

DOI:10.19651/j.cnki.emt.2209630

基于图像增广与迁移学习的输电线路 金具多目标实时检测方法^{*}

黄力^{1,2} 万旭东¹ 王凌云¹ 刘兰兰³

(1.三峡大学电气与新能源学院 宜昌 443002; 2.三峡大学湖北省输电线路工程技术研究中心 宜昌 443002;

3.带电巡检与智能作业技术国网公司实验室 长沙 421000)

摘要: 架空输电线路金具的状态评估工作对于线路的可靠运行至关重要,金具的检测是评估工作的重要一环。针对金具识别检测中数据集人工标注的工作量大,以及难以兼顾高精度和快速性问题,提出一种基于YOLOX网络改进的输电线路金具检测方法。将无人机拍摄的金具图像进行增广预处理丰富数据集,骨干网络采用在线Mosaic、Mixup增强方式,引入基于特征提取的迁移学习并采用余弦退火学习率进行两阶段模型训练。实验结果表明,改进后的方法对各类金具检测的平均精度均值提高了18.32%,与Faster R-CNN等5种主流检测模型相比,所提方法平均检测精度均值最高,且检测速度仅次于YOLOv3,能够更加快速、精准地识别各类金具,并在一定程度上减少人工标注的工作量。

关键词: 深度学习;卷积神经网络;迁移学习;实时检测;金具

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Multi-objective real-time detection method of transmission line fittings based on image augmentation and transfer learning

Huang Li^{1,2} Wan Xudong¹ Wang Lingyun¹ Liu Lanlan³

(1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;

2. Hubei Provincial Engineering Technology Research Center for Power Transmission Line, China Three Gorges University, Yichang 443002, China; 3. Live Inspection and Intelligent Operation Technology State Grid Corporation

Laboratory, Changsha 421000, China)

Abstract: The state assessment of overhead transmission line fittings is crucial to the reliable operation of the line, and the detection of the fittings is an important part of the assessment work. In response to the heavy workload of manual labeling of datasets in identification and detection of fittings, as well as the difficulty of balancing high precision and rapidity, an improved transmission line fittings detection method based on YOLOX network is proposed. The fitting images captured by UAV are augmented with preprocessing to enrich the datasets. The backbone network adopts the enhancement methods of online Mosaic and Mixup. The transfer learning based on feature extraction is introduced and the cosine annealing learning rate is used for two-stage model training. The experimental results show that the mean average precision of the improved method for the detection of all types of fittings is improved by 18.32%. Compared with five mainstream detection models such as Faster R-CNN algorithm, the mean average precision of proposed method is the highest, and its detection speed is lower than YOLOv3's, which can identify various types of fittings more quickly and accurately, and reduce the workload of manual labeling to a certain extent.

Keywords: deep learning; convolutional neural network; transfer learning; real-time detection; fittings

0 引 言

架空输电线路作为电能输送的载体,其运行状态关系

着电网的安全性与电能质量^[1-2]。架空输电线路中用于支撑、固定、悬挂的金具多安装在户外,容易出现锈蚀、错位、断裂等现象,产生电网安全隐患^[3]。采用深度学习的方法

收稿日期:2022-04-15

^{*} 基金项目:国家自然科学基金(51907104)、湖北省输电线路工程技术研究中心开放课题(2019KXL05)项目资助

对金具进行检测识别是状态分析和故障检测的基础,对实现自动化巡检具有重大意义。

在输电线路巡检中,使用无人机、带电作业机器人等设备的智能化巡检方式有着便捷、精确、安全的特点。随着智能电网的不断建设,采用智能化巡检替代传统人工巡检已经成为必然的趋势^[4]。在输电线路智能化巡检中对各目标进行检测识别是评估线路运行状态的重要一环,输电线路巡检中的目标检测方法主要分为基于人工提取特征的传统检测方法和基于深度学习的检测方法^[5],后者在多目标、精准性、快速性方面优势明显,更加适用于线路巡检。

近年来,人工智能技术不断走向成熟,深度学习方法在电力智能化检测领域也得到应用并取得了成效^[6-8]。基于深度学习的目标检测方法通过深度卷积神经网络(convolutional neural networks,CNN)提取目标特征,可分为两段和单段识别模型。两段识别模型需要预先提取候选区域再进行筛选,精度较高,速度相对较慢,以 R-CNN^[9]、Fast R-CNN^[10]、Faster R-CNN^[11] 模型为代表。白洁音等^[12]采用 Faster R-CNN 对输电线路部分故障进行检测,能够获得较高的精度,但速度上难以适应野外无人机等设备的实时巡检。赵振兵等^[13]结合形状约束和 KL 散度改进 Faster R-CNN,对多目标金具进行检测,一定程度上提高了整体精度,但仍存在检测速度慢、部分小目标金具检测精度不高的问题。单段识别模型对网络进行端到端的训练,识别速度快,主要有 YOLO(you only look once)系列^[14]、SSD(single shot multibox detector)^[15] 等模型。张永翔等^[16]运用迁移学习的方法训练 YOLOv3 模型,在小样本集上实现了防震锤和线夹两类金具识别精度的提高,但检测的目标类别较少,无法满足线路多种目标的巡检要求。赵

振兵等^[17]针对金具的遮挡问题对 SSD 进行改进,提高了部分遮挡情况下的检测精度,但 U 型挂环等小目标金具识别率较低。此外,上述方法均以大型样本数据集为模型学习的基础,在训练时数据集图像依赖于人工挑选与标注,无疑存在较大的人工工作量。

针对上述问题,本文提出一种基于新一代深度学习 YOLOX 网络的输电线路金具多目标实时检测方法。采用图像增广预处理方法,使用更小的数据集达到更好的训练效果,降低人工筛选数据集的工作量同时大幅提升网络检测精度;骨干网络中采用在线 Mosaic、Mixup 增强以提高模型对复杂背景下小目标金具的检测能力;训练时引入基于特征提取的迁移学习策略实现参数特征共享,降低训练成本同时增强模型特征提取能力;采用余弦退火学习率加速模型收敛,帮助跳出局部最优解。最终在实验结果中验证了改进方法的高精度与快速性,为复杂背景下的无人机等设备航拍图像中多目标金具的实时故障检测提供新的思路。

1 YOLOX 模型原理

1.1 整体结构

YOLO 系列模型是一种基于 CNN 的目标检测模型,以其快速实时性而广泛应用于自动驾驶汽车等各种前沿领域的实时对象检测场景。YOLOX 网络于 2021 年提出,是 YOLO 系列新一代实时目标检测网络^[18]。在与同类检测模型参数量相当的情况下,其在 MS COCO 数据集^[19]上有着比其他模型更高的检测精度以及更快的速度。

YOLOX 网络基于 YOLOv3^[20]进行改进,其网络结构由骨干网络、特征融合网络以及预测 3 个部分组成,如图 1 所示。

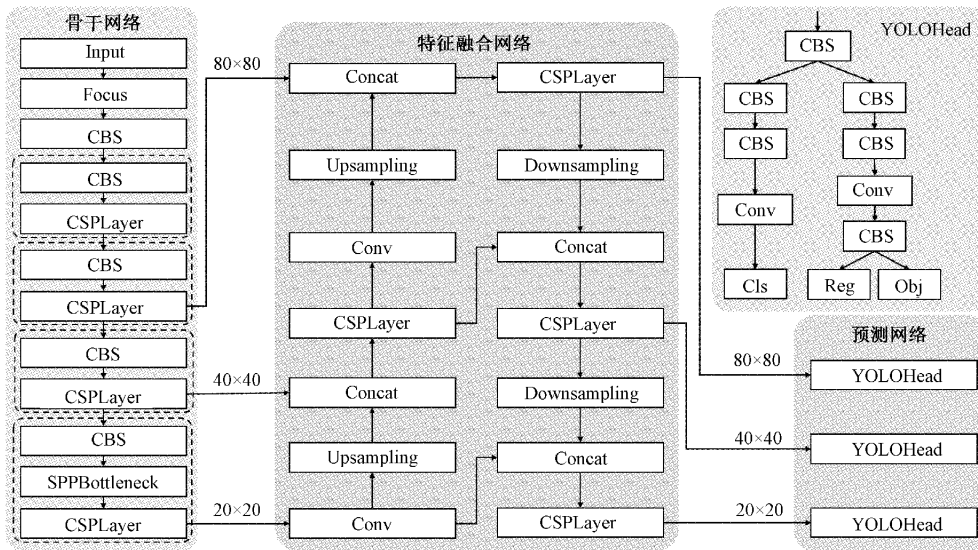


图 1 YOLOX 网络结构

1.2 骨干网络

YOLOX 骨干网络为 CSPDarknet53,主要包括 Focus、

CBS、CSPLayer 及 SPP Bottleneck 四种特征提取模块。当图片输入骨干网络后,首先进行自适应缩放将尺寸调整为

640×640 以避免图像信息丢失,然后进入不同模块进行特征提取。Focus 模块通过如图 2 的切片操作对输入的金具图像进行处理,实现对图像的无损下采样。CBS 卷积块由卷积层、批归一化^[21]层、SiLU 激活函数层 3 个部分组成。CBS 卷积块与 CSPLayer^[22] 残差结构共同构成了 Resblock_body 模块,该模块能够不断压缩图像尺寸扩充通道信息,增加网络深度同时缓解梯度消失问题。网络中前 3 个 Resblock_body 结构相同,最后一个加入了 SPP Bottleneck^[23] 结构,通过不同大小的最大池化核^[24] 扩大感受野,同时自适应不同金具图像的分辨率。在骨干网络中,不同深度的网络层对输入图像进行特征提取,最终输出 3 种尺寸的特征图,经过较少层网络输出的特征图感受野较小,包含更多整体特征,适合大目标的检测,而经过更多层网络得到的特征图包含更多细节特征,适合小目标的提取。

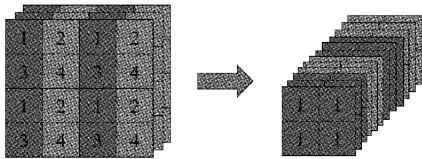


图 2 Focus 切片操作

1.3 特征融合网络

在特征融合网络中,YOLOX 基于 YOLOv3 采用 FPN+PAN 的双向融合方式,其结构如图 3 所示。特征金字塔网络(feature pyramid networks, FPN)能够提取更多深层网络语义信息^[25],路径聚合网络(path aggregation networks, PAN)能够提取更多浅层网络定位特征,两种方式融合特征信息能够增强网络多尺度语义表达和定位能力,最终输出 3 种尺寸、通道分别为 80×80×256、40×40×512 和 20×20×1 024 的特征图。

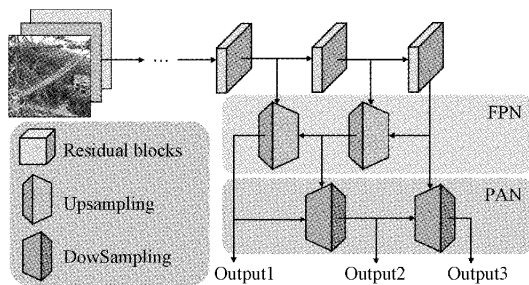


图 3 特征融合网络

1.4 预测

预测部分主要延用 YOLOv3 的方式对输出的 3 种特征图进行解码以及非极大值抑制等操作从而得到最终的目标框检测结果。与 YOLOv3 不同的是,YOLOX 对检测头进行解耦,分别对不同特征图分类、回归,更有效地提取目标信息,采用无锚框设计加快检测速度,引入 SimOTA^[26] 动态地匹配正样本,在不影响模型精度的情况

下使得模型更加轻量。

2 基于 YOLOX 改进的金具检测方法

YOLOX 模型主要以通用数据集中的常见物体为目标进行训练与检测。本文提出的检测方法是在 YOLOX 模型的基础上根据部分金具的小目标特点以及金具检测所需要的精确性和实时性进一步改进的。

2.1 基于图像增广的数据集预处理

选取输电线路中悬垂线夹、U 型挂环、均压环、防震锤、间隔棒等 10 类常见金具制作数据集。金具航拍图像由某输电检修公司提供,为适应 YOLOX 网络的数据集格式和输电线路巡检标准,在构建数据集时采用 VOC 数据集标注格式,并参考中国电力企业联合会标准制定的《架空输电线路设备缺陷影像标注规范》进行标注^[27]。

在目标检测任务中,数据集中图片的数量往往会影响模型训练效果,数据集图片数量较少可能会使模型在训练中无法找到最优解进而使得检测性能无法满足要求。同时,由于数据集的制作依赖于人工对图像中目标的标注,制作大量的数据集会耗费大量的时间。因此,通过图像增广的方式对数据集中图片及其标签同时扩充,将大大减轻人工标注耗时耗力的问题,减轻模型对某些属性的敏感性,提高模型的泛化能力^[28]。

通过对图像采用上下翻转、镜像、改变亮度、仿射变换以及添加高斯噪声的方式对原始图像进行增广操作^[29-30]。对于每张图片都会随机进行上述五种处理以保证扩充后数据集的多样性。当图像尺寸发生变化时,其周围出现的多余区域用 RGB 值均为 0 的黑色进行补充以最小化增广操作对后续特征提取造成的影响,并且,与图像对应的标签信息也将同步改变以适应扩充后的图像。

在增强操作中,翻转、镜像及亮度变化为基础的几何变换,仿射变换包含了图像的平移以及尺寸缩放变换,为提升模型训练泛化性,在操作中还引入高斯模糊核处理,该方法是一种将高斯分布用于图像处理的数据平滑技术。将高斯分布运用到图像的模糊处理时,首先将图片中的像素按中心点划分,指定越中心的像素获得更高的模糊权重,而离中心更远的像素获得更低的模糊权重,而模糊权重由高斯函数进行计算。高斯一维分布函数如式(1)所示。

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

其中, μ 为 x 的均值, σ 是 x 的方差,又称模糊半径。在计算均值时,中心点即为原点,故 $\mu=0$,由此可推导二位高斯分布函数如式(2)所示。

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

其中, $G(x, y)$ 即为高斯模糊核的权重计算方法,本文

设定模糊半径为 0~3 中的随机值对不同的图像进行随机模糊处理。将原始图像与模糊核进行卷积,得到经过高斯模糊处理的图像,在训练时使用该图像从而提高模型的泛化能力。

图 4 展示了金具图像的增广效果。数据集中原共有 614 张金具图片,使用 Labelimg 工具对选取的 10 类金具目标进行标注,生成与图像文件一一对应的 xml 标签文件,经过数据集增强后扩充为 3 070 张图片及其对应标签。为增强模型的泛化能力,数据集中还增加了部分没有对目标进行标注的图像作为负样本。

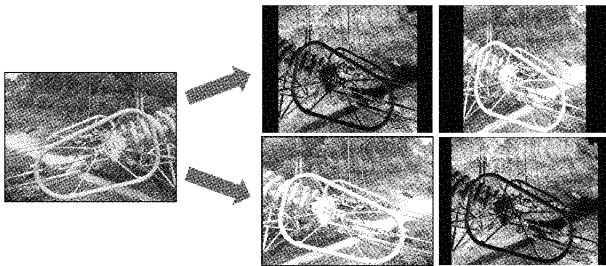


图 4 金具图像的增广

2.2 基于在线增强的骨干网络

为提高对复杂背景下小目标金具的检测能力,骨干网络中采用 Mosaic 和 Mixup 两种在线增强方法进行训练,每当数据集图像输入模型时,都会对图像执行上述两种增强方法,从而实现训练过程中的在线更新增强策略以提升模型鲁棒性,采用在线增强后的骨干网络结构如图 5 所示。

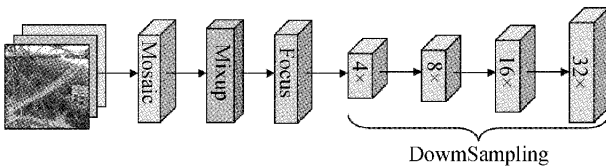


图 5 基于在线增强的骨干网络

如图 6 所示,Mosaic 方法在数据集中随机取出 4 张图片,并将 4 张图片及其标签同时组合,得到一张新的带标签图像。将该图像输入 CNN 中训练相当于同时将输入 4 张图片进行训练,能够在丰富训练集图像背景信息的同时加快训练速度。

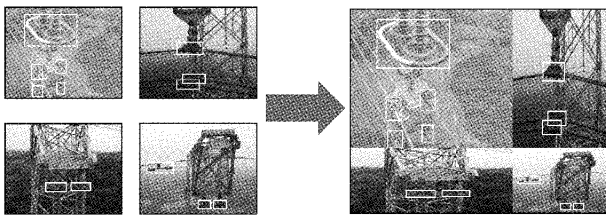


图 6 Mosaic 数据增强

Mixup 通过逐像素线性相加混合两张图片,同时对标签进行混合,起到标签平滑、增强鲁棒性的作用,并且能够

在几乎不增加模型运算量的情况下稳定提高 1% 的目标检测精度^[31]。其原理可由式(3)~(4)表示。

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \tag{3}$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \tag{4}$$

其中, $(x_i, y_i), (x_j, y_j)$ 为训练集中随机选取的两个图像样本及其对应标签,为经过融合之后的图像及其对应标签, λ 为一个服从 Beta 分布的参数, $\lambda \in Beta(\alpha, \alpha)$,其中 $\in [0, +\infty]$ 。图 7 展示了两张金具图像按照不同权重经 Mixup 处理前后的效果。

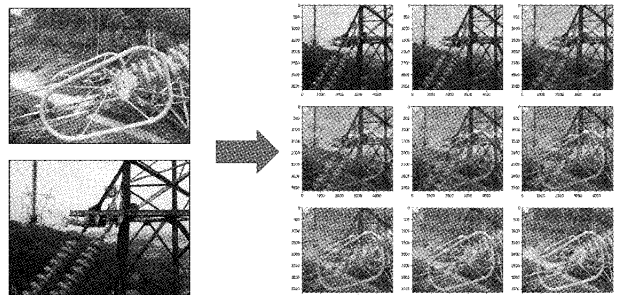


图 7 Mixup 数据增强

2.3 基于特征提取的迁移学习训练策略

在深度学习中,训练 CNN 往往需要大量的样本并耗费大量的时间。在训练过程中,模型将会逐步对图像中目标的整体、局部、边缘、细节等方面进行学习,不断调整寻找最优的网络参数。深度网络中存在大量参数,数据集中也存在大量图片需要进行学习,模型若从 0 开始训练,模型权重将过于随机,训练难以收敛,耗费大量时间且特征提取效果不明显。

此处引入基于特征提取的迁移学习策略进行模型训练,该方法采用训练成熟的 CNN 为基础进行目标特征的提取,在此基础上利用较小的数据集少量调整网络参数得到适用于该数据集的模型,从而提升模型泛化能力,加速模型收敛^[32]。基于此,本文将已有的大型通用数据集上训练得到的骨干网络参数作为金具检测模型的预训练权重进行训练,实现骨干网络提取能力的迁移,虽然数据集不同,但目标的特征通用,该方式能够实现特征与参数的共享,降低训练成本并获得更好的训练效果。

如图 8,采用迁移学习策略的训练过程分为冻结和解冻两个阶段。在冻结阶段,使用预训练权重初始化骨干网络参数并将其冻结,其他网络参数则进行随机进行初始化,从而在模型训练的开始就能提取到更加有效的目标特征,同时对主干以外的网络参数进行寻优调整,相当于调整模型的所有参数。在冻结训练后,模型已经充分学习了通用数据集的特征,此时,对模型的骨干网络参数解冻后继续训练,此过程将调整模型的所有参数,使模型充分学习金具图像特征,最终得到最优模型。

2.4 余弦退火学习率

在深度学习中,学习率(learning rate, LR)的设置影响

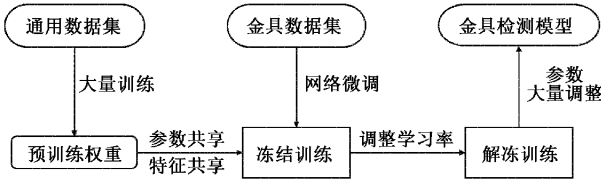


图8 迁移学习策略

着模型的训练速度及训练效果,设定合理的学习率能够使模型更加快速、高效地找到局部最优解。本文模型在冻结训练和解冻训练时均采用余弦退火(cosine annealing)^[33]控制学习率替代原模型中采用的等间隔调整学习率(step LR)帮助模型寻找最优解。

等间隔调整学习率在模型训练开始阶段采用较大学习率快速调整网络参数,在训练过程中按照相同倍数逐渐减小学习率使得模型更好的收敛,但此方式容易让网络陷入局部最优解,达不到网络训练的最好效果。采用余弦退火的方式能够控制学习率以类似余弦函数的形式进行周期性变化,在对学习率的最大值、最小值及初始值设定后,学习率先从初始值以余弦方式下降到最小值,网络开始寻找一个局部最优解并向其靠拢,然后学习率由最小值以余弦方式上升到最大值,此时网络开始寻找下一个局部最优解,在训练中重复该过程从而帮助模型找到全局最优解。

图9以冻结训练(前100次迭代)为例,展示了训练过程中余弦退火和等间隔调整两种学习率的变化趋势。此处设定余弦退火学习率在每个训练批次后略微降低,当减小到 1×10^{-5} 后学习率再缓慢增大到初始设定值,每10次迭代将会循环该过程。而间隔调整学习率在每次迭代之后将会衰减为原来的0.92倍,直到训练结束。在解冻训练中,学习率将会重新初始化,然后以相同的方式进行变化。

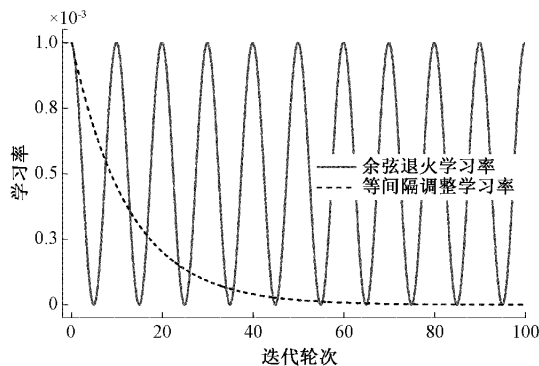


图9 不同学习率随迭代次数的变化趋势

3 实验与分析

3.1 实验介绍

1) 实验数据与平台

在实验过程中采用改进的YOLOX模型进行训练与测试,数据集图片为使用大疆Phantom 4 Pro无人机航拍

的614张线路金具图片,单张图片分辨率为 $4\ 864 \times 3\ 468$ 。将该数据集图片经过2.1节所述预处理后得到3 070张图片,将其按照9:1:1的比例划分为训练集、验证集和测试集。

实验所用软硬件配置如下:CPU为AMD Ryzen 7 4800H,GPU为NVIDIA GTX 1660Ti;采用Pytorch1.2深度学习框架,使用CUDA11.2为GPU加速库,在Python3.6环境下进行实验。

2) 网络训练

使用数据集种划分好训练集和验证集进行网络训练,设定冻结阶段每次传入模型的迭代量为16,解冻阶段设为2,冻结阶段迭代次数为100轮,解冻阶段为200轮,共300轮。设定冻结训练阶段与解冻训练阶段的初始学习率分别为 1×10^{-3} 和 1×10^{-4} ,训练和测试中网络的初始化图像尺寸为 640×640 。

损失函数(loss function)用于衡量模型的预测值与真实值的误差大小,损失函数越小则预测值越接近于真实值,模型的训练效果越好,当损失函数趋于不变时模型趋于收敛。训练过程的迭代损失曲线如图10所示,在模型冻结训练阶段,损失函数刚开始迅速下降,当迭代到80轮后趋于平缓;在模型解冻训练阶段开始时,损失函数迅速下降,由于采用了余弦退火算法调整学习率,故损失函数在收敛过程中不断波动,当迭代270轮后,损失函数虽仍在某范围内波动,但该范围已经几乎不变,模型已收敛。取迭代过程中总损失最低的权值作为模型参数对测试集图像进行测试。

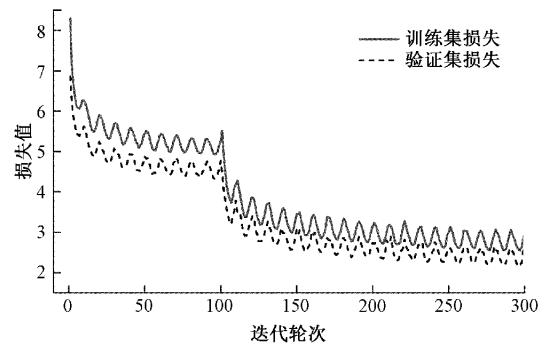


图10 训练过程中迭代损失曲线

3) 实验评价指标

针对各金具目标识别的性能评估,采用精确率(precision)、召回率(recall),平均精度(average precision, AP)、平均精度均值(mean average precision, mAP)、F1值(精确率和召回率的加权调和平均)等对金具检测的性能进行衡量^[34-35],如式(5)~(9)所示。

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \int_0^1 p(r)dr \tag{7}$$

$$mAP = \frac{\sum AP}{N(Class)} \tag{8}$$

$$F1 = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \tag{9}$$

其中, TP 为检测正确的正样本, FP 为检测错误的正样本, FN 为检测错误的负样本, N 为类别总数, 式(8)中的 p 与 r 分别为精确率与召回率, AP 表示 PR 曲线所占的面积。在进行本文实验时, 均采用式(5)~(9)计算相应指标进行检测方法的性能评价。

3.2 实验结果与分析

为提高小目标金具的检测精度, 同时保证金具检测的速度和精度, 本文提出改进的 YOLOX 检测方法, 通过采用图像增广预处理数据集, Mosaic, Mixup 在线增强, 以及引入迁移学习和余弦退火学习率训练策略的方式实现对金具的实时检测。为此, 本节设计了不同改进策略的消融实验评价改进的有效性, 详细展示了使用改进方法对各类金具检测的结果, 并设计了所提方法与其他主流识别方法的对比实验以验证其性能。

1) 不同改进策略的消融实验

为了更好地研究 3 种改进方式的有效性, 使用 614 张原始图片为原始数据集, 设置 4 组对照进行消融实验, 得到模型精度对比结果如表 1 所示。其中, 图像增广预处理

均扩充到 3 070 张图片进行实验, 其他实验环境与设置均相同, 表 3 中的“√”表示该实验组包含了表头所列项目。可以看出, 采用图像增广预处理之后 mAP 大幅提升, 该策略使用更少的数据集达到了更大数据集的训练效果, 相当于减少了人工标注的工作量。在增加 Mosaic, Mixup 在线增强以及余弦退火学习率两种改进策略后, 模型 mAP 也均有提升, 说明前者增加了模型的鲁棒性, 对模型精度提升有所帮助, 后者帮助模型跳出了局部最优解, 找到了更优模型, 故也提升了模型精度。最终模型对比原模型提升达到 18.32%, 所提改进均能有效提升模型检测性能。

表 1 不同改进策略的消融实验

原模型	图像增广 预处理	Mosaic, Mixup 在线增强	余弦退火 学习率	mAP/ %
√				68.15
√	√			80.09
√	√	√		82.79
√	√	√	√	86.47

2) 改进模型检测结果

使用所提的各改进方法进行训练, 将得到的最优权重文件载入模型, 对测试集中的金具图像进行识别, 得到实验结果, 图 11 展示了部分金具的识别效果, 表 2 给出了模型对不同金具的检测性能结果。

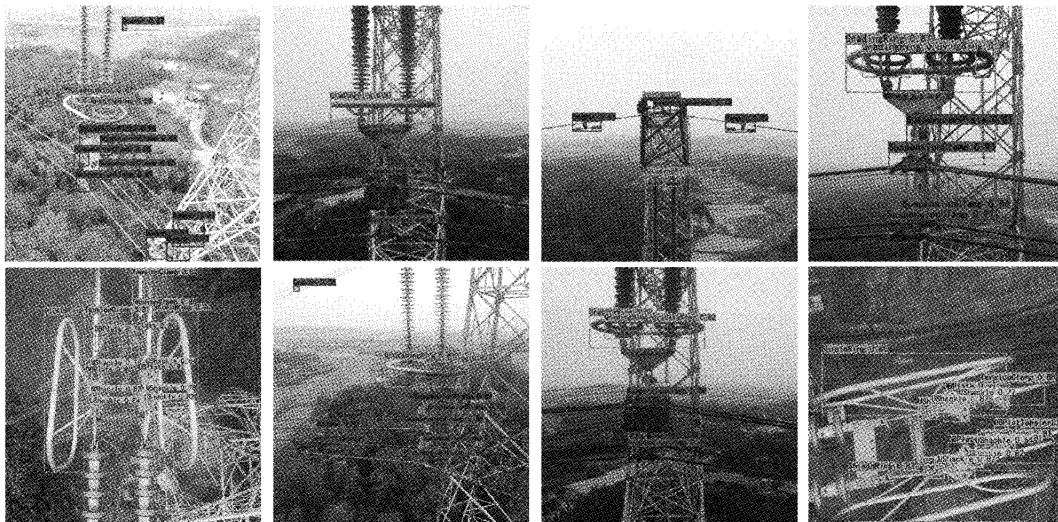


图 11 部分金具检测效果

从表 2 可以看出, 多数金具的 AP 值均高于 83%, 所有金具的识别精确率都在 87% 以上, 说明本文方法的误检率较低, 检测到的金具几乎均为正样本, 在召回率上有个别金具较低, 说明该类金具仍存在部分正样本未被检测到, 推测可能是由于该部分样本个数较少导致。除防震锤以外, 各金具的 $F1$ 值都在 0.8 以上, 说明大部分金具的精确率和召回率都能达到很好的效果, 防震锤的召回率较低,

可能是由于防震锤出现的场景多位于远处背景, 尺寸过小且比较模糊, 另外一个可能的原因是其在数据集中样本量不足。

3) 不同改进策略的消融实验

表 3 给出了在相同数据集、实验平台以及同样采用迁移学习的情况下, 本文方法与目前主流检测模型的对比。相比两段检测方法 Faster R-CNN, 本文方法精度更高, 速

表2 不同金具的检测结果

金具	AP/%	Precision/%	Recall/%	F1 值
提包式悬垂线夹	85.92	92.68	75.81	0.83
压缩型耐张线夹	76.70	96.34	69.30	0.81
U型挂环	88.85	93.13	83.44	0.88
均压环	99.29	98.36	99.17	0.99
屏蔽环	96.11	89.63	96.74	0.93
DB型调整板	85.11	95.38	81.85	0.88
悬垂联板	92.06	90.00	85.71	0.88
平行挂板	83.84	96.46	75.96	0.85
防震锤	71.24	87.91	52.63	0.66
间隔棒	85.57	89.68	76.87	0.83

度更快,与单段检测方法 SSD、YOLOv3、Efficientdet、Retinanet 相比,mAP 分别高出 40.42%、9.62%、40.26%、30.05%,且 F1 值也最高,说明综合考虑召回率、精确率和平均精度均值,本文方法表现更好。在检测时间上,相比目前的主流识别方法,本文方法对单张分辨率为 $4\ 864 \times 3\ 468$ 图片的检测时间为 45.90 ms,可见本文方法能够在高精度的同时保证快速性,在金具实时检测中有着更好的性能表现。

表3 不同模型的对比结果

检测模型	mAP/ %	F1 值 (各类别平均)	单张图片平均 检测时间/ms
Faster R-CNN	39.64	0.42	135.26
SSD	46.05	0.34	24.06
YOLOv3	76.85	0.78	45.23
Efficientdet	46.22	0.51	161.29
Retinanet	56.42	0.55	74.24
本文方法	86.47	0.85	45.90

4 结 论

基于新一代实时检测网络 YOLOX 提出一种多目标金具实时检测方法。与现有方法相比,YOLOX 检测方法兼具高精度与快速性特点。提出图像增广预处理方法,解决金具检测模型训练中样本量少,人工标注数据集工作量大的问题,在骨干网络采用 Mosaic、Mixup 在线增强方法,解决复杂背景下小目标金具难以检测的问题,使用迁移学习提升模型鲁棒性,增加余弦退火学习率控制方式帮助模型收敛。实验表明,所提改进对模型精度均有较大提升,本文方法检测平均精度均值达到 86.47%,且对 U 型挂环等复杂背景下的小目标检测效果好,同时,单张图片检测时间仅为 45.90 ms,能够在节省人力的情况下满足输电线路金具巡检的精度和速度要求。未来,可进一步使用更大数据集以及在更低分辨率图像上进行测试,不断提高检测

方法精度与鲁棒性。

参考文献

- [1] 周华良,宋斌,安林,等.特高压输电线路分布式故障诊断系统研制及其关键技术[J].电力系统保护与控制,2019,47(24):115-122.
- [2] YE H, STRUNZ K. Multi-scale and frequency-dependent modeling of electric power transmission lines[J]. IEEE Transactions on Power Delivery,2018,33(1):32-41.
- [3] 江秀臣,刘亚东,傅晓飞,等.输配电设备泛在电力物联网建设思路与发展趋势[J].高电压技术,2019,45(5):1345-1351.
- [4] 秦科技.智能高压输电线路巡检机器人设计[D].南昌:南昌大学,2017.
- [5] 赵振兵,齐鸿雨,聂礼强.基于深度学习的输电线路视觉检测研究综述[J].广东电力,2019,32(9):11-23.
- [6] 丁志敏,邢晓敏,董行,等.基于深度学习的输电线挂接地线状态目标检测[J].电子测量技术,2021,44(3):132-137.
- [7] 解尧婷,张丕状.基于改进的 YOLOv4 输电线路小目标检测[J].国外电子测量技术,2021,40(2):47-51.
- [8] 彭继慎,孙礼鑫,王凯,等.基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J].仪器仪表学报,2021,42(10):161-170.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2014:580-587.
- [10] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision (ICCV), 2015:1440-1448.
- [11] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39 (6): 1137-1149.
- [12] 白洁音,赵瑞,谷丰强,等.多目标检测和故障识别图像处理[J].高电压技术,2019,45(11):3504-3511.
- [13] 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,2020,334(9):3018-3026.
- [14] 邢姗姗,赵文龙.基于 YOLO 系列算法的复杂场景下无人机目标检测研究综述[J].计算机应用研究,2020,37(S2):28-30.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. Proceedings of the European Conference on Computer Vision,2016:21-37.

- [16] 张永翔,吴功平,刘中云.等.基于YOLOv3网络的输电线路防震锤和线夹检测迁移学习[J].计算机应用,2020,40(S2):188-194.
- [17] 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,2020,84(4):656-662.
- [18] GE Z, LIU S, WANG F, et al. YOLOx: Exceeding YOLO series in 2021 [J]. ArXiv preprint, 2021, ArXiv:2107.08430.
- [19] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C]. European Conference on Computer Vision Springer, 2014:740-755.
- [20] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:7263-7271.
- [21] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]. International Conference on Machine Learning. PMLR, 2015:448-456.
- [22] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020:390-391.
- [23] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1904-1916.
- [24] LEE C Y, GALLAGHER P W, TU Z. Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree [C]. Artificial Intelligence and Statistics. PMLR, 2016:464-472.
- [25] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [26] GE Z, LIU S, LI Z, et al. OTA: Optimal transport assignment for object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 303-312.
- [27] HSU W Y, LIN W Y. Adaptive fusion of multi-scale YOLO for pedestrian detection [J]. IEEE Access, 2021(9):110063-110073.
- [28] ENGSTROM L, TRAN B, TSIPRAS D, et al. A rotation and a translation suffice: Fooling cnns with simple transformations [J]. ArXiv Preprint, 2018, ArXiv:1712.02779.
- [29] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning [J]. Journal of Big Data, 2019, 6(1):1-48.
- [30] HE H, SIU W C. Single image super-resolution using Gaussian process regression [C]. CVPR 2011, IEEE, 2011:449-456.
- [31] ZHANG H, CISSE M, DAUPJIN Y N, et al. Mixup: Beyond empirical risk minimization [C]. Proceedings of the International Conference on Learning Representations, 2017.
- [32] MISRA S, JEON S, LEE S, et al. Multi-channel transfer learning of chest X-ray images for screening of COVID-19 [J]. Electronics, 2020, 9(9):1388.
- [33] LOSHCHILOV I, HUTTER F. Sgdr: Stochastic gradient descent with warm restarts [J]. ArXiv preprint, 2016, ArXiv: 1608.03983.
- [34] EVERINGHAM M, ESLAMI S, GOOL L V, et al. The pascal visual object classes challenge [J]. International Journal of Computer Vision, 2010, 88(2):303-338.
- [35] 管军霖,智鑫.基于YOLOv4卷积神经网络的口罩佩戴检测方法[J].现代信息科技,2020,4(11):9-12.

作者简介

黄力,工学博士,讲师,主要研究方向为输电线路智能检测。

E-mail:huangli@ctgu.edu.cn

万旭东(通信作者),硕士研究生,主要研究方向为输电线路智能检测。

E-mail:416208625@qq.com

王凌云,工学博士,副教授,主要研究方向为电力系统运行优化与控制。

E-mail:wly@ctgu.edu.cn

刘兰兰,博士,高级工程师,主要研究方向为智能输电运检。

E-mail:0603050307@163.com