



嵌入视觉关系掩码的多类别金具检测方法

王巍, 杨耀权, 王乾铭, 翟永杰, 赵振兵

引用本文:

王巍,杨耀权,王乾铭,翟永杰,赵振兵. 嵌入视觉关系掩码的多类别金具检测方法[J]. 智能系统学报, 2023, 18(3): 440–449.

WANG Wei,YANG Yaoquan,WANG Qianming,ZHAI Yongjie,ZHAO Zhenbing. A multi-category fitting detection method with embedded visual relation masks[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(3): 440–449.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202202021>

您可能感兴趣的其他文章

基于图知识推理的输电线路缺销螺栓识别方法

Pin-missing bolts recognition method for transmission lines based on graph knowledge reasoning
智能系统学报. 2023, 18(2): 372–380 <https://dx.doi.org/10.11992/tis.202205004>

基于遮挡关系推理的输电线路图像金具检测

Fittings detection in transmission line images with occlusion relation inference
智能系统学报. 2022, 17(6): 1154–1162 <https://dx.doi.org/10.11992/tis.202108036>

融合外部知识的输电线路多金具解耦检测方法

A multi-fitting decoupling detection method for transmission lines based on external knowledge
智能系统学报. 2022, 17(5): 980–989 <https://dx.doi.org/10.11992/tis.202107026>

融合共现推理的Faster R-CNN输电线路金具检测

Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection
智能系统学报. 2021, 16(2): 237–246 <https://dx.doi.org/10.11992/tis.202012023>

嵌入遮挡关系模块的SSD模型的输电线路图像金具检测

Fittings detection in transmission line images with SSD model embedded occlusion relation module
智能系统学报. 2020, 15(4): 656–662 <https://dx.doi.org/10.11992/tis.202001008>

输电线路巡检图像智能诊断系统

Intelligent diagnosis system for patrol check images of power transmission lines
智能系统学报. 2016, 11(1): 70–77 <https://dx.doi.org/10.11992/tis.201503043>

DOI: 10.11992/tis.202202021

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20230327.1809.006.html>

嵌入视觉关系掩码的多类别金具检测方法

王巍¹, 杨耀权¹, 王乾铭¹, 翟永杰¹, 赵振兵²

(1. 华北电力大学 自动化系, 河北 保定 071003; 2. 华北电力大学 电子与通信工程系, 河北 保定 071003)

摘要: 在对输电线路金具进行检测的过程中, 由于受到复杂背景的影响, 一些互相遮挡或者特征不明显的金具会隐匿在复杂环境难以精确检测。针对这一问题, 提出了基于视觉关系掩码的多类别金具检测模型, 通过挖掘和提取输电线路金具之间包含空间信息的视觉关系先验知识, 构建视觉关系掩码和视觉关系检测网络, 并将先验知识作为辅助信息融入视觉关系模块中, 最终实现多类别金具的精确定位与识别。对具有 14 类金具的数据集进行多种定性和定量实验, 结果表明, 改进后的模型平均检测精度能提高到 76.25%, 检测效果也优于其他先进目标检测模型。

关键词: 目标检测; 输电线路; 金具; 深度学习; 视觉关系; 先验知识; 空间信息; 辅助信息

中图分类号: TP183 **文献标志码:** A **文章编号:** 1673-4785(2023)03-0440-10

中文引用格式: 王巍, 杨耀权, 王乾铭, 等. 嵌入视觉关系掩码的多类别金具检测方法 [J]. 智能系统学报, 2023, 18(3): 440-449.

英文引用格式: WANG Wei, YANG Yaoquan, WANG Qianming, et al. A multi-category fitting detection method with embedded visual relation masks[J]. CAAI transactions on intelligent systems, 2023, 18(3): 440-449.

A multi-category fitting detection method with embedded visual relation masks

WANG Wei¹, YANG Yaoquan¹, WANG Qianming¹, ZHAI Yongjie¹, ZHAO Zhenbing²

(1. Department of Automation, North China Electric Power University, Baoding 071003, China; 2. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: In the process of detecting power transmission line fittings, due to the influence of complex backgrounds, some fittings with mutual obscuration or inconspicuous features will be hidden in complex environments and it is hard to detect them accurately. In response to this problem, this paper proposes a multi-category hardware detection model based on visual relationship masks. By mining and extracting the prior knowledge of visual relationships between power transmission line fittings that contain spatial information, the visual relationship mask (VRM), and visual relationship detection network (VRDN) are constructed, and the prior knowledge is integrated into the visual relationship module as auxiliary information, realizing the precise positioning and recognition of multi-category fittings. A variety of qualitative and quantitative experiments have been performed on a data set with 14 types of hardware. The results show that the average detection accuracy of the improved model can be increased to 76.25%, and the detection effect is better than other advanced target detection models.

Keywords: target detection; transmission line; fittings; deep learning; visual relationship; prior knowledge; spatial information; auxiliary information

近年来, 随着电网高速发展与全面覆盖, 长距

离输电线路日渐增多, 输电线路作为电力传输中的核心系统, 其稳定运行对电网的安全有着至关重要的影响^[1-2]。目前, 输电线路金具的缺陷日益增多, 运行维护需求量增长与巡检人员数量之间的矛盾逐渐凸显, 对自动化与智能化检测技术的

收稿日期: 2022-02-24. 网络出版日期: 2023-03-28.

基金项目: 国家自然科学基金项目 (U21A20486, 61871182); 河北省自然科学基金项目 (F2021502008).

通信作者: 翟永杰. E-mail: zhaiyongjie@ncepu.edu.cn.

需求越来越迫切,而如何精确定位输电线路金具仍是亟待解决的关键问题^[3-4]。输电线路金具主要起到支持、固定、防护和接续线路各个部件以及维持整个线路稳定的作用,如间隔棒、防震锤、均压环等^[5]。当前输电线路金具检测主要依靠无人机航拍获得的金具图像,以计算机视觉和图像处理技术为支撑,实现对金具目标的精确定位和识别^[6-7]。无人机采集的海量图像,人工逐张确认有无缺陷并不现实,而准确定位金具给判断有无缺陷带来便利,因此,如何对金具精确定位和识别具有极高的研究意义。

输电线路金具检测大致可以分为两类方法,第1种是基于传统图像处理算法,依靠人工设计特征提取器来定位金具,并结合机器学习算法进行分类识别^[8];第2种是依靠深度学习模型,自主提取多层次多角度的特征,依托强大的学习能力和表达能力对目标进行定位,由于其鲁棒性高、泛化性强,相比与传统图像处理算法更适用于复杂多变的现实环境^[9]。汤踊等^[10]采用 Faster R-CNN 网络结构,研究了不同网络主干对检测精度和检测速度的影响,并探究了网络第一层的卷积核大小对召回率的影响,发现较大的卷积核一定程度上能够提高检测精度。高强等^[11]基于 YOLOv3 框架,考虑目标间的遮挡问题,提出可聚合度的概念,将属于同一整体的目标检测框进行聚合,可以通过目标部分元素确定改目标的存在。赵振兵等^[12]针对目标互相遮挡的问题,设计一种带有注意力机制的遮挡关系模块,嵌入到单阶段多框检测器(single shot multibox detector, SSD)目标检测模型中,较原模型提升了4.46%。威银城等^[13]对金具图像进行不同角度的旋转和固定尺寸的自适应裁剪进行样本扩充,并使用改进的 IoU(intersection over union)替换原损失计算方式,一定程度上提高了检测精度。Wan 等^[14]在基于区域的全卷积网络基础上引入可变形卷积和挤压激励模块,有效提取输电线路金具之间细粒度特征和融合上下文信息。赵振兵等^[15]针对金具与金具之间、金具与背景之间相互干扰的问题,提出了一种结合 KL(kullback-leibler)散度^[16]和形状约束的 Faster R-CNN 金具检测方法,使用 KL 散度替代原有的 SmoothL1 损失函数^[17],同时将不同目标的形状特征作为约束加入损失函数中对模型训练进行优化,以提高模型检测精度。

然而上述的研究工作并未充分考虑金具之间的视觉关系,只是针对目标框重叠的情况做分析,金具间的视觉关系没有得以充分利用,并且

在金具图像背景复杂和重叠框数量较多的情况下,单单考虑目标框重叠信息是不够的,一些重叠区域较小的目标容易“淹没”在背景复杂的图像中,不容易被模型发现。针对以上问题,本文通过挖掘输电线路金具之间的空间信息和视觉关系,引入一种视觉关系掩码(visual relationship mask, VRM)的概念,并构建了一种视觉关系检测网络(visual relationship detection network, VRDN),融合上下文信息来学习目标框之间的视觉关系,辅助原模型对目标框的定位效果,从而提高检测精度,为输电线路的金具智能化巡检提供了新思路。

1 总体框架

本文总体结构如图1所示,首先通过分析输电线路金具航拍图像之间的空间信息和视觉关系,提取金具之间视觉关系先验知识,然后将金具航拍图像作为输入,通过卷积神经网络获取高层特征,在经区域提议网络(regional proposal network, RPN)网络^[18]进行感兴趣区域框筛选,并进行感兴趣区域(region of interest, ROI)池化将每个区域框整合为同一尺度的区域向量,在经过全连接层进行预测框坐标和类别预测。从所有的预测信息中筛选128个预测框进行目标的分类和定位损失的计算;同时从128个预测框中32个正样本中筛选 N 个预测框,通过掩码生成算法(mask generation algorithm, MGA)构建视觉关系掩码,并生成掩码特征,再输送到VRDN中进行掩码特征的高层语义信息提取,最后将视觉关系先验知识融入到网络输出端进行视觉关系损失计算。最终,由视觉关系检测器辅助分类、定位检测器对多类别金具进行精确检测。

2 视觉关系先验知识提取

2.1 视觉关系规则制定

在架空输电线路的设计和建造过程中,为了达到标准的电气要求和安装规范,输电线路金具之间往往具有较为固定的组合规则和结构^[19],区别于场景复杂、多类目标共有的上下文信息较少的单一目标检测模型,融合金具之间的视觉关系先验知识有助于更好地实现金具检测。由于无人机采集金具图像时相机视线相对平行于地面,并且金具的组装连接呈现出结构化、固定化的特点,所以金具各类别呈现在图像上会出现强烈的空间视觉关系,如纵向视觉关系和横向视觉关系。如图2,均压环、屏蔽环通常存在于整个组合结构的

上端或者下端,和悬垂线夹、联板、重锤等存在明显纵向视觉关系;成对的联板和重锤出现的时候会存在横向视觉关系,防震锤作为减轻输电线路

受微风影响的防护金具,通常出现在整个组合结构的周围,并构成横向视觉关系,调整板和 U 型挂环往往同时出现,并存在横向视觉关系。

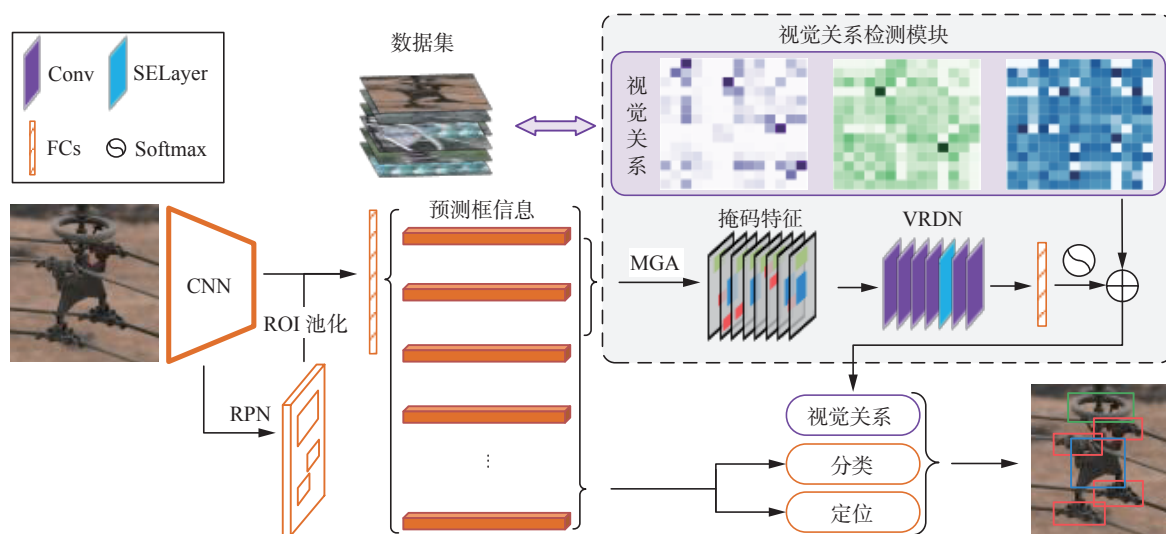
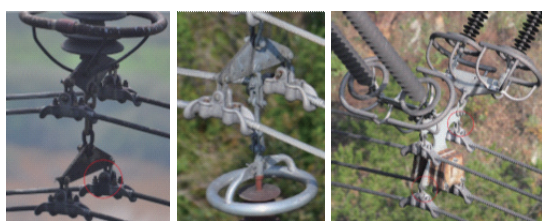
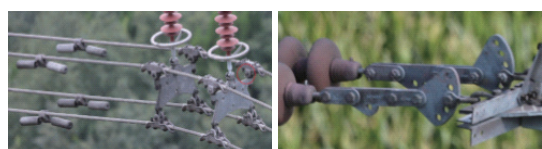


图 1 整体框架图

Fig. 1 Overall framework diagram



(a) 纵向视觉关系



(b) 横向视觉关系

图 2 金具的 2 种视觉关系

Fig. 2 Two visual relationships of hardware

除了上述 2 种视觉关系外,金具数据集中标注框之间的视觉重叠关系不可忽略,标注框之间的重叠区域影响模型对单个目标的识别度,尤其出现目标尺寸较小重叠区域较大的情况(如图 3a),而且有部分重叠区域是单个目标的关键特征部分(如图 3b),由于部分金具之间的固定组合结构,金具之间毗邻连接,标注框同样会出现重叠部分(如图 3c),这样的视觉重叠关系可以帮助本文判断金具之间的连接关系,提高检测精度。

基于以上 3 种视觉关系,本文从数据集标注框的相对位置以及相对空间角度出发,制定了提取 3 种视觉关系的具体规则。

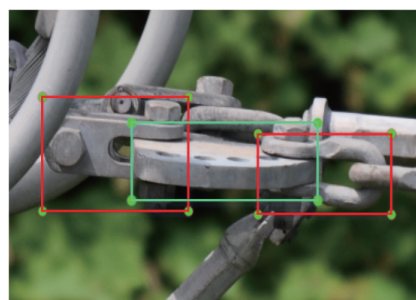
对于一个关系对 (S, O) , 本文使用 b_s 和 b_o 来

表示对应目标边界框归一化后的左上角坐标和右下角坐标。其中

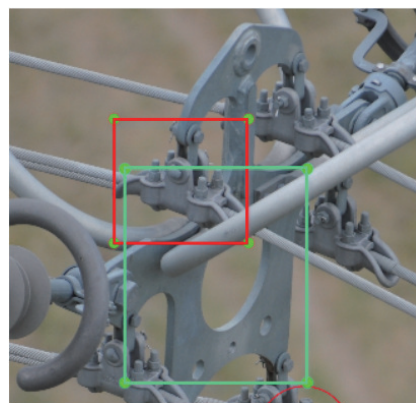
$$\begin{cases} b_o = (x_{1o}, y_{1o}, x_{2o}, y_{2o}) \\ b_s = (x_{1s}, y_{1s}, x_{2s}, y_{2s}) \end{cases} \quad (1)$$

1) 视觉重叠关系

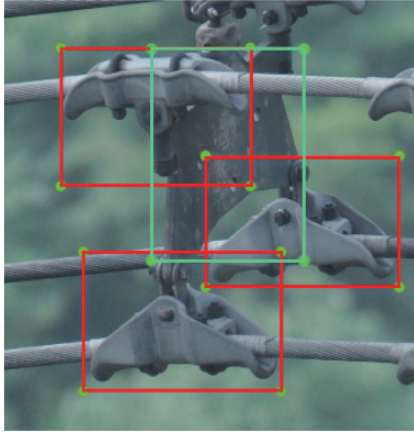
$$\nabla = [\min(x_{2i}) - \max(x_{1i})]_+ \times [\min(y_{2i}) - \max(y_{1i})]_+ \quad (2)$$



(a) 目标尺寸小重叠区域较大



(b) 重贴区域为关键特征



(c) 固定组合结构且互相重叠

图 3 视觉重叠关系

Fig. 3 Visual overlap relationship

使用式(2)计算 2 个对象框是否存在重叠区域, 其中 $i \in \{s, o\}$, $[\cdot]_+ = \max(0, \cdot)$ 表示仅保留输入为正数的部分, ∇ 代表重叠值, 如果计算重叠值大于 0, 那么就认为关系对 (S, O) 之间存在视觉重叠关系。

2) 纵向视觉关系和横向视觉关系

将关系对 (S, O) 的边界框坐标进行相应的转化得到边界框的中心坐标 $b'_s = (x_s, y_s)$ 和 $b'_o = (x_o, y_o)$, 样, 2 个坐标中心对 (b'_s, b'_o) 在二维平面可以确定一个向量 \mathbf{l}_{so} , 设向右为正方向水平向量为 \mathbf{l} , 计算两向量的夹角 θ , 那么空间关系 R_s 可以为

$$R_s = \begin{cases} 1, & 45^\circ < \theta \leq 135^\circ \\ 0, & \text{其他} \end{cases} \quad (3)$$

如果 R_s 为 1, 认为是纵向视觉关系, 否则是横向视觉关系。图 4 在二维平面中形象显示了这两种空间关系。以图像中心的重锤为观察对象, 黄色阴影部分为横向视觉关系, 蓝色阴影部分为纵向视觉关系, 2 个黑色圆弧夹角表示 θ 的大小, 红色框内的均压环和重锤为纵向视觉关系, 绿色框内的联板和重锤为横向视觉关系。

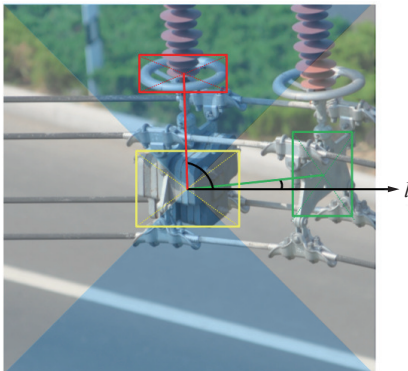


图 4 空间关系演示

Fig. 4 Demonstration of spatial relationship

2.2 先验知识提取

根据既定规则, 本文通过数据驱动的方式构建视觉关系概率矩阵, 即通过挖掘标签在数据集中的视觉关系来定义标签之间的相关性。首先, 假设单张图片中含有的标注信息为

$$I = \{(c_1, b_1), (c_2, b_2), \dots, (c_k, b_k)\} \quad (4)$$

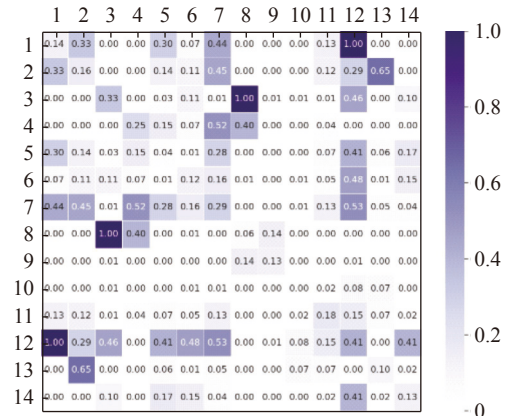
式中: k 为标注框个数, c 表示类别信息, b 表示位置信息, 然后从 k 个标注框中无重复的挑选 2 个标注框 $I_i = (c_i, b_i)$ 和 $I_j = (c_j, b_j)$ 做关系构建, 当满足任意一种关系时, 两个标注框类别相对应的关系计数加 1, 即

$$\Delta^R(c_i, c_j)^n = \Delta^R(c_i, c_j)^{n-1} + 1 \quad (5)$$

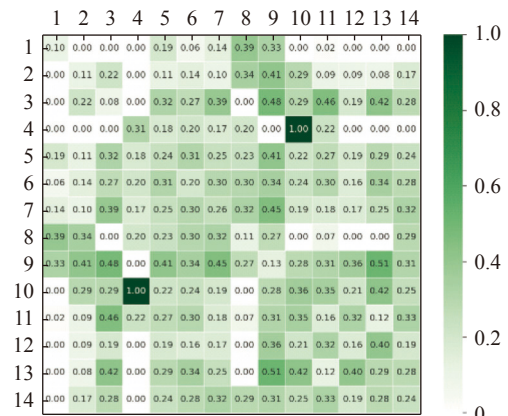
式中: $i \leq j \leq k$, $\Delta^R(c_i, c_j)$ 表示 (c_i, c_j) 属于某种关系。基于 2.1 节讨论, 本文优先考虑视觉重叠关系, 如果是视觉重叠关系, 则忽略其余 2 种关系, 最终统计单张图像所有的关系组合应为 C_k^2 个。统计数据集中所有关系组合, 按 3 种关系归一化概率和为 1 的方式进行矩阵归一化, 即

$$\sum P(\Delta^R(c_i, c_j)) = 1 \quad (6)$$

对于本文数据集, 共有 14 种类别, 3 种视觉关系, 对应的关系组合为 $14^2 \times 3 = 588$ 种, 如图 5, 将每种关系的概率得分以热力图的方式呈现, 从左到右依次是重叠、横向、纵向视觉关系, 热力图中横纵坐标序号对应的类别名称如表 1 所示。



(a) 重叠视觉关系



(b) 横向视觉关系

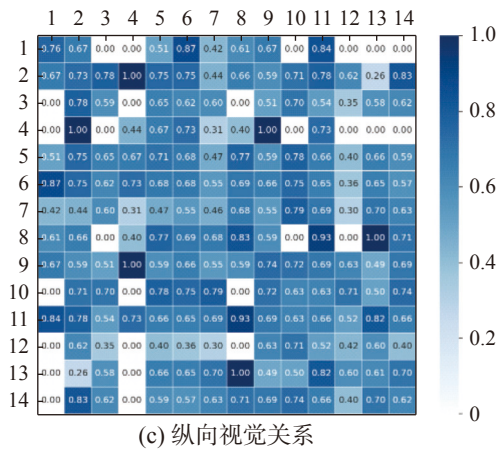


图 5 3 种视觉关系的概率矩阵

Fig. 5 Probability matrix of 3 kinds of visual relations

表 1 序号和类别名称对应关系

Table 1 Correspondence between serial number and category name

序号	类别名称	序号	类别名称
1	预绞式悬垂线夹	8	并沟线夹
2	提包式悬垂线夹	9	防震锤
3	压缩型耐张线夹	10	间隔棒
4	楔型耐张线夹	11	均压环
5	挂板	12	屏蔽环
6	U型挂环	13	重锤
7	联板	14	调整板

3 视觉关系表达

3.1 视觉关系掩码

为了更好地融合视觉关系知识,本文引入一种 VRM 的概念,下面本文将详细介绍 MGA。

给定 2 个归一化后的目标框的位置信息为 $b_s = (x_{1s}, y_{1s}, x_{2s}, y_{2s})$ 和 $b_o = (x_{1o}, y_{1o}, x_{2o}, y_{2o})$ 。

1) 初始化一个 64×64 像素大小的全零矩阵,作为背景;

2) 把位置信息映射到背景上,得到新的位置信息 M_{bs} 和 M_{bo} ;

3) 将新的边界框信息 M_{bs} 和 M_{bo} 内部元素置为 1, 得到新的掩码 M ;

4) 最终的掩码 M 为视觉关系掩码。

假设一张图像中有 k 个预测框,则会出现 C_k^2 个视觉关系掩码 M , C_k^2 个视觉关系掩码 M 共同决定了所有对象之间的视觉关系,将 C_k^2 个掩码拼接在一起,用 $[M]$ 表示,称做掩码特征。掩码特征能够反应多个对象之间相互作用的视觉关系,而视觉特征描述的是单个对象的高层语义信息,因

此本文认为视觉关系掩码和视觉特征可以共同促进对图像的整体理解。

3.2 视觉关系检测

为了有效地利用掩码特征 $[M]$, 本文构建了一种 VRDN。在 VRDN 之前,输入图像经过卷积神经网络提取高层语义信息,经过 RPN 和 ROI 池化得到固定尺度的特征图,在经过全连接层得到最终的预测框信息,在这些预测框信息中,挑选 128 个预测框进行分类损失 \mathcal{L}_c 和定位损失 \mathcal{L}_l 计算,其中正负样本数量按 1:3 划分。

为了和模型的预测阶段有效的联系起来,本文采用和预测阶段相同的操作,取这些正样本中置信度大于 0.05 的预测框,并进行阈值为 0.5 的非极大值抑制,在过滤完的预测框中按置信度排序,挑选前 N 个预测框构建掩码特征 $[M]$, 然后送入视觉关系检测网络中。对于 N 的选取,本文统计了数据集中每张图像标注的目标框个数,并以横轴为每张图片上目标框数量,纵轴为对应的图像数量做直方图,如图 6 是目标框数量直方图,并选取 N 为 $\{8, 16, 32\}$ 进行实验对比,实验结果将会在 4.1 节给出。

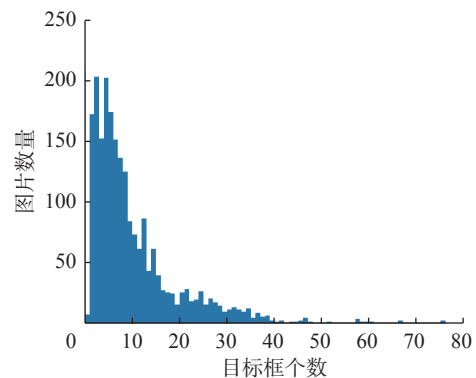


图 6 目标框个数直方图

Fig. 6 Histogram of the number of target frames

VRDN 由 6 个卷积层、一个注意力模块和全连接层组成。由于较大的卷积核尺寸在增大感受野的同时能获得图像更多的全局特征,为了提高网络的表征能力,本文将前 4 个卷积层的卷积核尺寸大小设置为 5; 后 2 个卷积层卷积核尺寸大小设置为 1, 主要起到降维的作用; 全连接层的最后一个节点个数设置为 3, 对应 3 种视觉关系。为了让网络能够关注局部特征信息,受 Hu 等^[20] 的启发,本文引入了一个注意力模块 (squeeze and excitation layer, SELayer), 该模块能够让网络重新校准学习特征,有选择的强调有用信息,抑制无用信息,这对于学习图像中局部目标框之间的相对空间位置是非常有用的。

通过融入视觉关系先验知识,本文设计了一个损失函数 \mathcal{L}_R 用于拉近对象之间的视觉关系与真实世界的距离。函数设计为

$$\mathcal{L}_R = \min \left(\sum_{k=1}^N \mathcal{L}_M^k \right) \quad (7)$$

式中: N 为 $[M]$ 中 M 的个数, \mathcal{L}_M^k 表示第 k 个 M 的计算损失,每一个 M 的计算损失 \mathcal{L}_M 为

$$\mathcal{L}_M = \psi \{ \phi [\phi [O^R(c_i, c_j)] + P(\Delta^R(c_i, c_j))], G^R(c_i, c_j) \} \quad (8)$$

其中: $\psi\{\cdot, \cdot\}$ 表示 $\{\cdot, \cdot\}$ 做交叉熵损失, $\phi[\cdot]$ 表示经 softmax 归一化处理, $O^R(c_i, c_j)$ 表示 M 经过 VRDN 后的输出, $P(\Delta^R(c_i, c_j))$ 表示 M 中 2 个预测框的先验概率, $G^R(c_i, c_j)$ 表示两目标类别的真实视觉关系。通过视觉先验知识进行辅助校正 VRDN 的预测输出,增强对应类别的预测得分,最大化的促进对象之间不同关系的共生。

最终,将视觉关系损失 \mathcal{L}_R 、分类损失 \mathcal{L}_c 和定位损失 \mathcal{L}_l 累加,作为金具检测模型的总体损失,3 种损失在反向传播的过程中共享主干网络和 RPN 网络参数,通过外加的 VRDN 提高模型检测精度,此外,VRDN 将训练阶段和预测阶段有效地联系起来,提高模型的准确率。

4 实验结果及分析

本文的实验数据集为 5760 像素×3840 像素和 4288 像素×2848 像素的高分辨率图像,选取 14 类金具作为实验对象,共包含 20 804 个实例对象,其中训练集和验证集按 4:1 划分。本文所述模型采用

NVIDIA 1080Ti 专业加速卡进行训练与测试;采用的操作系统为 Ubuntu16.04.6 LTS,利用 CUDA10.0 加速训练;使用的计算机语言为 Python3.6,网络开发框架为 Pytorch。在训练阶段,将批处理大小设为 2,使用了随机梯度下降^[21](SGD)算法,初始学习率设为 0.004,每经过 3 个 epoch,学习率下降为原来的 33%,最多迭代训练 16 个 epoch。

采用目前目标检测模型中常用的评价指标平均精度均值(mean average precision, mAP)对模型进行评估,通过度量目标检测框与真值框间的 IoU 计算各类目标平均精度(average precision, AP),以所有类目标的 AP 平均值(mAP)作为目标检测模型的最终评价指标。由于金具目标尺度的多样性,本文以 Faster R-CNN^[22]加特征金字塔结构^[23](feature pyramid network, FPN)进行多尺度检测,并作为基础框架加以改进,以此来验证视觉关系检测模块的有效性。

4.1 消融实验

为了提高基线模型的检测精度,首先使用在 COCO^[24]数据集上预训练的 Faster R-CNN+FPN 模型参数进行训练,将训练完毕后的模型再次作为预训练权重进行加载,以此来提升基线模型的检测精度。为保证视觉先验知识的有效性,训练阶段并没有使用任何图像增强技术,因为一些图像增强技术(如:旋转、平移、翻转等)会破坏金具之间的视觉关系。表 2 为 14 类金具的 AP₅₀ 测试结果对比,表 3 为 N 为不同值时的金具的 mAP₅₀ 以及 mAP_{50:95} 的测试结果对比。

表 2 14 类金具测试结果
Table 2 Test results of 14 types of fittings

模型	预绞式悬垂线夹	提包式悬垂线夹	并沟线夹	防震锤	U型挂环	挂板	联板
Baseline _{coco}	70.59	90.98	32.20	88.16	53.11	62.25	72.00
Baseline _{ours}	80.91	91.94	39.27	87.05	55.77	65.40	75.99
Ours-8	77.80	93.37	43.06	87.15	56.45	66.36	77.51
Ours-16	78.58	92.63	46.21	86.85	60.11	68.33	76.66
Ours-32	77.29	92.40	40.65	87.50	58.88	66.48	76.20
模型	压缩型耐张线夹	楔型耐张线夹	间隔棒	均压环	屏蔽环	重锤	调整板
Baseline _{coco}	52.97	32.48	82.02	89.25	68.73	98.33	78.66
Baseline _{ours}	52.51	54.79	87.78	88.62	67.24	99.57	79.42
Ours-8	57.25	66.27	86.33	90.53	70.28	99.04	80.92
Ours-16	59.71	69.99	88.67	90.52	69.41	97.52	82.44
Ours-32	58.98	74.57	85.16	89.97	68.31	99.19	82.77

表 3 不同 N 值的检测结果
Table 3 Detection results for different N values %

模型	mAP ₅₀	mAP _{50:95}
Baseline _{coco}	69.41	37.10
Baseline _{ours}	73.30	40.96
本文-8	75.17	41.35
本文-16	76.25	42.50
本文-32	75.60	42.26

其中, Baseline_{coco}表示使用 COCO 的预训练权重, Baseline_{ours}表示加载使用 COCO 权重训练一次的预训练权重, Ours- K 中的 K 代表计算掩码特征时筛选前 K 个正样本。从表 2 可以看出, 本文模型对于绝大多数金具的检测效果都有一定的提升, 其中对于楔型耐张线夹、压缩型耐张线夹、并沟线夹、U 型挂环以及挂板的检测效果有显著的提升。由于并沟线夹目标体积小, 传统模型没有考虑视觉关系, 并沟线夹容易“淹没”在复杂背景中, 本文考虑并沟线夹和其他金具的视觉关系, 提升了对并沟线夹的检测; U 型挂环和挂板由于目标体积小, 目标框重叠部分区域较大, 容易重复学习相同特征, 从而造成传统模型的误判; 楔型耐张线夹和压缩型耐张线夹形状及纹理特征和导线相似, 且楔型耐张线夹在数据集中出现次数较少, 传统模型没有得到充分学习。本文的方法充分考虑了数据集中金具之间的视觉关系, 较传统模型有了较大的提升, 从楔型耐张线夹案例还可以看出本文的模型对样本数量分布不均有一定

的适应能力。

从表 3 可以看出, 将在 COCO 数据集上将训练完毕后的模型参数再次作为新的预训练权重能有效提高金具的检测精度, 给一些检测精度较低的模型提供了新的方法。加入视觉关系检测模块后能够提高模型的检测准确率, 与基线模型相比, mAP₅₀ 和 mAP_{50:95} 分别提高了 2.95%、1.54%; N 的大小一定程度上决定了 VRDN 的效果, N 过大容易包含过多的负样本, 干扰 VRDN 对视觉关系的判断; N 过小不能包含较多的正样本, 很多预测正确的目标框没有得以利用, VRDN 的性能不能得以体现。其中 N 为 16 时模型的检测精度最高, 这说明 N 值过大或者过小都会影响检测精度。

4.2 模型检测结果可视化对比实验

为了定性地分析本文模型的检测效果, 图 7 显示了本文所述模型与基线模型的分析对比, 第 1 行为基线检测结果, 第 2 行为改进后模型检测结果。由于楔型耐张线夹纹理特征较其他金具不明显, 在图 7(a)中基线模型出现了漏判的情况, 在图 7(b)中出现了误判的情况, 考虑金具之间的视觉关系后, 使用本文算法进行检测后提高了楔型耐张线夹的检测置信度和检测精度; 图 7(c)显示的是对具有遮挡关系的小目标的检测对比, 基线模型出现了误判的情况, 错将调整板识别为挂板, 本文算法准确识别出了各类金具; 图 7(d)中出现了联板多检测的情况, 说明基线模型在检测的过程中对金具目标定位不准, 本文算法在考虑视觉关系后, 能优化基线模型对目标的定位效果。

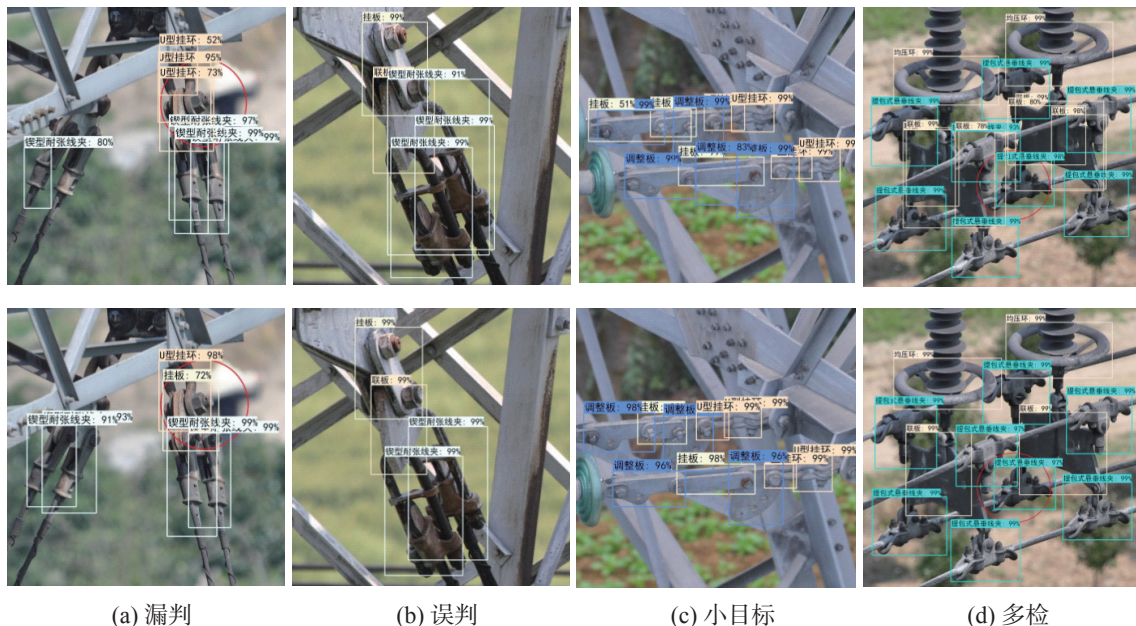


图 7 金具检测定性结果对比

Fig. 7 Comparison of qualitative results of hardware detection

4.3 与先进模型对比实验

为了验证本文模型的先进性, 本文在金具数据集上实现了多种目标检测模型, 分别与 SSD^[25]、YOLOv3 以及 YOLOv3-SPP^[26] 进行比较。

表 4 显示了所有模型 mAP_{50} 、 $mAP_{50:95}$ 以及部分金具 AP_{50} 的检测结果的检测结果, 为了保证实验的可靠性, 在与其他模型做对比的过程中, 同样没有采用数据增强技术。

表 4 所有模型的检测结果对比
Table 4 Comparison of test results of all models

模型	图像尺寸/像素	金具类别/%					$mAP_{50}/\%$	$mAP_{50:95}/\%$
		楔型耐张线夹	挂板	并沟线夹	U型挂环	间隔棒		
Baseline _{coco}	608×608	32.48	62.25	32.20	53.11	82.02	69.41	37.10
Baseline _{ours}	600×800	54.79	65.40	39.27	55.77	87.78	73.30	40.96
SSD	600×600	56.20	61.14	43.88	59.80	71.49	72.33	40.74
YOLOv3	608×608	59.55	55.87	29.15	50.08	80.04	64.96	35.86
YOLOv3-SPP	608×608	61.42	65.93	35.53	56.46	83.80	72.55	41.17
本文-16	600×800	69.99	68.33	46.21	60.11	88.67	76.25	42.50

从表 4 中可以看出, 本文模型相比于现阶段主流模型的 mAP_{50} 指标有较大的提升, 在部分金具检测结果上同样有较大提升, 尤其对密集遮挡小目标 U 型挂环、刮板之类的金具提升较大, 这说明加入视觉关系检测模块同样能够应对目标的密集遮挡问题; 由于数据集中楔型耐张线夹、并沟线夹目标特征不明显, 其他的主流模型没有充分考虑到视觉关系, 本文加入视觉关系检测模块后, 能有效提高检测精度, 一定程度上说明了本文算法的泛化能力优于其他主流模型。

5 结束语

本文提出了一种基于视觉关系掩码的多类别金具检测模型, 实现了金具的精确定位与识别, 通过对比实验分析, 得到如下结论:

1) 针对输电线路金具图像背景复杂和目标特征不明显及相互重叠的问题, 本文通过挖掘金具之间的视觉关系先验知识, 将其作为辅助信息融入视觉关系模块中, 实现了多类别金具的精确识别。

2) 在定性对比实验中, 针对基线模型和改进模型的检测效果进行可视化对比, 改进模型对小目标漏判、错判、多判以及互相遮挡问题均有改善。

3) 在定量对比实验中, 本文改进后的模型相较于基线模型 mAP_{50} 指标提高了 2.95%, 较其他先进模型 mAP_{50} 和 $mAP_{50:95}$ 指标均有一定的提升, 并在部分金具检测效果上提升明显。

本文使用金具之间视觉关系和先验知识来改

善金具的检测效果, 为输电线路设备部件智能巡检技术提供了新的思路。未来工作拟深入研究视觉关系知识与模型的融合, 使用知识推理的方法使视觉关系知识得到更有效的表达, 从而实现更准确的多类别金具检测。

参考文献:

- [1] 王宾, 倪江, 王海港, 等. 输电线路弧光高阻接地故障单端测距分析 [J]. 中国电机工程学报, 2017, 37(5): 1333–1341.
WANG Bin, NI Jiang, WANG Haigang, et al. Single-end fault location for high impedance arc grounding fault in transmission line[J]. Proceedings of the CSEE, 2017, 37(5): 1333–1341.
- [2] 刘传洋, 吴一全. 基于深度学习的输电线路视觉检测方法研究进展 [J/OL]. 中国电机工程学报, 2022: 1–24. (2022–08–31). <https://kns.cnki.net/kcms/detail/11.2107.TM.20220831.1053.002.html>.
LIU Chuanyang, WU Yiquan. Research progress of vision detection methods based on deep learning for transmission lines[J/OL]. Proceedings of the CSEE, 2022: 1–24. (2022–08–31). <https://kns.cnki.net/kcms/detail/11.2107.TM.20220831.1053.002.html>.
- [3] 马忠梅, 韩宝卿, 李文娟, 等. 智能识别技术在无人机输电线路巡检中的应用 [J]. 集成电路应用, 2022, 39(8): 126–127.
MA Zhongmei, HAN Baoqing, LI Wenjuan, et al. Application of intelligent identification in UAV inspection for transmission lines[J]. Application of IC, 2022, 39(8): 126–127.

- [4] 王杨杨, 曹晖, 莫文昊. 基于深度学习的改进型 YOLOv4 输电线路鸟巢检测与识别 [J]. 智慧电力, 2023, 51(1): 101–107.
WANG Yangyang, CAO Hui, MO Wenhao. Bird's nest detection and identification on improved YOLOv4 transmission line based on deep learning[J]. Smart power, 2023, 51(1): 101–107.
- [5] 李轶. 基于深度学习的无人机输电线路巡检故障检测研究 [J]. 流体测量与控制, 2022, 3(6): 56–58, 82.
LI Yi. Research on fault detection of UAV transmission line inspection based on deep learning[J]. Fluid measurement & control, 2022, 3(6): 56–58, 82.
- [6] 翟永杰, 杨旭, 赵振兵, 等. 融合共现推理的 Faster R-CNN 输电线路金具检测 [J]. 智能系统学报, 2021, 16(2): 237–246.
ZHAI Yongjie, YANG Xu, ZHAO Zhenbing, et al. Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection[J]. CAAI transactions on intelligent systems, 2021, 16(2): 237–246.
- [7] 和敬涵, 罗国敏, 程梦晓, 等. 新一代人工智能在电力系统故障分析及定位中的研究综述 [J]. 中国电机工程学报, 2020, 40(17): 5506–5516.
HE Jinghan, LUO Guomin, CHENG Mengxiao, et al. A research review on application of artificial intelligence in power system fault analysis and location[J]. Proceedings of the CSEE, 2020, 40(17): 5506–5516.
- [8] 刘召, 张黎明, 耿美晓, 等. 基于改进的 Faster R-CNN 高压线缆目标检测方法 [J]. 智能系统学报, 2019, 14(4): 627–634.
LIU Zhao, ZHANG Liming, GENG Meixiao, et al. Object detection of high-voltage cable based on improved Faster R-CNN[J]. CAAI transactions on intelligent systems, 2019, 14(4): 627–634.
- [9] 宋万潼, 李冰峰, 费树岷. 基于先验知识的航拍绝缘子检测方法 [J]. 计算机工程, 2021, 47(8): 301–307, 314.
SONG Wantong, LI Bingfeng, FEI Shumin. Prior knowledge-based insulator detection method using aerial images[J]. Computer engineering, 2021, 47(8): 301–307, 314.
- [10] 汤踊, 韩军, 魏文力, 等. 深度学习在输电线路中部件识别与缺陷检测的研究 [J]. 电子测量技术, 2018, 41(6): 60–65.
TANG Yong, HAN Jun, WEI Wenli, et al. Research on part recognition and defect detection of transmission line in deep learning[J]. Electronic measurement technology, 2018, 41(6): 60–65.
- [11] 高强, 廉启旺. 航拍图像中绝缘子目标检测的研究 [J]. 电测与仪表, 2019, 56(5): 119–123.
GAO Qiang, LIAN Qiwan. Research on target detection of insulator in aerial image[J]. Electrical measurement & instrumentation, 2019, 56(5): 119–123.
- [12] 赵振兵, 江爱雪, 戚银城, 等. 嵌入遮挡关系模块的 SSD 模型的输电线路图像金具检测 [J]. 智能系统学报, 2020, 15(4): 656–662.
ZHAO Zhenbing, JIANG Aixue, QI Yincheng, et al. Fittings detection in transmission line images with SSD model embedded occlusion relation module[J]. CAAI transactions on intelligent systems, 2020, 15(4): 656–662.
- [13] 戚银城, 江爱雪, 赵振兵, 等. 基于改进 SSD 模型的输电线路巡检图像金具检测方法 [J]. 电测与仪表, 2019, 56(22): 7–12, 43.
QI Yincheng, JIANG Aixue, ZHAO Zhenbing, et al. Fittings detection method in patrol images of transmission line based on improved SSD[J]. Electrical measurement & instrumentation, 2019, 56(22): 7–12, 43.
- [14] WAN Neng, TANG Xuming, LIU Siyan, et al. Transmission line image object detection method considering fine-grained contexts[C]//2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference. Chongqing: IEEE, 2020: 499–502.
- [15] 赵振兵, 李延旭, 甄珍, 等. 结合 KL 散度和形状约束的 Faster R-CNN 典型金具检测方法 [J]. 高电压技术, 2020, 46(9): 3018–3026.
ZHAO Zhenbing, LI Yanxu, ZHEN Zhen, et al. Typical fittings detection method with faster R-CNN combining KL divergence and shape constraints[J]. High voltage engineering, 2020, 46(9): 3018–3026.
- [16] JOYCE J M. Kullback-leibler divergence[M]//International encyclopedia of statistical science. Berlin: Springer, 2011: 720–722.
- [17] SENHAJI K, RAMCHOUN H, ETTAOUIL M. Training feedforward neural network via multiobjective optimization model using non-smooth L1/2[J]. Neurocomputing, 2020, 410: 1–11.
- [18] FAN Qi, ZHOU Wei, TANG Chikeung, 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. Seattle: IEEE, 2020: 4013–4022.
- [19] 刘兴亚, 周洋. 特高压交流输电线路跳线安装施工技术 [J]. 集成电路应用, 2021, 38(12): 94–95.
LIU Xingya, ZHOU Yang. Installation and construction technology of jumper for UHV AC transmission line[J]. Application of IC, 2021, 38(12): 94–95.

- [20] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [21] BOTTOU L. Stochastic gradient descent tricks[J]. Neural networks: tricks of the trade (second edition), 2012: 421–436.
- [22] 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.
- [23] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 936–944.
- [24] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[M]//Computer Vision – ECCV 2014. Cham: Springer International Publishing, 2014: 740–755.
- [25] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21–37.
- [26] REDMON J, FARHADI A. YOLOv3: an incremental im-

provement[EB/OL]. (2018-04-08)[2022-02-24].
<https://arxiv.org/abs/1804.02767>.

作者简介:



王巍, 硕士研究生, 主要研究方向为图像处理和电力视觉。



杨耀权, 教授, 博士, 主要研究方向为计算机视觉、智能测控技术。主持河北省科技计划项目 1 项, 主持横向科研项目 12 项。发表学术论文 80 余篇。



翟永杰, 教授, 博士, 主要研究方向为电力视觉。主持国家自然科学基金面上项目 1 项、河北省自然科学基金项目 1 项, 主持横向科研项目 16 项, 参与国家重点研发计划项目 1 项, 授权发明专利 10 项, 获山东省科技进步一等奖 1 项。发表学术论文 30 余篇, 编著 1 部, 参编著作 3 部, 参编教材 1 部。