Vol.43 No.7 Jul. 2023

规范透视场景的半监督目标检测及其在保护压板巡检上的应用

韩畅1,韩笑2,陈虹1,钟杰3,戈洋1,曹灿3,马杰3

(1. 电子科技大学 计算机科学与工程学院,四川 成都 611731;2. 南京工程学院 电力工程学院,江苏 南京 211167;
 3. 国家电网江苏省电力有限公司连云港供电公司,江苏 连云港 222000)

摘要:如何快速准确地对继电保护压板的异常状态进行识别,是变电站二次设备巡检工作中亟待解决的技术 难题。基于深度学习的通用目标检测算法在向诸如继电保护屏压板检测等特殊化场景的迁移中,不能够充 分利用保护屏场景中的规范透视先验特征;此外,人工标注大数据集的困难性一直以来都是通用检测模型迁 移至特殊场景时的挑战。针对上述问题,提出了一种适用于保护压板规范化分布特征的半监督目标检测模 型,该模型根据压板识别场景的特点对模型框架进行了一系列适应性改进。在模型的半监督训练阶段,使用 一致性正则化方法生成伪标签,并基于保护屏压板图像特征,以边缘吸附和点阵行列拟合等方式,优化或剔 除伪标签,从而突破了数据标注困难性带来的限制。改进后的模型达到平均精度均值为98.12%的应用级精 度,并额外输出图像的逆透视变换参量。该模型被应用于便携式智能终端,辅助工作人员进行继电保护压板 状态的巡检工作;模型输出的逆透视变换参量,也可为3D人机交互等下游视觉任务提供技术支撑。 关键词:继电保护压板;电力系统智慧化;半监督学习;目标检测算法;逆透视变换 **户图分类号:TP391.41 文献标志码:**A DOI:10.16081/j.epae.202209026

0 引言

继电保护压板是一种安装于继电保护屏上的重 要机械零件。工作人员通过投入或断开保护压板, 以选择继电保护的具体功能或决定保护装置是否动 作于断路器,因此压板是变电站二次回路的"关键"。 在变电检修及日常运行过程中,部分压板的投入或 断开状态需要按照调度指令进行人工切换,而为了 消除压板状态错误造成的事故隐患,工作人员需要 通过人工巡检的方式核对所有压板状态的正确性。 被巡检压板的数量极为可观,保护屏上压板布置方 式错综复杂,巡检人员的专业基础水平不一,加之缺 乏有效的辅助核对工具,导致工作效率低下,误检、 漏检时有发生。

在以变电站为代表的一次设备巡检工作中,智 能化的技术与装备正不断得到推广应用^[1]。然而, 在以继电保护压板为代表的二次设备巡检过程中, 受到识别准确率不足、复杂背景识别能力弱、图像采 集成本高等因素的制约,相关设备难以有效推广应 用,因此,如何在便携式智能终端对压板状态进行准 确的可视化识别,及时提醒巡检人员压板状态异常 的信息,提高压板巡检的数字化水平,已成为变电站 二次设备智能巡检过程中亟待解决的技术难题。

收稿日期:2022-03-29;修回日期:2022-07-20 在线出版日期:2022-10-19 基金项目:江苏省自然科学基金资助项目(BK20181021)

Project supported by the Natural Science Foundation of Jiangsu Province(BK20181021)

近10年来,基于深度学习的目标检测技术高速 发展。以基于锚框的检测方法为例,2014年提出的 基于区域的卷积神经网络(region-based convolutional neural networks, R-CNN)模型^[2]和2016年提出的 YOLO模型^[3]分别奠定了两阶段和单阶段目标检测 模型的技术基础。此后,基于这2类框架的改进与 优化层出不穷:单阶段检测技术中,经过多轮改进的 YOLOR-D6在保证检测速度的前提下,在COCO2017 测试集上的平均精度均值(mean average precision, mAP)达到了57.3%;两阶段目标检测技术也产生了 Cascade R-CNN^[4]、Libra R-CNN^[5]等改进方案。以 深度学习为基础的目标检测模型在2017年前后超 越了人眼识别的极限,并逐步在各高精度自动化检 测领域得到应用。

基于深度学习的目标检测算法曾在变电设备红 外图像识别^[6]、输电线路巡检^[7]等电力生产管理过 程中得到有效的工程应用,然而,与上述电力一次设 备检测工作不同的是,继电保护压板的检测过程仅 需满足准实时性标准,但对于精确度要求较高,且保 护屏集成的压板数量普遍较多,这给大规模数据集 的人工标注带来了极大的困难;同时,继电保护压板 的检测场景具有独特的先验特征,或可被用于提升 检测方法的精度或扩充数据集样本。

现存的压板状态检测方法以传统的图像特征提 取手段为主^[89],此类方法避免了大规模数据集的标 注问题,却难以满足精确性的要求,且通常需要苛刻 的图像采集条件,因此难以在移动设备上有效应用, 以文献[8]所提方法为例,该方法仅能适应本文第4 节所述的正面清晰图像集;此外少部分基于深度学 习目标检测算法的应用方案¹⁰⁰难以解决大规模数据 集标注的成本问题,从而不可避免地因数据集规模 过小导致模型的严重过拟合现象,也未能充分、良好 地利用压板图像的规范分布特征。

针对上述问题,本文提出了一种适应于继电保 护压板规范化分布特征的半监督目标检测模型,并 在较小的标注压板图像数据集和较大的未标注数据 集上联合训练,最终使模型的mAP达到98.12%的 应用级精度。该模型被应用于便携式终端中,辅助 工作人员进行继电保护压板状态的高效率巡检工 作。同时,该技术可被推广应用于保护端子排连接 正确性、断路器控制开关位置正确性等变电站二次 设备状态的核检工作中。

1 检测场景、对象与方法

218

1.1 检测场景:以保护屏为例的规范透视场景

本文所述的规范透视场景是指待检测目标主体 具有显著透视规律的目标检测场景。继电保护压板 检测场景具有鲜明的规范透视场景特征,具体体现 在:①压板安装于室内保护屏上,排布规范,呈矩形 阵列状,具有显著透视规律;②压板形制、尺寸统一, 相对于图像尺寸较小,且密集度较高;③压板之间不 存在包含、依附关系。

本文旨在设计一种适应复杂光照与拍摄条件的 压板状态检测模型,充分利用规范透视场景特征缓 解标注监督数据缺失所造成的一系列问题,从而有 效提升检测方法的精度。

1.2 检测预期及对象

对压板状态检测模型的预期是能够准确地从通 用拍摄设备在工作场景下获取到的压板屏图像中标 注出各压板的位置及状态。模型预期与压板分类见 附录A图A1,模型的检测对象为图像中的压板位置 与状态,压板位置以矩形框边界坐标表示;压板状态 可分为停用状态(连片被取下)、投入状态(压板正常 连通)与退出状态(压板未连通)3种。大部分采集 到的压板图像见图A1(a),其中少部分将由人工标 注图像中压板的位置及状态类别,形成标注数据集, 剩余的大部分图像不进行任何形式的标注组成未标 注数据集,本文使用上述2个数据集对模型进行半 监督训练,使其产生如图A1(c)所示的预测结果。

1.3 检测方法:半监督目标检测模型

已有的相关研究和预测试结果显示,在本文的 保护屏检测场景下,经典的通用目标检测算法即使 仅在较小数据集上监督训练(500张)亦能够确保 mAP在80%以上,故可以认为训练后的经典模型的 大部分预测是"基本精确"的。基于这一假设,本文 通过改进经典两阶段目标检测方法,使用同一增强 的模型框架构建教师模型与学生模型,在较小的监 督数据集上训练教师模型^[11];之后使用一致性正则 化方法对教师模型的预测输出进行修正,并结合保 护屏图像在透视关系、压板尺寸、排布规律等方面的 先验特征,剔除和纠正部分错误的预测框,从而为未 标注的数据生成一系列伪标签,并结合监督数据与 伪标签联合训练学生模型。反复迭代两学习过程, 最终达到提高训练精度的目的。

2 模型框架及其适应性改进

教师模型与学生模型在架构和训练流程上基本 一致。考虑到所使用模型的稳定性、精确性、代表性 和拓展性,本文改进增强Faster R-CNN作为模型框 架,同时作为对照参考半监督训练效果。本节将介 绍模型框架概要、实现细节及其在本文场景下做出 的适应性改进。

2.1 模型框架:增强的Faster R-CNN

Faster R-CNN是两阶段目标检测的经典模型, 通过调整区域生成网络(region proposal network, RPN),使其与检测头共享同一组图像卷积特征从而 大幅提升模型效率。

增强的 Faster R-CNN 框架见附录 A 图 A2。增强框架使用残差神经网络^[12]作为骨干网络以提取特征图,并使用平衡特征图金字塔网络(balanced feature pyramid network, BFPN)融合提取多尺度的图像特征;图像特征一方面作为 RPN 的输入,以滑动窗口的方式生成一系列锚框,并经由分类分支与回归分支生成候选框,经由非极大值抑制(non maximum suppression, NMS)等后处理后与图像特征进行兴趣区域调优(region of interest align, RoI Align)^[5],将结果输入目标检测头;增强框架使用级联边界框回归以提升检测精确度,通过使用多个级联的回归器^[4],设置递增的交并比(intersection over union, IoU)样本阈值将检测框不断回归至更精确的位置;最终合并级联的检测头的分类结果和末端检测头的回归结果以生成预测框,完成目标检测过程。

2.2 模型框架的实现细节

在模型框架中,本文综合考虑任务复杂性和网络能力,选择了Resnet50-vd作为骨干网络,其网络结构如附录A图A3所示,其对Resnet的主要改进是使用平均池化层代替了原有降采样过程从而减少信息损失。考虑到压板图像的比例特征,在RPN阶段选取(1.0,2.0,3.5)的长宽比及(12,24,48,96,192)的尺寸生成锚框,使用平衡交并比(IoU-balanced)采样以缓解样本"难度"之间的不平衡性,根据IoU的不同将采样阈值内的区间均匀划分为若干个采样单元,并在每单元内随机采样,检测头中的分类与回归过程与RPN过程类似。模型的RoI Align阶段中位于以(x_1, y_1)与(x_N, y_N)为顶点的候选框内的任意一点(x, y)处的特征向量F(x, y)可由式(1)插值产生。

$$F(x, y) = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{F(x_i, y_j)}{D((x, y), (x_i, y_j))MN}$$
(1)

式中:**F**为该坐标位置的特征向量;*M*、N分别为产生 插值的区域在横向与纵向的坐标范围;*D*(·,·)为两 坐标之间的距离函数。

2.3 模型框架中的适应性改进

相对于通用目标检测,压板状态检测存在一定 特殊性,即待检测物体之间不存在包含或依存关系, 且各物体尺寸相对一致。基于这种特殊性,可对模 型框架进行针对性改进。

首先,使用完整交并比(complete IoU,CIoU)^[13] 损失函数*L*_{CLoU}增强IoU损失函数,该损失函数由传统 IoU损失函数、框A与框B的距离损失函数、两框宽 高比损失函数三部分构成,如式(2)所示。

$$L_{\text{CIoU}} = L_{\text{IoU}} + L_{\text{Distance}} + L_{\text{WH}} = 1 - r_{\text{IoU}} + \frac{\rho^2 (A, B)}{d_{\text{Bound}}} + \alpha \nu \quad (2)$$

$$\begin{cases}
\nu = \frac{4}{\pi^2} \left(\arctan \frac{w_A}{h_A} - \arctan \frac{w_B}{h_B} \right)^2 \\
\alpha = \frac{\nu}{(1 - r_{\text{IoU}}) + \nu}
\end{cases} \quad (3)$$

式中: L_{IoU} 为传统 IoU 损失函数; $L_{Distance}$ 为框 A 与框 B 的距离损失函数; L_{WH} 为两框宽高比损失函数; r_{IoU} 为 两框的 IoU; ρ 为两框中心点的距离函数; d_{Bound} 为两 框最小外接矩形的对角线长度; ν 为衡量两框宽高 比一致性的函数; α 为调整宽高比损失占全局损失 的权重调节函数; w_A 、 h_A 和 w_B 、 h_B 分别为框 A 和框 B 的 宽、高。该损失函数可以增强原损失函数在特殊情 况下的适应性和精确性。

然后,使用"压板区域不相互包含"的先验特征 改进NMS过程。考虑如附录A图A4所示的情况,图 中错误交叠的预测框之间的IoU值均小于0.5,而二 者置信度又均高出阈值(如0.6),传统NMS过程不能 正确处理此类反例,本文使用前述先验特征增强 NMS判别条件,当某可信的预测框可能包含另一个 时,倾向于抛弃内部的预测框。框A与框B发生包 含的概率P_{inelude}为:

$$P_{\text{include}} = \left| \frac{S_{1}}{S_{A}} - \frac{S_{I}}{S_{B}} \right| \frac{\beta}{\left(1 - \alpha \max\left\{S_{1}/S_{A}, S_{I}/S_{B}\right\}\right)^{\gamma}} \quad (4)$$

式中: S_A 、 S_B 、 S_1 分别为框A、框B及两框相交部分的面积; α 、 β 与 γ 分别为 $P_{include}$ 值域范围及其各部分所占比重的调节参数。

3 伪标签生成、优化与纠错

本节在教师模型预测的基础上,结合一致性正则化相关方法,为未标注数据生成伪标签;分析利用 压板图像的先验特征,使用边缘吸附方法优化生成的伪标签坐标;最后利用压板的排列分布特征,使用 点阵的行列拟合以及消失点计算的方式,纠正或剔 除部分错误预测的伪标签。

3.1 基于一致性正则化的伪标签生成

平滑假设是一致性正则化乃至半监督学习中的 重要假设^[11],该假设认为:若两样本输入在样本空间 中足够相似,则其标签也应相似。具体至本文场景, 以 Auto-Augment^[14]、Mixup 和 Mosaic 为代表的通用 图像样本增强手段所产生的系列输出(见附录A图 A5),应当对相同的物体产生相同的分类和相近的 预测框位置。基于上述假设,本文使用教师模型对 同一张图片的各类数据增广样本进行预测,经逆变 换后合并预测结果并执行目标物聚类,之后对于不 同样本的预测结果,根据其相对目标物的距离和置 信度进行"投票",最终选取拥有足够"选票"的区域 的外接矩形作为软标签。标签的边界框伪监督信息 可通过计算某图像坐标(x, y)处的像素属于目标物i边界框的概率 $P_i(x, y)$ 产生, $P_i(x, y)$ 的计算公式为: $P_i(x, y)=$

$$\frac{1}{N_{\text{Aug}}} \sum_{k=1}^{N_{\text{Aug}}} \frac{1}{N_k} \sum_{j=1}^{N_k} \frac{f_j[(x, y) \in B_j] f_j[B_j \subseteq P_{\text{pred}k}] C_{\text{conf}}(B_j)}{D(c_i, c_{B_j})/D_{\text{avg}}}$$
(5)

式中: N_{Aug} 为增强样本数量; N_k 为第k个增强样本预测结果中有效候选框的数量; f_j [·]为条件判别函数, 当内部条件成立时函数值为1,否则为0; B_j 为第j个 预测框; P_{predk} 为第k个增强样本的预测结果; $C_{conf}(B_j)$ 为预测框 B_j 的置信度;c为物体或目标框的中心点; D_{avg} 为所有合法预测框与目标物中心点距离的调和 平均值。对于物体i的边界框,汇总概率 $P_i(x,y)$ 高 于阈值的像素区域,并以该区域的外接矩形作为软 标签坐标,并以区域内像素归属i的概率均值作为软 标签置信度,根据目标物所有预测结果中各类别的 占比生成软标签各分类的置信度。

可视化的软标签生成过程见附录A图A6。以图像中某一压板为例,对该压板的预测结果如图A6的 左下图所示。对各图进行变换后再进行预测,所得 的预测框经逆变换映射至原图,不同的预测框颜色 代表不同的类别。区域内各像素归属于该压板区域 的概率由图A6的右图中不同颜色的矩形半透明区 域叠加后的不透明度进行可视化,软标签各分类的 置信度则由区域内各颜色的比例进行可视化展示。

3.2 基于边缘吸附的伪标签优化

压板相对保护屏底板在图像上一般呈现明显的 梯度变化,因而本文使用边缘吸附方法优化3.1节生 成的伪标签坐标,该方法亦可以作为模型输出的后 处理过程以提升模型精度。

边缘吸附过程中,当判别边界框未经过、未贴合 边缘时,尝试将边界框范围调整至贴合附近的边缘。 本文使用一种基于引力梯度的自适应阈值Canny算 子^[15]以从边缘信息丰富的保护屏图像中提取边缘。 经典Canny算子中使用的梯度向量计算方法表示为:

220

$$\begin{cases} \left| \vec{g} \right| = \sqrt{\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2} \approx \sqrt{\left| E_x \right|^2 + \left| E_y \right|^2} \\ \psi = \arctan\left(\frac{\partial}{\partial y} / \frac{\partial}{\partial x}\right) \approx \arctan\left|\frac{E_x}{E_y}\right| \end{cases}$$
(6)

式中: $|\vec{g}|$ 为灰度梯度值; ψ 为梯度方向角; E_x 、 E_y 分别 为x、y方向的偏导值,由2×2的邻域差分算子近似计 算得到。本文引入灰度值引力场改进这一计算方 法,点i处的近似梯度向量 E_i 计算如下:

$$\boldsymbol{E}_{i} = \sum_{j} \frac{G(\boldsymbol{m}_{j} - \boldsymbol{m}_{i}) \left| \boldsymbol{m}_{j} - \boldsymbol{m}_{i} \right|}{\left\| \boldsymbol{r}_{ij} \right\|^{2}} \frac{\boldsymbol{r}_{ij}}{\left\| \boldsymbol{r}_{ij} \right\|}$$
(7)

式中:G为引力常数;r_{ij}为点*i、j*之间的坐标向量;m_i, m_j分别为点*i、j*的灰度值。E_i是对周围所有像素灰 度差产生的"场"的合并从而近似梯度向量。使用区 域梯度阈值自适应的方法,根据像素点位置周围区 域的梯度分布调整局部阈值以保证不损失边缘细 节。在执行边缘吸附算法时,若预测框的某个边界 未与任何边缘相交,则判别该边界处于目标压板图 像外侧,应向内调整;为避免发生错误的边缘吸附, 算法不向外部调整边界框。边缘检测效果的对比以 及边缘吸附流程如附录A图A7所示。

3.3 基于行、列拟合与消失点计算的伪标签纠错

保护屏压板通常以横向水平、纵向竖直的矩形 阵列排布,每行、列压板的坐标均可拟合为直线,并 汇于2个消失点,基于拟合生成的行、列直线以及消 失点可以排除预测框位置的显著错误。需说明的 是,图像的消失点检测问题已有许多有效解决方案, 例如使用Hough变换方法¹⁶¹等,这类方案大多需要 从图像内正确检测到足够多的直线或直线对,然而 当图像背景复杂时,在边缘信息丰富的保护屏图像 中检测到的直线精度将被大量噪声所干扰,类似 Hough变换的方法易产生部分错误的估计结果,该 类方法也浪费了模型的预测结果所包含的信息。

经过一致性正则化和边缘吸附的教师模型预测 结果附录A图A8所示,但其仍包含数个错误检出的 负样本,其中2个远离保护屏区域,另一个则混杂于 大量压板图像之间。本节所述的处理过程旨在从该 预测结果中排除类似错误。为方便表述,将附录A 图A8所示的预测结果抽象为一个带有标签的点阵, 需注意不是所有状态压板的图像中心均能充分靠近 压板的实际中心(空与投状态压板图像的中心一般 落在2个连接旋钮之间,而退状态压板图像的中心 则偏向一侧),因此这一抽象过程是通过抽取预测框 右上角顶点实现的,抽象点阵如图1所示。

行点集初步筛选的步骤如下:①统计点阵中各 点到最近邻点的距离,取中值作为压板间距估测值,



Fig.1 Dot matrix of prediction results

并将其乘以比例系数(如1/5)得到阈值 T_0 ,计算点阵规模乘以比例系数(如0.95)记作 T_1 ;②对于点阵中的每一点,均与其他所有点构成直线,统计与这些直线的距离小于阈值 T_0 的点构成点集S,选择合适的直线使S的规模最大,对于多个相同规模者按照S内各点的最大间距 d_{max} 对各直线进行升序排列;③根据各直线所对应 d_{max} 按比例对步骤②中选择的点集S进行投票,统计点阵中各点的投票,选择票数最高且互不重叠的点集 $S_0 - S_k$,使其总规模高于阈值 T_1 ,从而得到行点集的初步划分结果。

上述方法要求点阵的行列之间的规模存在显著 的差异,否则将无法确定步骤③所得投票结果为行 点集或列点集;且该方法要求图像的倾角不过分大, 否则将存在步骤③为斜向点集投票的可能。上述要 求并不会显著减少无监督数据集对模型的训练质 量,不符合要求的图像由人工判别并剔除,不参与训 练过程。图2为上述过程输出的行点集划分示例。



图 2 行点集划分示例 Fig.2 Division example of dot matrix in lines

使用完全最小二乘法(total least squares,TLS) 分别对划分出的行点集进行直线拟合。点 $p_i = (x_i, y_i)$ 到直线 $ax+by+c=0(a^2+b^2=1)$ 的距离可以表示为 $|ax_i+by_i+c|$,因此期望寻找直线参数(a,b,c)以最小 化规模为n的行点集 $R = \{p_i = (x_i, y_i) | i \in [1, n]\}$ 到直 线的距离平方和误差 δ,δ 的计算公式为:

$$\delta = \sum_{i=1}^{n} (ax_i + by_i + c)^2$$
(8)

令 δ 对c的偏导为0,则有:

$$\frac{\partial \delta}{\partial c} = \sum_{i=1}^{n} \left[-2 \left(a x_i + b y_i + c \right) \right] = 0$$
(9)

求解可得c为:

$$c = \frac{a}{n} \sum_{i=1}^{n} x_i + \frac{b}{n} \sum_{i=1}^{n} y_i = a\bar{x} + b\bar{y}$$
(10)

式中: \bar{x} 、 \bar{y} 分别为 $x_1 - x_n$ 、 $y_1 - y_n$ 的均值。

故δ可表示为如下形式:

$$\delta = \left\| \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \right\|^2 = (UN)^{\mathrm{T}} (UN) \quad (11)$$

式中:N=[a,b]^T,为待求解的参数矩阵。

令 δ 对N的导数为0,则有:

$$\frac{\mathrm{d}\delta}{\mathrm{d}N} = U^{\mathrm{T}}UN + N^{\mathrm{T}}U^{\mathrm{T}}U = 2(U^{\mathrm{T}}U)N = 0 \qquad (12)$$

式中: $U^{T}U$ 为关于x,y的二阶矩。 $U^{T}U$ 为对称矩阵, 又由于 $||N||^{2}=1$ 的限定,故 $U^{T}U$ 最小特征值对应的 特征向量即为所求N。直线拟合结果如图3所示。 使用 DBSCAN^[17]对TLS 过程所求得的行直线簇产生 的交点进行聚类,从而生成行消失点。



图 3 直线拟合结果 Fig.3 Fitting results by using straight lines

在此基础上,本文使用以下最优化算法完成对 点阵的列直线拟合。算法的输入为点集的划分结果 $R_1 - R_n$ 与过其消失点 $V_r = (x_{V_r}, y_{V_r})$ 拟合生成的直线 $l_{r1} - l_m$,划分中点的集合 $S = \bigcup_{i=1}^{k} R_i = \{s_1, s_2, \dots, s_k\}$,待 拟合的列直线条数为m;输出为列消失点 $V_c = (x_{V_c}, y_{V_c})$ 及对S拟合的m条列直线 $l_{c1} - l_{cm}$ 。

规定矩阵 $X_{a\times b}$ 的p范数为:

$$\left\| \boldsymbol{X}_{a \times b} \right\|_{p} = \left(\sum_{i=1}^{a} \sum_{j=1}^{b} x_{ij}^{p} \right)^{\frac{1}{p}}$$
(13)

直线
$$l_{r1}$$
— l_{rn} 、 l_{c1} — l_{cm} 均以直线倾斜角表示为:

$$\begin{cases}
F(\theta_i, x_v, y_v) = (x - x_v)\sin \theta_i + (y - y_v)\cos \theta_i = 0 \\
x\sin \theta_i + y\cos \theta_i - x_v\sin \theta_i - y_v\cos \theta_i = 0
\end{cases}$$
(14)

式中:*θ*_i为第*i*条行直线倾斜角,按升序排列。

同样地,按照升序排列列直线倾斜角 β_i ,本条列 直线与前一条列直线倾斜角之差 $\Delta\beta_i = \beta_{i+1} - \beta_i$,从而 构成m = 1行的矩阵 $\Delta\beta$,如式(15)所示。 $\Delta \beta = [\Delta \beta_1, \Delta \beta_2, ..., \Delta \beta_{m-1}]^{T}$ (15) 联立行、列直线方程(若平行,则向列直线方程 中的 β_i 值添加一极小量 γ),解得 $m \times n$ 个交点坐标, 记为 S_c ,如式(16)所示。

$$S_{c} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & & \vdots \\ p_{n2} & p_{n2} & \cdots & p_{nm} \end{bmatrix}$$
(16)

式中:*p_{ij}为第i*行直线与第*j*列直线相交形成的交 点。若*S*的规模为*k*,则定义每个交点*p_{ij}*与*S*中每一 点的距离向量*D_i*为:

 $D_{ij} = [D(p_{ij}, s_1), D(p_{ij}, s_2), \dots, D(p_{ij}, s_k)]^{T}$ (17) 优化式(18)所示的目标函数 F_{OPT} 以拟合列直 线簇。

$$F_{\text{OPT}} = \sum_{i=1}^{n} \sum_{j=1}^{m} \left\| \boldsymbol{D}_{ij} \right\|_{\omega} + \lambda \sigma (\Delta \boldsymbol{\beta}) + \eta \left\| \Delta \boldsymbol{\beta} \right\|_{\kappa}^{-1} \quad (18)$$

$$\sigma(\Delta \boldsymbol{\beta}) = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m-1} (\Delta \boldsymbol{\beta}_i - \overline{\Delta \boldsymbol{\beta}})}$$
(19)

式中: $\sum_{i=1}^{n} \sum_{j=1}^{m} \| D_{ij} \|_{\omega}$ 为所有交点到S内所有点距离的 ω 范数; λ, η 为超参数; $\sigma(\Delta \beta)$ 为相邻直线间夹角的均 方差; $\overline{\Delta \beta}$ 为相邻直线间夹角的均值; $\| \Delta \beta \|_{\kappa}^{-1}$ 为直线间夹角、范数的倒数。

值得指出的是,式(18)中的 $\sum_{i=1}^{n} \|D_{ij}\|_{\omega}$ 用于衡量直线簇 $l_{c1} - l_{cm}$ 对点集S的拟合程度。在式(18)中设置 $\sigma(\Delta \beta)$ 的目的是防止直线倾角间的差距过大,设置 $\|\Delta \beta\|_{\kappa}^{-1}$ 的目的是避免直线簇间的夹角过小而仅拟合某—列点集。在实践中,可以取 $\omega = -2, \kappa = -2$ 。

具体而言,最优化计算过程分为2个阶段。

1)第一阶段统计给定行的点集规模,选择规 模为中值的点集 R_i ,随机选定列消失点位置 $V_e^{(0)}$,过 点 $V_e^{(0)} = R_i$ 中各点分别构造 $|R_i|$ 条直线以构造列直 线簇参数矩阵计算交点矩阵 S_e ,此时目标函数是 对变量 $V_e = (x_{ve}, y_{ve})$ 的二元函数,分别对 x_{ve}, y_{ve} 求偏 导,使用梯度下降法优化该目标函数,迭代生成消 失点 V_e^* 。

2)第二阶段固定列消失点 V_e^* ,以第一阶段中的 直线构造方法生成列直线倾角参数 $\theta_e^{(0)}$,作为迭代初 始值输入式(18),此时目标函数是对直线簇倾角 θ_e 的m元函数,使用梯度下降法优化目标函数,迭代生 成列直线簇倾角参数 θ_e^* 。

使用上述最优化计算过程,可以计算点到行、列 拟合直线的距离,以剔除位置错误的预测点,拟合结 果与错误点剔除示例如图4所示。

综合使用本节所提方法,可以实现无监督数据 集的伪标签的生成、优化与错误剔除。



图4 拟合结果与错误点剔除示例

Fig.4 Example of fitting result and error dots elimination

4 实验与评估

222

4.1 数据集及数据增强方法

根据图像拍摄环境及识别的复杂程度,对采集 到的压板图像进行差异化选择并标注1000张图像 归入4个难度类别,其中500张作为教师模型的监督 训练数据,其余500张作为测试集评估模型能力,并 针对各类数据集各采集2500张作为未标注数据以 训练学生模型。各数据集采集要求和样本参数如附 录A表A1所示。特别地,所采集的无监督样本需确 保图像包含且仅包含1个保护屏及其上的所有压 板,否则将导致伪标签生成错误。本文在教师模型 训练阶段与伪标签一致性正则化阶段中均结合应用 Auto-Augment与Mosaic数据增强方法,另外根据半 监督通用目标检测相关研究的经验,使用 Auto-Augment方法进行数据增强时,教师模型训练阶段 设置相较于生成伪标签时相等数量的图像变换方式 和更高的数据增强强度,一致性正则化阶段的强度 被启发式地设置为教师模型训练阶段的40%。

4.2 半监督训练过程

半监督训练由监督和无监督过程交替迭代完成,为确保教师模型的初始准确率,采用 AI Studio 开源的 ResNet50-vd-SSLDv2-FPN 模型参数进行初始 化^[18],然后完全在监督数据集上完成第一次 Fine-Tune 过程,此时模型在测试集整体上的 mAP 已达 91.89%。再从无监督数据集中按类别随机抽取与 当前监督数据集规模大致相同的图像,使用第3节 所述方法为这些图像生成伪标签并加入监督数据集 内,使用新的监督数据集继续训练该模型,迭代上述 过程直至所有无监督数据均被标注。每阶段训练所 使用的图像均以4:1的比例划分为训练集与验证 集,与表A1所示的测试集不存在重叠。

实验的训练与测试均运行于2张RTX A6000显 卡,使用CUDA11.2与CUDNN-v8.1.1.33架构进行并 行化加速。教师模型的第一次训练使用监督数据集 中的500张图像,共计250个阶段、迭代7813次,开 始阶段使用线性warmup策略执行250次,之后以 0.0025的初始学习率和余弦学习率衰减策略调整学 习率;使用基于 Adam 的优化器完成梯度更新。各 训练阶段均选取前一阶段在测试集上拥有最佳表现 的模型作为教师模型。之后的训练阶段中,除不使 用 warmup 策略外,仅对学习率与训练阶段数进行调 整,其余策略及参数配置不作改变,各训练阶段具体 参数及其训练用时见附录 A 表 A2。另考虑到初次 训练后的教师模型在大倾角复杂测试集上的 mAP 仍不理想,因此第一次伪标签生成时所选择的无监 督数据集不包括大倾角复杂数据集。

4.3 模型效果评估

各训练阶段的全局模型损失及在4类测试集上的mAP如图5所示。图中:P_{mAP05}为IoU阈值为0.5时的mAP;波形截断了少量训练初期的极高模型损失数据,如图中黑色点线所示,使用k(k=10)移动平均对损失数据进行平滑处理。由图可见,每个新的无监督训练阶段开始时全局模型损失会有小幅上升,但模型在测试集上的表现基本不受此影响,每个阶段的全局模型损失及模型在验证集和测试集上的表现最终均趋于稳定。





为验证本文所增强模型及本文所使用的方法所 带来的效果提升,本文使用监督数据训练了包括经 典Faster R-CNN与YOLOv3在内的各类目标检测模 型,并在4类测试集上进行评估测试。表1为以每类 测试集下不同 IoU 阈值的 mAP 为评估指标的测试结 果。表中:P_{mAP-075}为IoU阈值为0.75时的mAP。由 表可见:模型本身的适应性、后处理中的边缘吸附方 法效果以及无监督训练过程效果均有了明显的提 升,其中边缘吸附的后处理流程带来了0.35个半分 点的提升,验证了边缘吸附过程在本文场景下的有 效性;半监督训练过程带来了5.88个百分点的mAP 提升,验证了本文使用的半监督方法对数据集扩充 的重要意义;本文方法的P_{mAP.05}最终达到了98.12%, 在正面清晰与大倾角清晰测试集上的效果尤为突 出,分别达到了99.72%和99.17%,这意味着适当保 证图像的拍摄条件时,本文方法可保证压板识别结 果的精确性与稳定性;模型的预测速度可达9帧/s, 满足准实时性的检测要求;此外,并无迹象表明无监



表1 不同模型及训练方法的测试结果

Table 1 Test Results of different models and training methods

	P _{mAP-0.5} / %				P _{mAP-0.75} / %					空洞中平 座 1		
模型及训练方法		正面 清晰	正面 复杂	大倾角 清晰	大倾角 复杂	均值	正面 清晰	正面 复杂	大倾角 清晰	大倾角 复杂	均值	⁻ 顶侧速度 / (帧·s ⁻¹)
YOLOv3		85.96	80.77	82.86	76.58	81.54	79.91	73.82	76.81	69.63	75.04	32
Faster R-CNN(Baseine)		90.67	85.79	87.79	81.98	86.56	84.20	81.12	82.36	77.02	81.18	17
Cascade R-CNN		92.99	88.11	90.12	83.49	88.68	87.35	83.90	86.24	79.71	84.30	11
Cascade R-CNN+FPN+ssld		95.23	90.24	92.31	85.81	90.90	91.12	86.56	89.60	81.51	87.20	14
	监督数据	96.60	91.36	92.87	86.72	91.89	93.23	89.65	90.51	82.72	89.03	12
本文模型	监督数据+边缘吸附	96.94	91.49	93.49	87.03	92.24	93.93	90.45	92.21	85.59	90.55	9
	全部数据+边缘吸附	99.72	97.94	99.17	95.64	98.12	98.02	95.48	96.19	93.03	95.68	9

督学习阶段对模型效果的提升达到饱和,更大数据 规模的无监督学习极有可能进一步提升模型效果。

4.4 压板图像的可视化比对

在压板图像识别结果的基础上,利用3.3节中拟 合所得的2组直线簇及其消失点,使用逆透视变换 方法^[19-20],完成图像内参数的标定,从而消除图像中 的透视形变,进而完成压板图像的可视化比对,方便 工作人员检查压板的变动。压板状态识别、逆透视 变换与结果对比如附录A图A9所示,其直观展示了 3.3节所述的行、列直线拟合以及消失点计算结果应 用于下游计算机视觉任务的可行性。

图 A9(a)、(c)分别展示了本文模型对包含同一 块保护屏的2张不同图像的状态识别结果,这2张图 像的拍摄角度与光照条件均不相同,其中有6个压 板的状态发生了改变,应用第3.3节所述的行、列直 线拟合以及消失点计算方法,并进行逆透视变换还 原其正面图像分别见图 A9(b)、(d),构建2张图像的 压板状态矩阵,从而对比状态发生变化的压板并标 记压板位置,图 A9(e)展示了抽象化的对比结果。

5 结论

本文以继电保护压板屏为研究对象,重点研究 规范化透视场景下的目标检测精度提升方法,在对 Faster R-CNN模型进行适应性增强的基础上,提出 一种使用一致性正则化与保护屏透视先验特征生成 伪标签的半监督模型训练方法,并在较小规模监督 数据集(500张)和较大规模无监督数据集(10000张) 上进行联合训练,所得结论如下。

1)采用本文所提训练方法的模型识别精度达 到了不低于98.12%的应用级水准,基于该方法的 智能移动终端应用在实际测试中取得良好的检测效 果,验证了本文方法的稳定性、精确性和鲁棒性。

2)本文所使用的半监督训练方法具有广泛的拓展性,该方法可被迁移到绝大多数支持Fine Tune的通用目标检测模型的训练中,例如其可在其他变电站二次设备(如端子排)的巡检工作中推广应用。

3)本文模型对透视参数的估计可直接用于图像 的内外参数标定或被识别对象的姿态估计,并可为 增强现实或3D人机交互等下游任务提供技术支撑。

附录见本刊网络版(http://www.epae.cn)。

参考文献:

- [1]任宝军,高志勇.一种基于边缘计算的分散式站所终端方案的 设计与实现[J].综合智慧能源,2022,44(6):59-69.
 REN Baojun, GAO Zhiyong. Design and implementation of a terminal configuration scheme in a decentralized distribution station based on edge computing[J]. Integrated Intelligent Energy,2022,44(6):59-69.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014; 580-587.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] //2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016:779-788.
- [4] CAI Zhaowei, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection [C] //IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018;6154-6162.
- [5] PANG Jiangmiao, CHEN Kai, SHI Jianping, et al. Libra R-CNN: towards balanced learning for object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019:821-830.
- [6]朱惠玲,牛哲文,黄克灿,等.基于单阶段目标检测算法的变电 设备红外图像目标识别及定位[J].电力自动化设备,2021,41 (8):217-224.
 ZHU Huiling,NIU Zhewen,HUANG Kecan, et al. Identification and location of infrared image for substation equipment
- based on single-stage object detection algorithm[J] Electric Power Automation Equipment, 2021, 41(8): 217-224.
- [7]林刚,王波,彭辉,等. 基于改进Faster-RCNN的输电线巡检图 像多目标检测及定位[J]. 电力自动化设备,2019,39(5):213-218.
 LIN Gang, WANG Bo, PENG Hui, et al. Multi-target detection and location of transmission line inspection image based on improved Faster-RCNN[J]. Electric Power Automation Equipment,2019,39(5):213-218.
- [8] 付文龙,谭佳文,吴喜春,等.基于图像处理与形态特征分析的 智能变电站保护压板状态识别[J].电力自动化设备,2019,39 (7):203-207.
 FU Wenlong, TAN Jiawen, WU Xichun, et al. Protection platen status recognition based on image processing and morphologi-

status recognition based on image processing and morphological feature analysis for smart substation [J]. Electric Power Automation Equipment, 2019, 39(7):203-207.

[9] 卜强生,高磊,闫志伟,等. 智能变电站继电保护软压板防误操 作策略及实现[J]. 电力自动化设备,2016,36(12):156-160,168. BU Qiangsheng,GAO Lei,YAN Zhiwei,et al. Strategy and im-

- ment,2016,36(12):156-160,168.
 [10] 张森,万吉林,王慧芳,等.基于注意力机制的卷积神经网络指针式仪表图像读数识别方法[J].电力自动化设备,2022,42 (4):218-224.
 ZHANG Sen,WAN Jilin,WANG Huifang, et al. Convolutional neural network based on attention mechanism for reading recognition of pointer-type meter images[J]. Electric Power Automation Equipment,2022,42(4):218-224.
- [11] XIE Qizhe, LUONG M T, HOVY E, et al. Self-training with noisy student improves ImageNet classification[C]//IEEE Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA; IEEE, 2020; 10687-10698.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] //2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, NV, USA: IEEE, 2016:770-778.
- [13] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence. California, Palo Alto, USA: AAAI, 2020:12993-13000.
- [14] CUBUK E D, ZOPH B, MANE D, et al. AutoAugment: learning augmentation strategies from data[C]//IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 113-123.
- [15] RONG Weibin, LI Zhangjing, ZHANG Wei, et al. An improved CANNY edge detection algorithm [C] //2014 IEEE International Conference on Mechatronics and Automation. Columbus, OH, USA: IEEE, 2014:577-582.
- [16] LUTTON E, MAITRE H, LOPEZ-KRAHE J. Contribution to

the determination of vanishing points using Hough transform [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(4): 430-438.

- [17] KHAN K, REHMAN S U, AZIZ K, et al. DBSCAN: past, present and future [C]//The 5th International Conference on the Applications of Digital Information and Web Technologies (ICADIWT). Bangalore, India: IEEE, 2014:232-238.
- [18] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[J]. Computer Science, 2015, 14(7): 38-39.
- [19] NIETO M, SALGADO L, JAUREGUIZAR F, et al. Stabilization of inverse perspective mapping images based on robust vanishing point estimation [C] //2007 IEEE Intelligent Vehicles Symposium. Istanbul, Turkey: IEEE, 2007:315-320.
- [20] BERTOZZ M, BROGGI A, FASCIOLI A. Stereo inverse perspective mapping: theory and applications[J]. Image and Vision Computing, 1998, 16(8):585-590.

作者简介:



韩 畅(1999—),男,硕士研究生,主要 研究方向为计算机视觉、区块链与网络安全、 联邦学习、隐私计算(E-mail:hanchang@std. uestc.edu.cn);

韩 笑(1969—),男,教授,主要研究
 方向为电力系统继电保护、电力系统自动化
 (E-mail:hanxiao@njit.edu.cn);

韩 畅

陈 虹(1974—),女,副教授,博士,主 要研究方向为隐私计算、数据安全、区块链

与网络安全(**E-mail**:chenhong8010@uestc.edu.cn)。

(编辑 任思思)

Semi-supervised object detection for normative perspective scenes and its application on protective connecting pieces

HAN Chang¹, HAN Xiao², CHEN Hong¹, ZHONG Jie³, GE Yang¹, CAO Can³, MA Jie³

(1. School of Computer Science and Engineering, University of Electronic Science and Technology of China,

Chengdu 611731, China; 2. School of Electrical Engineering, Nanjing Institute of Technology, Nanjing 211167, China;

3. Lianyungang Power Supply Company of State Grid Jiangsu Power Co., Ltd., Lianyungang 222000, China)

Abstract: Efficient and accurate identification of abnormal status of protective connecting piece is a pressing technical challenge in the inspection of substation secondary equipment. Deep learning-based generic target detection algorithms are not able to take full advantage of the canonical perspective priori features in protection screen scenarios when migrating to specialization scenarios such as relay protection panel connecting piece detection. The difficulty of manually labeling large data sets has been a challenge when migrating generic detection models to special scenarios. To address the above issues, a semi-supervised object detection model adapted to the normative distribution features is proposed, and a series of adaptations have been made to the model according to the connecting pieces identification scenes. In the semi-supervised training phase of the model, pseudo-labels are generated by using the consistent regularization method, and optimized or rejected by using edge adsorption and dotted ranks fitting based on features of protective screen, thus breaking through the limitations imposed by the difficulty of data annotation. The improved model can achieve an application-level accuracy of 98.12 % mean average precision and can additionally output the inverse perspective transformation parameters of the image. The model is applied to a portable intelligent terminal to assist staff in connecting piece inspection. The inverse perspective transformation parameters output by the model can also provide technical support for downstream vision tasks such as 3D humancomputer interaction.

Key words: protective connecting pieces; intellectualization of power systems; semi-supervised learning; object detection algorithm; inverse perspective transformation



图 A2 增强的 Faster R-CNN 框架 Fig.A2 Enhanced Faster R-CNN



图 A4 原 NMS 逻辑可能导致的错误案例 Fig. A4 Bad instances that primal NMS process may lead to



图 A5 可使用的图像增强变换方法

Fig.A5 Available image augmentation transformers



图 A6 一致性正则化过程可视化 Fig.A6 Visualization of consistency regulation process



图 A7 边缘提取方法与边缘吸附过程

Fig.A7 Edge detection method and edge attachment process



图 A8 错误预测示例 Fig.A8 Example of wrong prediction



(a)保护屏状态1识别结果

(b)状态1逆透视变换

(c)保护屏状态 2 识别结果

(d)状态 2 逆透视变换

(e)压板变动检测 抽象化结果

图 A9 保护屏状态变动后的检测结果对比 Fig.A9 Comparison between detection results after state change of protection screen

表 A1 保护屏图像数据集参数

Table A1 Parameters of protective base board image dataset											
厉伤米动	I	网络公律田子			数据集	駳模		压板样本数量			
图像尖型		图像木	果安水	有监督	无监督	测试集	共计	状态	有监督	测试集	共计
正面 清晰	镜头俯仰 倾角范围	± 10°	_充分光照,无大面积阴影, 无明显失焦模糊或运动模 _ 糊,10M 像素以上图像	125	2500	125	2750	投	2426	2386	4812
	镜头左右 倾角范围	± 10°						退	2012	2127	4139
	镜头旋转角范围	±5°						空	1872	1961	3833
	镜头俯仰 倾角范围	± 15°	至少满足下列条件之一: (1)阴暗关照(2)大面积 阴影投射(3)玻璃板反光 (4)运动或失焦模糊	125	2500	125	2750	投	2293	2766	5059
正面 复杂	镜头左右 倾角范围	± 15°						退	2423	2278	4701
	镜头旋转角范围	$\pm 10^{\circ}$						空	1749	1550	3299
大倾角 清晰	镜头俯仰 倾角范围 ^{±15°}	±15° ±20°	充分光照,无明显失焦模糊 或运动模糊,10M 像素以上 图像	125	2500	125	2750	投	2809	2096	4905
		115~150						退	1904	2014	3918
	镜头左右	±15° ± 45°						空	1385	1544	2929
大倾角 复杂	倾角范围	113~143	至少满足下列条件之一: (1)大面积阴影投射(2) 玻璃板反光(3)运动或失 焦模糊		2500	125	2750	投	2756	2645	5401
	镜头旋转角范围	$\pm 10^{\circ}$ ~ $\pm 15^{\circ}$		125				退	2020	2349	4369
								空	1873	1869	3742
注1:无出	监督保护屏图像应得	确保压板行列	刘规模的差值高于 2								
注 2: 当打	注 2:当拍摄倾角导致压板图像重叠时,两图像 IoU 不超过 50%,否 500 10000 500 11000 总计 25522 25585 51107										
则弃用该	则弃用该图像										

表 A2 各训练过程配置及参数

Tuble 112 Configuration and parameters of ouch manning process										
训练过程	训练集+验证集 (规模/张)	图像来源	初始学习率	训练阶段 (Epoch)	训练步数(Step)	训练用时(min)				
监督训练	500	4×监督集	0.0025	250	7813	83.6				
半监督训练-1	1000	前一阶段+3×无监督集 (不包括大倾角复杂)	0.0020	125	7812	85.8				
半监督训练-2	2000	前一阶段 +大倾角复杂	0.0016	60	7489	81.4				
半监督训练-3	4000	前一阶段	0.0012	30	7477	80.9				
半监督训练-4	10000	前一阶段	0.0009	15	10981	117.3				