标的样本数, m_i 和 m 分别为 χ_{FD}^i 和 χ_{FD} 的特征均值,具体为

$$m_i = \frac{1}{n_i} \sum_{\chi \in \chi_{FD}^i} \chi \quad (6)$$

$$m = \frac{1}{n} \sum_{\chi \in \chi_{FD}} \chi \quad (7)$$

式中, n 表示 χ_{FD} 中的样本数量。

与 Cheng 等人 (2016) 的方法不同,本文定义费舍尔损失 ℓ_{Fisher} 为

$$\ell_{Fisher} = \frac{\text{tr}(S_w)}{\text{tr}(S_b)} \quad (8)$$

式中, $\text{tr}(\cdot)$ 表示矩阵的迹。

t-SNE 降维算法 (Van Der Maaten 和 Hinton, 2008) 可以使降维后的特征保持高维空间中的分布,即在高维空间距离近的特征降维之后也很近;在高维空间距离远的特征降维之后也很远。为了观察改进的费舍尔损失对样本特征的约束效果,本文使用 t-SNE 算法对特征进行降维可视化,如图4所示。图4(a)是基线模型(未使用改进的费舍尔损失)中 FC2 层的特征降维后的结果,图4(b)是 FD 层的特征降维后的结果。从两幅图的对比中可以明显看出,改进的费舍尔损失可以使特征的类内距离较小、类间间隔较大。

1.3 基于 K 近邻的难例挖掘

从图4(b)可以看出,样本特征整体呈现出类内相聚和类间相离的特点,但仍有一些特征处于类别交界处,较难区分。受 Focal loss (Lin 等, 2017b) 和 K 近邻用于小样本学习 (Wang 等, 2019)、图像分类 (Jia 等, 2021) 的启发,本文提出使用 K 近邻学习识别这些难例,并引导模型关注它们。如图2所示,可分为以下4个步骤(沿用 1.2 节中的符号描述):

1) 寻找 K 近邻。首先计算 χ_{FD} 中任意两个特征之

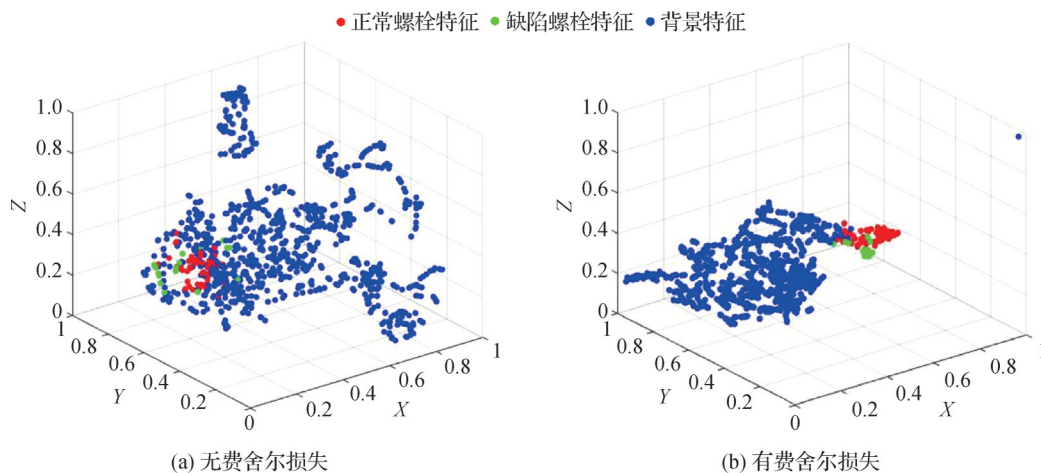


图4 特征降维可视化

Fig. 4 Visualizing dimensionality reduction features ((a) without Fisher loss; (b) with Fisher loss)

间的欧氏距离 $d(\chi_1, \chi_2)$, $\forall \chi_1, \chi_2 \in \chi_{\text{FD}}$ 。然后为每个 χ 选择 K 个除了它本身之外, 距离最近的特征, 记为 M 。 M_c 表示 M 中属于类别 c 的特征集合。

2) 计算相似度。两个特征之间的距离越大, 相似度越低; 距离越小, 相似度越高。定义相似度计算函数为

$$S(\chi_1, \chi_2) = e^{-d(\chi_1, \chi_2)} \quad (9)$$

采用指数衰减的形式主要是考虑放大近邻之间的差异。只要近邻距离变大, 相似度迅速减少。这有助于降低较远的邻居对后续概率计算的影响。

3) 计算概率分布。样本特征 χ 与其 K 个近邻的距离被转换为 C 个类别上的概率分布。样本 χ 被预测为类别 c 的概率为

$$p_{\text{knn}}(c|\chi) = \frac{\sum_{\chi_k \in M_c} S(\chi, \chi_k)}{\sum_{c=1}^C \sum_{\chi_k \in M_c} S(\chi, \chi_k)} \quad (10)$$

4) 难例挖掘。样本被预测为真值类别的概率 p_{knn} 作为难易样本的指示。 p_{knn} 越大, 样本越容易区分; p_{knn} 越小, 样本越难区分。结合图4(b), 相比其他样本, 处于类别交界处的样本的 p_{knn} 会相对较小, 易被模型误判。为引导检测模型更加关注这些误判样本, 本文定义调节因子 $w = (1 - p_{\text{knn}})^2$ 来调整简单样本和困难样本对总损失的贡献。如图5所示, p_{knn} 越大, 调节因子 w 越小, 简单样本对总损失的贡献越小; p_{knn} 越小, 调节因子 w 越大, 困难样本对总损失的贡献越大。最终式(3)可表示为

$$\ell = (1 + w)\ell_{\text{cls}} + \ell_{\text{Fisher}} + \ell_{\text{reg}} \quad (11)$$

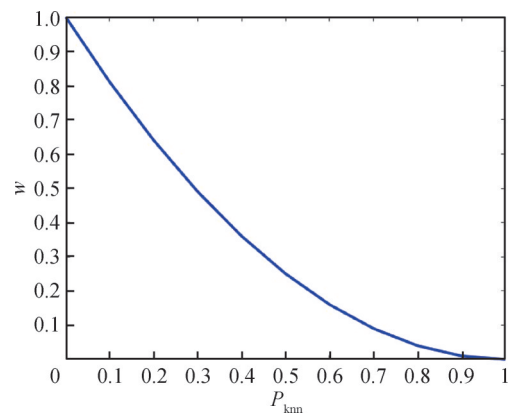


图5 调节因子曲线

Fig. 5 Regulatory factor curve

2 实验结果及分析

2.1 数据集

本文使用的所有数据都是来自华中地区一家电力公司的真实航拍巡检图像。电力杆塔上的螺栓可分为带销螺栓和不带销螺栓。带销螺栓起到连接和固定关键部件的作用, 需要承受较大的作用力, 相比不带销螺栓更容易发生缺陷, 因此本文的研究对象是带销螺栓。根据螺栓上开口销的状态, 将带销螺栓分为正常螺栓和缺陷螺栓, 其中缺陷螺栓是指开口销脱落或松动的螺栓。数据集中有 28 887 幅图像, 每幅图像分辨率约为 $5\,000 \times 3\,000$ 像素, 按照 8:1:1 的比例划分训练集、验证集和测试集。

2.2 环境设置和评价指标

本文模型使用 Python 3.7.12 和 PyTorch 1.8.1 实现, 并在配置为 NVIDIA GeForce GTX 2080Ti GPU

的本地服务器上训练和测试。采用 Faster R-CNN+FPN(feature pyramid network)(Lin 等, 2017a)作为基线模型,使用在 ImageNet(Deng 等, 2009)数据集上预训练的 ResNet-50(deep residual network)(He 等, 2016)作为特征提取网络。训练时设置 batch size 大小为 8,迭代 12 个 epoch,初始学习率为 0.02,并在第 8 个 epoch 和第 11 个 epoch 时衰减为原来的 1/10。

目标检测中两个常用的评价模型性能的指标是平均精度(average precision, AP)和平均精度均值(mean average precision, mAP)。Everingham 等人(2010)给出了 AP 的详细计算方法, mAP 是各类别 AP 的平均值。一般来说,某一类目标的 AP 值越高,模型对该类别的检测效果越好。

2.3 实验结果

为了验证本文模型的性能,将螺栓数据集在其他模型上进行了实验,包括两阶段模型 Cascade R-CNN(Cai 和 Vasconcelos, 2018)、Double Head R-CNN(Wu 等, 2020),单阶段模型 YOLOv5,基于 Transformer 的检测模型 Deformable DETR(deformable detection Transformer)(Zhu 等, 2020)和以 Swin Transformer(Liu 等, 2021)、ResNest(deep residual network with split-attention)(Zhang 等, 2022)作为骨干网络的检测模型。实验结果如表 1 所示,本文模型相比于其他模型在正常螺栓识别 AP、缺陷螺栓识别 AP 和 mAP 等 3 个指标上均有更好的表现。特别地,相比于基线模型,本文模型使缺陷螺栓识别 AP 提高了 12%。与两阶段模型 Double Head R-CNN、单阶段模型 YOLOv5、Deformable DETR 和以 ResNest

作为骨干网络的检测模型相比,本文模型使缺陷螺栓识别 AP 分别提高了 5%、3.8%、2.4% 和 1.3%。这说明本文模型针对螺栓类间相似和类内多样问题所做的改进是有效的。

表 1 本文模型与其他模型的对比
Table 1 Comparison between the model in this paper and other models

模型	骨干网络	AP		mAP
		正常螺栓	缺陷螺栓	
基线模型	ResNet-50	88.3	71.9	80.1
Faster R-CNN	Swin Transformer	89.2	82.1	85.7
Faster R-CNN	ResNest	89.1	82.6	85.9
Cascade R-CNN	ResNet-50	89.0	76.5	82.8
Double Head R-CNN	ResNet-50	89.1	78.9	84.0
YOLOv5	Focus+C3Net	88.7	80.1	84.4
Deformable DETR	Transformer	89.2	81.5	85.4
本文	ResNet-50	89.2	83.9	86.6

注:加粗字体表示各列最优结果。

P-R 曲线是以准确率(precision, P)为纵坐标,以召回率(recall, R)为横坐标绘制的,可用于衡量目标检测模型性能。P-R 曲线与横纵坐标轴围成的面积表示 AP 值。图 6 展示了基线模型和本文模型的 P-R 曲线。从图 6 中可以看出,本文模型在两个类别上的 P-R 曲线均高于基线模型,特别是在缺陷

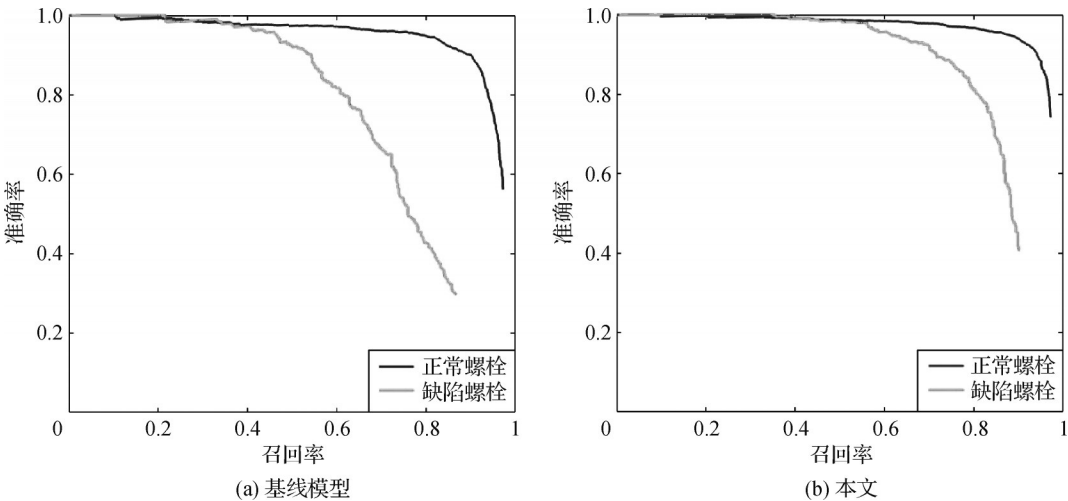


图 6 不同模型的 P-R 曲线对比
Fig. 6 Comparison of PR curves of different model ((a) baseline model; (b) ours)

螺栓的P-R曲线上。这表明在相同的召回率下,本文模型的准确率更高,检测结果更好。

图7展示了基线模型和本文模型检测的可视化结果,包括预测框、预测类别和置信度分数。对于图7中的第1幅图像,基线模型将铁架上的两个不带销螺栓误检为销丢失螺栓,而本文通过帮助模型建模不同类螺栓安装在不同位置的先验信息避免了这种误判;一个正常螺栓由于拍摄角度等的原因仅能看到开口销的头部,视觉分类特征不明显,基线模型将这个正常螺栓误判为缺陷螺栓,而本文模型则将其正确分类且获得了较高的置信度分数。图7第2幅图像中的两个销孔不明显的缺陷螺栓,基线模型虽然分类正确但置信分数较低,但本文模型则给出了较高的置信分数。

2.4 消融实验

为了验证各个方法的有效性,在螺栓数据集上对各个方法进行实验,结果如表2所示。由于基于K近邻的难例挖掘是在改进的费舍尔损失的基础上提出的,因此模型Ⅲ同时使用这两种方法。

1)由表2知,与使用均匀裁剪的基线模型相比,聚类裁剪使正常螺栓识别AP和缺陷螺栓识别AP分别提升了0.9%和7.2%,mAP提高了4.1%。这表明利用基于密度的空间聚类算法对原始航拍图像进行预裁剪,帮助模型聚焦关键区域,可以有效地提高检测模型对缺陷螺栓的识别精度。

在图3中,均匀裁剪获得了5个目标稀疏的子图,分别包含3、2、2、2、1个螺栓,且存在螺栓被割裂

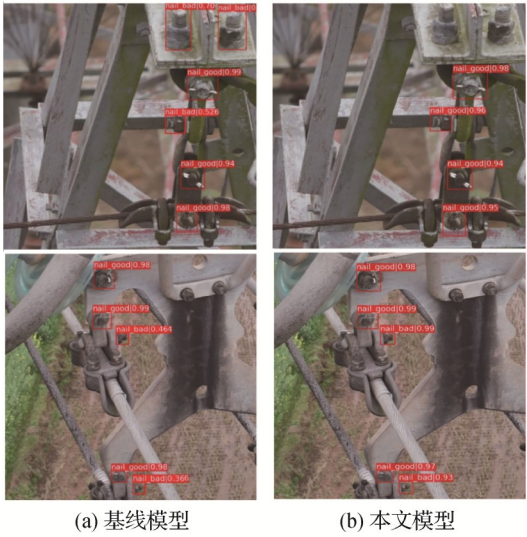


图7 检测结果可视化
Fig. 7 Visualization of detection results
((a) baseline model; (b) the proposed model)

的情况(红色椭圆圈出的螺栓不完整);而聚类裁剪仅生成了3个子区域,分别包含5、2和6个螺栓。均匀裁剪破坏了图3中原始巡检图像右边的三角挂板结构,将其分割到3幅子图中;而聚类裁剪则完整地保留了这一区域结构信息。

图8展示了使用这两种预裁剪方法处理训练集后获得的子图情况。均匀裁剪产生了86 199幅子图,其中约43%的子图仅含有1个带销螺栓;而聚类裁剪产生了59 676幅子图,相比均匀裁剪减少了约31%。这导致基线模型的训练时间比模型Ⅰ要长,造成计算资源的浪费。

表2 消融实验
Table 2 Ablation experiment

模型	聚类裁剪	费舍尔损失	基于K近邻的难例挖掘	AP/%		mAP/%
				正常螺栓	缺陷螺栓	
基线	×	×	×	88.3	71.9	80.1
模型Ⅰ	√	×	×	89.2	79.1	84.2
模型Ⅱ	×	√	×	89.0	78.9	83.9
模型Ⅲ	×	√	√	89.2	81.1	85.1
本文	√	√	√	89.2	83.9	86.6

注:加粗字体表示各列最优结果,“√”表示当前模型采用该方法,“×”表示当前模型未采用该方法。

通过上述分析可知,基于密度空间聚类的预裁剪算法不仅可以减少模型后续处理的数据量,节省计算资源,而且能够提取目标的上下文区域,帮助模

型聚焦关键信息,提高模型的识别能力。

2)由表2知,相比于基线模型,使用改进的费舍尔损失训练的模型使正常螺栓识别AP和缺陷螺栓

识别 AP 分别提高了 0.7% 和 7%, mAP 提高了 3.8%。费舍尔损失通过对特征施加约束, 使特征类内相聚、类间相离。这增大了不同类螺栓的特征差异度和同类螺栓的特征相似度, 更有助于螺栓分类。

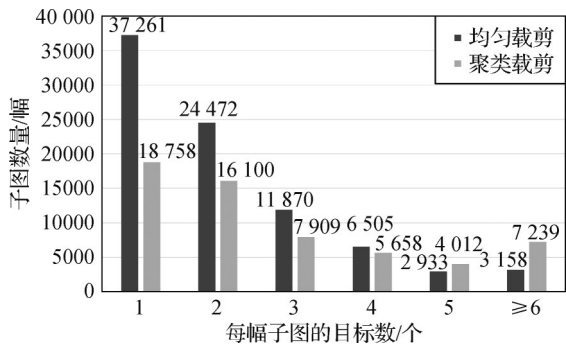


图 8 两种裁剪方法获得的子图对比
Fig. 8 Comparison of subgraphs obtained between two cropping methods

3) 由表 2 知, 基于 K 近邻的难例挖掘使模型 II 的性能进一步提升, 缺陷螺栓识别 AP 达到了 81.1%, 提高了 2.2%。由此可知, 通过 K 近邻学习引导模型关注难例可以提高不易辨识螺栓的识别效果。

近邻数 K 的取值会对模型性能产生影响。本文对 K 为 3~8 的取值进行了实验, 结果如表 3 所示。可以看出, 当 $K=5$ 时模型达到最佳效果; 而随着 K 值的增大, 模型效果开始下降。这说明选择合适的 K 值对基于 K 近邻的难例挖掘方法是有必要的, 若选取的 K 值过大, 这时无关样本也会影响 K 近邻概率。

4) 在统计分类问题中, 混淆矩阵是一种特定的表格布局, 用于可视化分类算法的性能。检测任务的混淆矩阵与分类任务的非常相似, 区别在于分类

表 3 近邻数 K 的取值对检测结果的影响

Table 3 The influence of the value of the nearest neighbor number K on the detection results

K	mAP/%
3	84.8
4	85.0
5	85.1
6	84.7
7	84.7
8	84.1

注: 加粗字体表示最优结果。

任务的对象是整幅图像, 而检测任务的对象是图像中的各个目标。因此, 为了能够绘制检测任务中混淆矩阵的正负例, 需要根据预测框和真实框的交并比 (intersection over union, IoU) 区分哪些预测结果是正确的, 哪些是错误的。

为了验证费舍尔损失和基于 K 近邻的难例挖掘方法对模型分类效果的提升, 本文绘制了基线模型和模型 III 检测结果的混淆矩阵, 结果如图 9 所示。在绘制过程中, IoU 阈值设置为 0.5, 且只统计置信度分数大于 0.3 的预测框。测试集中正常螺栓有 2 540 个, 缺陷螺栓有 556 个。与分类任务的混淆矩阵不同, 该混淆矩阵某一行的数量之和不代表该类目标的真实数量。因为模型的检测结果中存在对螺栓的重复检测, 即一个螺栓对应两个类别不同的预测框, 这导致非极大值抑制算法无法剔除掉任何一个检测框。

从图 9(a)(b) 的对比中可以看出, 模型 III 相比

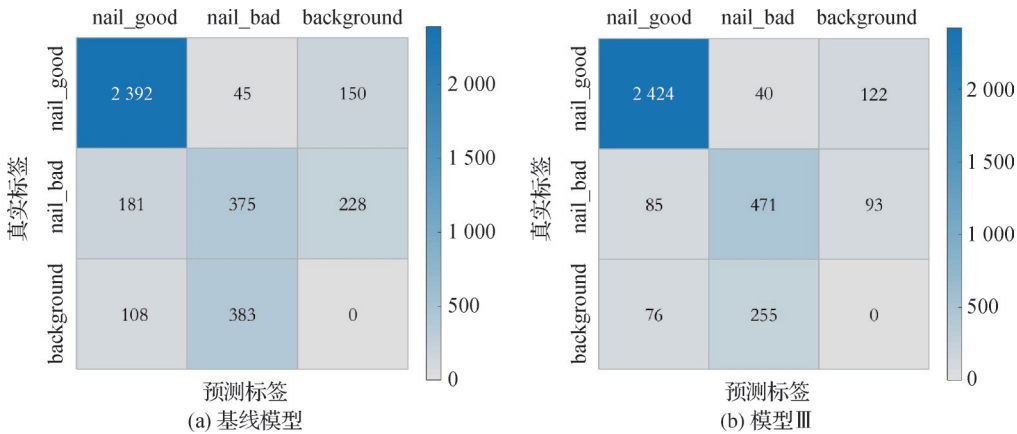


图 9 检测任务的混淆矩阵

Fig. 9 The confusion matrix of the detection task ((a) baseline model; (b) model III)

基线模型在正常螺栓和缺陷螺栓的召回率和准确率方面均有提升。基线模型在缺陷螺栓上的召回率和准确率分别为 67.4% 和 46.7%, 模型Ⅲ相比基线模型在缺陷螺栓上的召回率和准确率分别提高了 17.3% 和 14.7%。

3 结 论

本文工作为电力公司基于航拍图像进行螺栓故障巡检提供了一种有效的方法, 能够减少螺栓类间相似性和类内多样性对螺栓故障检测的干扰。主要贡献有以下几个方面:

1) 在数据预处理阶段, 利用 DBSCAN 算法对巡检图像进行预裁剪来获取感兴趣区域。预裁剪不仅可以加快模型训练速度, 帮助模型聚焦关键区域, 而且能够将不同类螺栓安装在不同位置的先验信息引入模型, 以此来辅助模型分类。

2) 为了学习到更有区分性的特征, 引入并改进费舍尔损失来约束样本特征的学习过程, 使得最终的样本特征类内分散小、类间分隔大。

3) 针对不容易分类的螺栓, 提出利用 K 近邻学习将样本特征与近邻特征的距离分布转化为概率分布, 并将其作为难易样本的指示以引导模型在训练时更多地关注难样本, 以提高难样本的识别效果。

本文方法仍存在一些问题, 如目前仅通过空间聚类算法将不同类螺栓安装在不同位置的先验信息引入模型, 但电力部件间存在的空间结构知识尚未充分挖掘利用。未来的工作将围绕这一课题展开, 探究部件间结构知识的提取方法和基于结构知识的网络建模方法。

参考文献 (References)

- Bochkovskiy A, Wang C Y and Liao H Y M. 2020. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. [2022-11-01]. <https://arxiv.org/pdf/2004.10934.pdf>
- Cai Z W and Vasconcelos N. 2018. Cascade R-CNN: delving into high quality object detection//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 6154-6162 [DOI: 10.1109/CVPR.2018.00644]
- Cheng G, Zhou P C and Han J W. 2016. RIFD-CNN: rotation-invariant and fisher discriminative convolutional neural networks for object detection//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 2884-2893 [DOI: 10.1109/CVPR.2016.315]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. Imagenet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]
- Ester M, Kriegel H P, Sander J and Xu X W. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise//Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining. Portland, USA: AAAI Press: 226-231. [DOI: 10.5555/3001460.3001507]
- Everingham M, Van Gool L, Williams C, Winn J, Zisserman A. 2010. The Pascal Visual Object Classes (VOC) challenge. International Journal of Computer Vision, 88 (2): 303-338 [DOI: 10.1007/S11263-009-0275-4]
- Feng M, Luo W, Yu L, Zhang P, Hao X L, Fan Q, Peng Q W, Zhang T B and Cao L L. 2018. A bolt detection method for pictures captured from an unmanned aerial vehicle in power transmission line inspection. Journal of Electric Power Science and Technology, 33(4): 135-140 (冯敏, 罗旺, 余磊, 张佩, 郝小龙, 樊强, 彭启伟, 张天兵, 曹玲玲. 2018. 适用于无人机巡检图像的输电线路螺栓检测方法. 电力科学与技术学报, 33(4): 135-140) [DOI: 10.3969/j.issn.1673-9140.2018.04.019]
- Girshick R. 2015. Fast R-CNN//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 1440-1448 [DOI: 10.1109/ICCV.2015.169]
- Girshick R, Donahue J, Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: 580-587 [DOI: 10.1109/CVPR.2014.81]
- 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]
- Jia M L, Chen B C, Wu Z X, Cardie C, Belongie S and Lim S N. 2021. Rethinking nearest neighbors for visual classification [EB/OL]. [2022-11-01]. <https://arxiv.org/pdf/2112.08459.pdf>
- Jiao R H, Liu Y Z, He H, Ma X H and Li Z Y. 2022. A deep learning model for small-size defective components detection in power transmission tower. IEEE Transactions on Power Delivery, 37 (4): 2551-2561 [DOI: 10.1109/TPWRD.2021.3112285]
- Jocher G, Stoken A and Borovec J. 2021. ultralytics/yolov5: v6.0 YOLOv5-x6 1280 models [EB/OL]. [2022-11-21]. <https://github.com/ultralytics/yolov5>
- Li X F, Liu H Y, Liu G H and Su H S. 2021. Transmission line pin defect detection based on deep learning. Power System Technology, 45(8): 2988-2995 (李雪峰, 刘海莹, 刘高华, 苏寒松. 2021. 基于深度学习的输电线路销钉缺陷检测. 电网技术, 45(8):

- 2988-2995) [DOI: 10.13335/j.1000-3673.pst.2020.1028]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017a. Feature pyramid networks for object detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 936-944 [DOI: 10.1109/CVPR.2017.106]
- Lin T Y, Goyal P, Girshick R, He K M and Dollar P. 2017b. Focal loss for dense object detection//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2999-3007 [DOI:10.1109/ICCV.2017.324]
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y and Berg AC. 2016. SSD: single shot multibox detector//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 21-37 [DOI: 10.1007/978-3-319-46448-0_2]
- Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, Lin S and Guo B N. 2021. Swin transformer: Hierarchical vision transformer using shifted windows//Proceedings of 2021 IEEE/CVF International Conference on Computer vision. Montreal, Canada: IEEE: 9992-10002 [DOI: 10.1109/ICCV48922.2021.00986]
- 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, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]
- Redmon J and Farhadi A. 2017. YOLO9000: better, faster, stronger//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 6517-6525 [DOI: 10.1109/CVPR.2017.690]
- Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement [EB/OL]. [2022-11-01]. <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]
- Toth J and Gilpin-Jackson A. 2010. Smart view for a smart grid — unmanned aerial vehicles for transmission lines//Proceedings of the 1st International Conference on Applied Robotics for the Power Industry. Montreal, Canada: IEEE: 1-6 [DOI: 10.1109/CARPI.2010.5624465]
- Van Der Maaten L and Hinton G. 2008. Visualizing data using t-SNE. Journal of Machine Learning Research, 9(11): 2579-2605
- Wang Y, Chao W L, Weinberger K Q and Van Der Maaten L. 2019. Simpleshot: revisiting nearest-neighbor classification for few-shot learning [EB/OL]. [2022-11-01]. <https://arxiv.org/pdf/1911.04623.pdf>
- Wu Y, Chen Y P, Yuan L, Liu Z C, Wang L J, Li H Z and Fu Y. 2020. Rethinking classification and localization for object detection//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 10186-10195 [DOI: 10.1109/CVPR42600.2020.01020]
- Xu D, Yu L, Chen T L and Wang J Q. 2017. Detection of catenary rotation binaural based on LBP-HOG feature. Journal of Railway Science and Engineering, 14(2): 370-378 (徐丹, 于龙, 陈唐龙, 王佳祺. 2017. 基于LBP-HOG特征的接触网旋转双耳区域识别. 铁道科学与工程学报, 14(2): 370-378) [DOI: 10.19713/j.cnki.43-1423/u.2017.02.024]
- Zeng Y B, Wang X H, Peng X G, Huang J L, Jian S C and Lu D. 2020. Research on risk modeling and forecasting method of transmission line defects. Power System Protection and Control, 48(10): 91-98 (曾勇斌, 王星华, 彭显刚, 黄景林, 简胜超, 鲁迪. 2020. 输电线路缺陷风险建模及其预测方法研究. 电力系统保护与控制, 48(10): 91-98) [DOI: 10.19783/j.cnki.pspc.190935]
- Zhang H, Wu C R, Zhang Z Y, Zhu Y, Lin H B, Zhang Z, Sun Y, He T, Mueller J, Manmatha R, Li M and Smola A. 2022. ResNeSt: split-attention networks//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans, USA: IEEE: 2735-2745 [DOI: 10.1109/CVPRW56347.2022.00309]
- Zhao W Q, Jia M Y, Zhai Y J and Zhao Z B. 2022a. Detection method for pin-losing-bolts in transmission lines based on improved YOLOv5s [J/OL]. [2022-11-01]. Journal of North China Electric Power University, 1-10 (赵文清, 贾梦颖, 翟永杰, 赵振兵. 2022. 基于改进YOLOv5s的输电线路螺栓缺陷检测方法[J/OL]. [2022-11-01]. 华北电力大学学报, 1-10 <http://kns.cnki.net/kcms/detail/13.1212.TM.20220920.0949.002.html>)
- Zhao W Q and Xu M F. 2022b. Bolt missing pins detection of transmission lines based on inter-level cross feature fusion. Journal of Image and Graphics, 27(11): 3222-3231 (赵文清, 徐敏夫. 2022. 基于隔级交叉特征融合的输电线路螺栓缺陷检测. 中国图象图形学报, 27(11): 3222-3231) [DOI: 10.11834/jig.210581]
- 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]
- Zhao Z B, Li Y X, Zhen Z, Zhai Y J, Zhang K and Zhao W Q. 2020. Typical fittings detection method with faster R-CNN combining KL divergence and shape constraints. High Voltage Engineering, 46(9): 3018-3026 (赵振兵, 李延旭, 甄珍, 翟永杰, 张珂, 赵

文清. 2020. 结合 KL 散度和形状约束的 Faster R-CNN 典型金具检测方法. 高电压技术, 46(9): 3018-3026 [DOI: 10.13336/j.1003-6520.hve.20200507023]

Zhao Z B, Qi H Y, Qi Y C, Zhang K, Zhai Y J and Zhao W Q. 2020. Detection method based on automatic visual shape clustering for pin-missing defect in transmission lines. IEEE Transactions on Instrumentation and Measurement, 69(9): 6080-6091 [DOI: 10.1109/TIM.2020.2969057]

Zhu X Z, Su W J, Lu L W, Li B, Wang X G and Dai J F. 2020. Deformable DETR: deformable transformers for end-to-end object detection [EB/OL]. [2022-11-01]. <https://arxiv.org/pdf/2010.04159.pdf>

作者简介

阎光伟,男,副教授,硕士生导师,主要研究方向为计算机图形/图像技术在电力系统中的应用。
E-mail: yan_guang_wei@126.com

焦润海,通信作者,男,教授,博士生导师,主要研究方向为能源数据(负荷、电量、新能源)预测与分析技术、图像目标检测技术。E-mail: runhaijiao@ncepu.edu.cn

周香君,女,硕士研究生,主要研究方向为人工智能在电力系统中的应用。E-mail: zhou_xiang_jun@163.com

何慧,女,讲师,硕士生导师,主要研究方向为目标识别和新能源预测。E-mail: huihe@ncepu.edu.cn