

DOI:10.19651/j.cnki.emt.2212004

## 改进 YOLOv5 的绝缘子旋转目标检测\*

张珂 余鲲鹏 石怀涛

(沈阳建筑大学机械工程学院 沈阳 110000)

**摘要:**为实现输电线路绝缘子的精确定位与快速识别,提出了一种基于 YOLOv5 的旋转目标检测算法。针对绝缘子自爆故障,提出了一种三标签的旋转矩形框标注方法,能够提升模型的检测效果和收敛速度;引入 Hardswish 激活函数加速计算,同时提高模型的数值稳定性;用 Ghost module 代替 Conv 结构,以此来降低模型的参数量,提升模型的检测效率;将位置损失使用的 CIOU Loss 函数改进为 EIOU Loss 函数,添加一个角度损失函数 SLL,用于矩形框角度的回归,使模型能更精准的定位绝缘子的位置,提高模型的检测效果。实验表明,改进的绝缘子故障检测算法相较于 YOLOv5s 模型的浮点计算量减小了 48.7%,模型大小降低了 44.5%,推理速度提升 2.9%,模型的整体检测精度能达到 97.7%,同时还能满足移动端部署的实时性要求。

**关键词:** 精确定位;快速检测;主干网络轻量化;损失函数;旋转目标检测

**中图分类号:** TP   **文献标识码:** A   **国家标准学科分类代码:** 520.6040

## Rotating insulator detection based on improved YOLOv5

Zhang Ke Yu Kunwu Shi huaitao

(School of Mechanical Engineering, Shenyang Jianzhu University, Shenyang 110000, China)

**Abstract:** To realize the precise positioning and quick recognition of insulators on a line that transmits electricity. A rotating target detection algorithm was proposed based on YOLOv5 (You Only Look Once v5). A rotating rectangular frame containing three types of labels was proposed for self-exploding insulators. It can improve the detection effect of the model and the convergence rate of the model. Use a Hardswish activation function to accelerate the computation and improve the numerical stability of the model. Replace the Conv structure with the Ghost module structure, to reduce the number of parameters in the model and improve the detection efficiency of the model. Replace the CIOU Loss (Complete IoU loss) function with the EIOU Loss (Efficient IoU Loss) function and add a SLL (Smooth L1 Loss) about angle for regression of rectangular box angle. It can position the insulator more accurately and improve the detection effect of the model. The experimental results show that the improved insulator fault detection algorithm reduce the calculation amount of YOLOv5s model by 48.7% and the model size by 44.5%, Inference speed increase 2.9%, The overall detection accuracy of the model is 97.7%. In addition, it can also meet the real-time requirements of mobile deployment

**Keywords:** precise localization; quick recognition; backbone network lightweight; loss function; rotating target detection

## 0 引言

绝缘子作为电力系统的重要组成部分,对输电线路的安全稳定运行具有十分重要的作用。由于其长期工作在户外,会不可避免地产生故障或损坏。常见的故障种类有绝缘子自爆、绝缘子污秽和绝缘子裂缝等。传统的故障检测方法主要通过巡检人员登上塔杆人工巡检,这种巡检方法不仅辛苦和耗时,并且存在一定的危险性。随着近年来人工智能和深度学习的快速发展,搭载目标检测算法的无人机巡检逐渐进入人们的视线。

目前主流的目标检测算法可以分为两类,一类是通过特征提取来生成候选区域,然后进行分类的两阶段方法,其中的代表算法主要包括有空间金字塔池化网络<sup>[1]</sup> (spatial pyramid pooling networks, Spp-Net)、区域卷积神经网络<sup>[2]</sup> (regions convolutional neural networks, R-CNN)、Fast-CNN、Faster-CNN 等,另一类是端到端的一阶段方法,其中的代表算法是单步多框目标检测<sup>[3]</sup> (single shot MultiBox detector, SSD)和你只需看一次<sup>[4]</sup> (you only look once, YOLO)系列。这两类方法在当前的目标检测领域都

收稿日期:2022-11-09

\* 基金项目:辽宁省科技重大专项(2020JH1/10100012)资助

有所应用。杨焰飞等<sup>[5]</sup>通过对基础的YOLOv3模型增加InceptionResnet模块、改进损失函数和利用K-means算法进行锚点框选择。使模型在保持检测速度的同时还能保持较高的识别率。针对航拍图像多尺度、排布密集的特点,Azimi等<sup>[6]</sup>在FPN<sup>[7]</sup>的基础上提出了图像级联网络(image cascade network, ICN),其通过FPN与图像金字塔的结合,能够获得包含更多细节信息的多尺度检测层从而提高航拍图像的小目标检测效果。

虽然以上研究在取得了相应进展,但都有各自的局限性,在工业部署中的效果并不理想。例如基于区域提取的Two stage方法大多有较好的检测效果,但模型臃肿,参数量过大,难以部署于移动端或嵌入式设备。而基于位置回归的One stage方法则大多难以满足识别准确率的要求,此外它们大多无法实现目标的精确定位。针对上述情况,本文提出了一种基于YOLOv5的旋转目标检测算法,通过提出一种针对绝缘子自爆故障的三标签旋转标注方法,改进模型的主干网络以及损失函数,能够实现目标的精确定位,并且在保持其较高检测精度的同时还能满足实时性的要求,能够满足工程实际的需要。

## 1 改进YOLOv5算法

### 1.1 YOLOv5算法原理

YOLOv5的网络结构包括4个部分,分别为Input(输入端)、Backbone(骨干网络)、Neck(多尺度特征融合模块)、Prediction(预测端)<sup>[8]</sup>。输入端进行的工作主要包括数据增强、自适应锚框计算、自适应图片缩放。骨干网络主要用来聚合,提取3种尺度的特征图。和之前的YOLOv4<sup>[9]</sup>一样都采用跨级部分网络<sup>[10]</sup>(cross stage partial network, CSP)结构,但值得注意的是骨干网络和多尺度融合模块的BottleNeck结构并不相同。多尺度融合模块的作用就是对不同尺度的特征进行特征融合,多尺度特征融合模块的金字塔注意力网络<sup>[11]</sup>(pyramid attention network, PAN)结构中加入了CSP,作用是减少网络的计算量以及对显存的占用,同时保证网络的能力不变或者略微提升。YOLOv5的输出端包括损失计算及非极大值抑制(non maximum sup-pression, NMS)。YOLOv5的损失计算包括类别损失(classes loss)、置信度损失(objectness loss)和位置损失(location loss),其中类别损失和置信度损失采用的是BCE Loss(binary cross entropy),而在位置损失处采用的是CIoU loss函数。CIoU loss函数综合考量了预测框与真实框的重叠面积,中心点距离和长宽比。非极大值抑制用来筛选目标框,对图像特征预测生成边界框并预测类别。

### 1.2 卷积模块的改进

提出的旋转矩形框标注方法虽然能够更精准的检测出目标,但由于其引入了矩形框旋转角度这一额外参数,模型的计算量会有所增加,推理速度会下降。为保证模型推理的

实时性要求,将卷积模块的卷积批处理标准化(convolution batch normalization SiLU, CBS)结构进行改进,提出了一种GBH的轻量化结构,由Ghost module+Batch normalization+Hard swish组成,GBH的结构如图1所示。

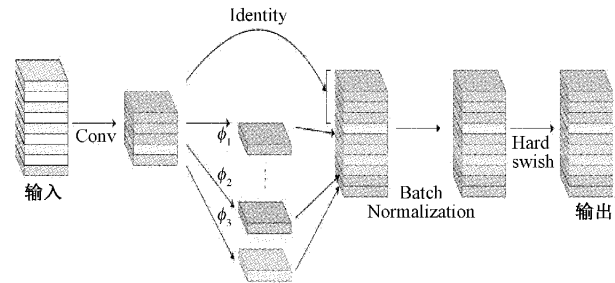


图1 GBH结构

Ghost卷积来自于GhostNet<sup>[12]</sup>,Ghost卷积将普通的卷积分为两部分,首先进行一个 $1 \times 1$ 的卷积,这个卷积的通道数是普通卷积的一半,用以整合特征,生成输入特征层的特征浓缩。然后进行深度可分离卷积,这个深度可分离卷积是逐层卷积,用以生成特征浓缩的相似特征图,将相似特征图与特征浓缩图拼接即可得到与普通卷积相同结构的输出特征图。在一个优秀的CNN模型中,特征图存在冗余是非常重要的,因为这有助于得到良好的训练结果。但这无疑会增加模型的训练量,而通过使用一些计算量更低的去生成这些冗余的特征图是一个很好的解决思路。因此通过将YOLOv5模型结构中的CBS模块改进为GBH模块,能在保持模型精度的前提下大幅减小模型的参数量。从而降低模型训练所需要的配置要求,进而便于模型在移动端和嵌入式设备的部署。

Hard swish激活函数在MobileNetV3<sup>[13]</sup>架构中被提出,相较于YOLOv5使用的SiLU函数,具有数值稳定性好,计算速度快等优点。尽管SiLU函数提高了检测精度,但由于其求导较为复杂,量化计算较慢,因此会增加部署的计算成本,而Hardswish函数的计算却简单很多,而在准确性上却几乎没有带来损失,Hardswish函数曲线如图2所示。

$$\text{Hardswish}(x) = \begin{cases} 0, & x \leq -3 \\ x, & x \geq 3 \\ x \cdot (x + 3) / 6, & -3 < x < 3 \end{cases} \quad (1)$$

### 1.3 CSP结构的改进

YOLOv5s的CSP结构是将输入分成两个分支,分别进行卷积操作使得通道数减半,其中一个分支额外进行Bottleneck $\times N$ 的操作,然后对两个分支进行拼接,使得输入输出的数据结构相同,这样能使模型学习到更多的特征,YOLOv5中存在两种CSP结构,它们分别存在于Backbone处的C3结构和Neck处的C3结构中,两者的区别主要体现在BottleNeck上,特征提取部分的BottleNeck由经过两个CBS结构和初始输入通过残差结

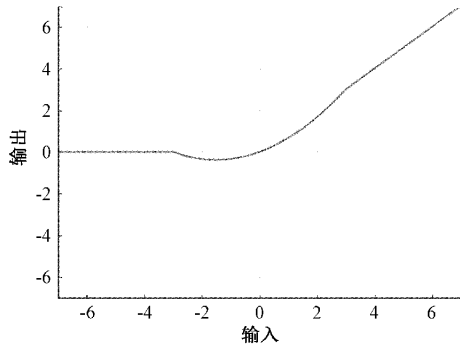


图 2 Hardswish 激活函数

构相加而成,而特征融合部分的 Bottleneck 只线性连接两个 CBS 结构。这样的设计能减少网络的计算量以及对显存的占用,同时保证网络的能力不变或者略微提升。为了进一步减少参数量和优化计算,将 C3 结构中的 CBS 结构换成 GBH 结构,改进后的 C3Ghost 结构如图 3 所示。

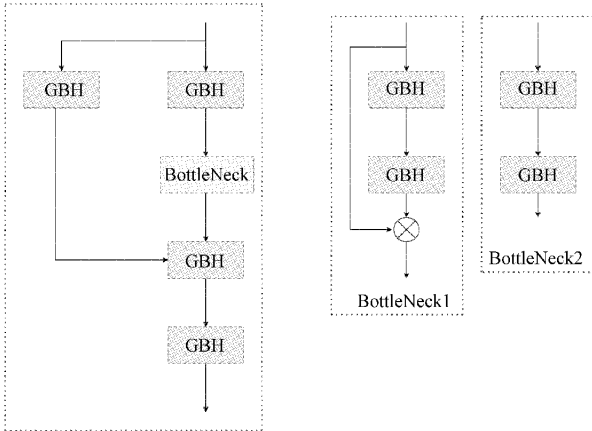


图 3 改进的 C3Ghost 结构

### 1.4 损失函数的改进

YOLOv5 的类别损失和置信度损失均采用二值交叉熵损失函数,其中类别损失只计算正样本的损失,而置信度损失计算的是所有样本的损失,这两类损失计算都比较简单,并且表现效果也很不错。而一个好的目标框回归函数应该综合考虑长宽比、重叠面积和目标框和预测框之间的中心点距离。YOLOv5 的位置损失函数采用的是 CIOU loss<sup>[14]</sup>。该损失函数的计算公式如下。

$$CIOU = IOU - \left[ \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \right] \quad (2)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

$$\alpha = \frac{v}{(1 - IOU)} + v \quad (4)$$

$$CIOU_{Loss} = 1 - CIOU \quad (5)$$

上式中的 IOU 为预测框和真实框的交并比,IOU 通过计算预测框与真实框之间的交集除以预测框与真实框之间的并集得来,中括号左半部分部分是真实框和预测框中心

点距离的平方与两者最小框对角线长度平方的比值,右半部分中的  $v$  为预测框和真实框长宽比例差值的归一化,  $\alpha$  为权衡长宽比例和 IOU 部分造成损失的平衡因子。虽然 CIOU 在 DIOU 的基础上通过添加预测框和真实框综合考量了长宽比的影响,但还存在着两方面的问题。首先 CIOU 使用的是高宽的相对比例,一旦预测框和真实框的宽高纵横比呈现线性比例时,CIOU 中添加的相对比例的惩罚项便不再起作用,另外预测框的宽高无法同时进行缩放,由梯度公式可知,其中一个参数增大,另一参数必定减小。针对这种问题,EIOU<sup>[15]</sup> 提出了直接对高宽的预测值进行惩罚计算的损失函数,EIOU 的计算公式如下:

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (6)$$

从上式可以看出,虽然 EIOU 同样是由长宽比、重叠面积和中心点距离组成,但其在计算长宽比时是直接将边长作为惩罚项的,这可以应对 CIOU 中的两个问题,加速模型的收敛。

YOLOv5 的 Prediction 结构的检测层通道维度为  $3x(C+5)$ ,通过新增一个旋转角度预测通道即可实现旋转矩形框的预测,这个新增的通道用来预测角度回归参数  $\theta$ ,通过 Smooth L1 Loss 损失函数进行优化,Smooth L1 Loss 的优点是当预测框与真实框的差别很小时,梯度值足够小,当预测框与真实框的差别过大时,梯度又不至于过大,训练更加稳定,Smooth L1 Loss 的计算公式如下:

$$smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (7)$$

训练时使用 EIoU Loss 用于优化旋转矩形框的位置和形状,Smooth L1 Loss 则用于优化该矩形框的旋转角度,推理时将两者进行结合便可获得位置信息的完整预测结果,最终总体的损失函数如下:

$$Loss = \lambda_1 L_{cls} + \lambda_2 L_{obj} + \lambda_3 L_{loc} + \lambda_4 L_{theta} \quad (8)$$

### 1.5 最终模型

将前面所使用的包括改进模块、激活函数、损失函数等在内的所有改进策略添加进 YOLOv5 基模型中,最终改进模型的模型结构图如图 4 所示。

## 2 数据集构建

实验使用的数据集来自中国电力线路绝缘子数据集 (CPLID),数据集由 248 张自爆故障绝缘子图片和 600 张正常绝缘子图片组成,故障图片为通过图像分割技术合成的图片,在目标检测中,模型训练使用的训练集越大,泛化能力就越强,训练后模型的实用性就越好。由于原始数据比较少,因此通过上下翻转、左右翻转、裁剪、擦除像素块的方式将数据集进行扩充至 3 400 张,同时添加 50 张不包含检测对象的背景图,这有助于减小将背景错误识别为检测

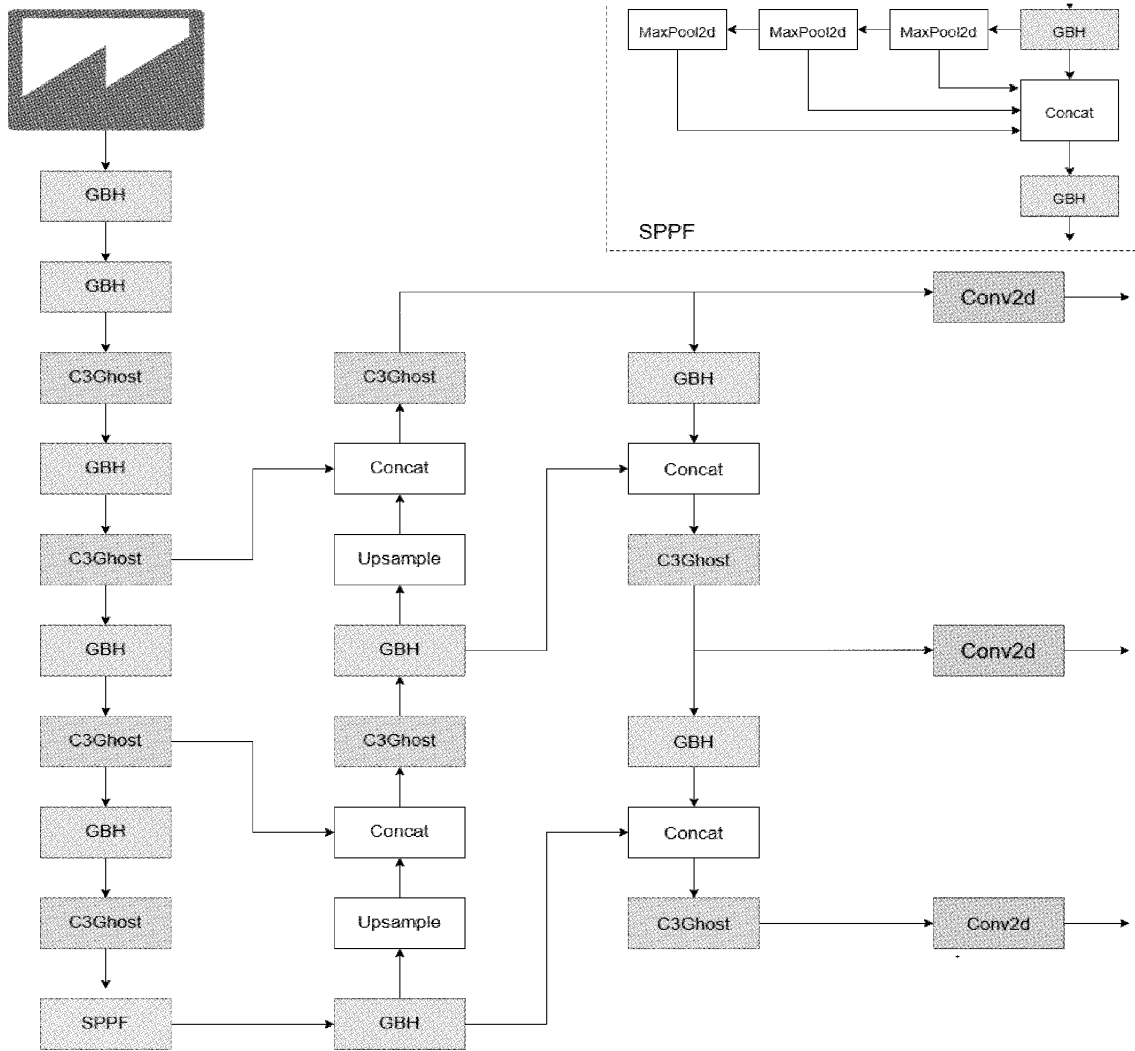


图4 改进后的YOLOv5结构

对象的可能,因此最终数据集为3450张,训练集与验证集按照9:1的比例进行划分。

通过RoLabelImg图片标注工具对数据进行标注,设置三类标签,分别为正常绝缘子normel、故障绝缘子damaged和故障位置zibao,生成的标签格式为xml文件,这种格式的标签无法直接应用于YOLOv5,需要通过开源项目RoLabelImg\_Transform将其转化为txt格式,生成的为矩形框标注法中的八标签表示法。

目前旋转矩形框主要的表示方式有两种,分别为五参数表示法和八参数表示法。五参数表示法在传统矩形框的基础上新增了一个用于描述旋转角度的参数 $\theta$ ,五参数表示法根据取值范围的不同又可分为OpenCV表示法和长边表示法,其中OpenCV表示法的表示形式为 $(x, y, w, h, \theta)$ 其中 $x, y$ 表示的是矩形框中心点的坐标, $w$ 为 $x$ 轴逆时针旋转遇到的第一条边, $w$ 的邻边定义为 $h, \theta$ 表示 $x$ 轴逆时针旋转到 $w$ 所经的角度范围,因为是逆时针旋转所以 $\theta$ 的符号为负, $\theta \in [-\pi/2, 0)$ 。而长边表示法和OpenCV

表示法的区别为 $w$ 定义为最长边, $w$ 的邻边则定义为 $h, \theta$ 表示 $x$ 轴旋转到 $w$ 所经的角度范围, $\theta \in [-\pi/2, \pi/2)$ 。八参数表示法是用四顶点来表示矩形框的,并从左上点开始以顺时针方式排序。本方法最终采用的旋转矩形框标注方法为五参数表示法中的长边表示法,它相对于八参数表示法有更少的参数量,能够保证模型的训练速度,并且它和正矩形框的标注方法更加相似,便于对YOLOv5模型进行结构上的适应性改进。

### 3 实验与结果分析

#### 3.1 实验环境及配置

实验采用的操作系统为Windows10, GPU的型号为GTX 1650,显存大小为4GB,内存大小为16G,训练模型时的批处理尺寸为4,整体配置中等偏下,能够满足模拟移动端性能的条件,为满足移动端视频检测的实时性要求,算法检测的速度应保证大于或等于24帧/s。实验以mAP作为模型检测精度评估指标,以FPS作为速度评估指标,

以模型内存大小和浮点计算量作为模型尺寸评价指标。

### 3.2 标注方式对比

目前常见的针对绝缘子自爆缺陷检测数据标注方法有两种,一种只标注故障位置,另一种分为两类标签,包括故障位置标签和正常标签。这两种标注方式存在的主要问题是绝缘子发生自爆的部位只占据绝缘子的一小部分,在无人机拍摄的照片中属于小目标,自爆位置的检测效果并不好,故障绝缘子整体的检测效果是优于故障位置的。并且由于故障类和自爆类是独立两类标签,因此实际检测中更不容易漏检故障绝缘子,图 5 为各种标注方法在 YOLOv5 模型下的检测效果图。另外通过对各类标签的检测精度进行实验对比,发现故障类这类标签的检测精度是要高于其他两类标签,这对于实际检测故障是很有价值的,各类标签在验证集上的检测对比具体如表 1 所示。

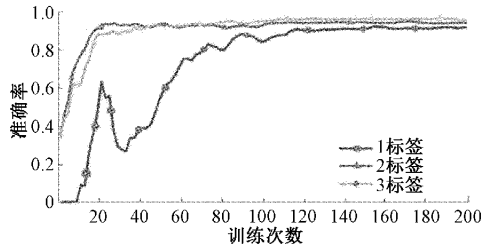


图 5 不同标注方法检测对比图

表 1 不同类别标签检测对比

标签	故障类	自爆类	正常类
准确率	0.995	0.932	0.972

针对旋转矩形框标注方法,将其与采用正矩形框标注的相同数据集进行对比,比较结果如表 2 所示,通过对比可以发现采用旋转矩形框标注的 R-YOLOv5s 相对于 YOLOv5s 虽然在推理速度上损失了 14.6%,但整体精度 mAP 提升了 1.7%,并且训练生成模型的浮点计算量和内存大小几乎没有受到影响。

表 2 不同矩形框标注方法对比

标注方法	mAP <sub>50</sub>	FPS	模型量/M	GFLOPs
正矩形框	0.966	26.7	13.7	15.7
旋转矩形框	0.983	22.8	13.7	15.7

### 3.3 轻量化实验对比

旋转矩形框标注方法虽然更适应无人机航摄绝缘子图像具有的尺度变化大,角度各异的特点,能够更精准的定位目标,但由于其引入了旋转角度这一额外参数,增加了模型推理过程中的计算量,因此需要对模型进行轻量化。为了验证模型主干网络轻量化改进的有效性,以及所引入的 *Hardswish* 激活函数与 *Ghost module* 轻量化模块分别所起的作用,将原始 YOLOv5s 模型与单独添加

*Hardswish* 激活函数、单独添加 *Ghost module* 轻量化模块以及同时添加 *Hardswish* 激活函数与 *Ghost module* 模块的模型做对比实验,对比结果如表 3 所示,对比实验均采用正矩形框标注方法。

表 3 不同矩形框标注方法对比

模型	mAP <sub>50</sub>	FPS	模型量/M	GFLOPs
YOLOv5s	0.966	26.7	13.7	15.7
YOLOv5s-H	0.968	22.6	13.7	15.9
YOLOv5s-G	0.942	34.5	7.5	8.0
YOLOv5s-G-H	0.949	34.1	7.5	8.1

通过实验对比发现,*Ghost module* 模块对于模型轻量化具有十分优秀的效果,添加 *Ghost module* 模块后的算法训练出的模型相较于原 YOLOv5s 训练出的模型体量基本减小了一半,并且模型的浮点计算量也基本减少了一半,但 *Ghost module* 会对模型造 2.4% 的精度损失,通过添加 *Hardswish* 激活函数可以稍稍缓解这一损失。

### 3.4 损失函数实验对比

针对损失函数的改进包括改进位置损失函数和添加额外的角度损失函数,额外引入的角度损失函数最终的影响也会反应到算法的位置损失上,由于旋转矩形框标注算法必须添加角度损失,为了验证各部分改进所起的作用,将采用正矩形框标注的 YOLOv5s 模型、采用正矩形框标注的改进位置损失为 *EIOU Loss* 的 YOLOv5s-E 模型、采用旋转矩形框标注的添加 *Smooth L1 Loss* 的 YOLOv5s-S 模型和采用采用旋转矩形框标注的同时添加 *Smooth L1 Loss* 与 *EIOU Loss* 的 YOLOv5s-E-S 在位置损失上做对比实验,实验结果如图 6 所示。

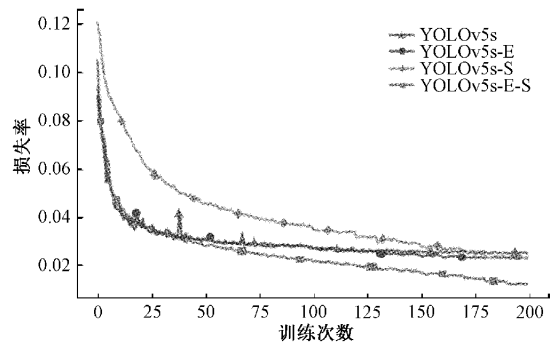


图 6 不同损失函数对比图

通过实验对比发现,相比于正矩形框标注的基线模型 YOLOv5s,改进位置函数为 *EIOU* 的模型位置损失收敛更加的平稳,并且精度也有少量提升,添加角度损失函数 *Smooth L1 Loss* 的模型虽然收敛较慢,但模型平稳,而最终的 YOLOv5s-E-S 模型在精度和平稳性上都要优于初始模型。

### 3.5 消融实验

采用控制变量的方法研究对YOLOv5s模型各项改进所起的作用,所提的三项改进分别为激活函数(*hardswish*, H)、轻量化模块(*ghost module*, G)和ES(*Smooth L1 Loss*与*EIOU Loss*损失函数),为了控制变量,所有的改进模型都

使用三标签标注的旋转矩形框标注数据进行实验。为了直观对照,实验一共包含8组,分别为不添加任何改进的基础YOLOv5s模型、单独添加3种改进策略的3组模型、3种策略两两结合的3组模型以及同时添加3种改进策略的模型。实验结果如表4所示。

表4 消融实验对比

算法模型	H	G	ES	mAP <sub>50</sub>	FPS	模型量/M	GFLOPs
YOLOv5s				0.966	26.7	13.7	15.8
YOLOv5s-H	✓			0.984	22.6	13.7	15.9
YOLOv5s-G		✓		0.964	30.8	7.5	8.0
YOLOv5s-ES			✓	0.985	19.5	13.8	16.0
YOLOv5s-H-ES	✓		✓	0.986	21.7	13.8	16.1
YOLOv5s-G-ES		✓	✓	0.973	28.8	7.5	8.1
YOLOv5s-H-G	✓	✓		0.969	30.6	7.5	8.1
YOLOv5s-H-G-ES	✓	✓	✓	0.977	27.5	7.6	8.1

通过实验对比发现,提出的各项改进策略均能起到相应的效果,*Hardswish* 激活函数能提升算法的稳定性,*Ghost module* 在牺牲部分精度的前提下,能够大幅减少模型的参数量和浮点计算量,实现模型的轻量化。改进的损失函数能够大幅提高模型的检测精度。

### 3.6 最终模型的检测效果

相较于原始的YOLOv5模型,改进后的模型在对绝缘子检测的精度、速度和框选的准确性上均有明显提升,精度上提升了1.1%,推理速度提升了2.9%,模型大小和浮点计算量分别减小了44.5%和48.7%,表现效果较为理想。初始模型和改进模型的检测对比如图7所示,无人机拍摄的绝缘子图像具有尺度变化大,角度各异的特点,且本身大多呈长条状。相较于正矩形框,旋转矩形框能够更好的贴合绝缘子,减少标注框中的冗余信息,加速模型的训练速度。除此之外,当目标过于紧密时,精准的标注方式可以避免一些正确的矩形框被非极大值抑制‘错杀’。

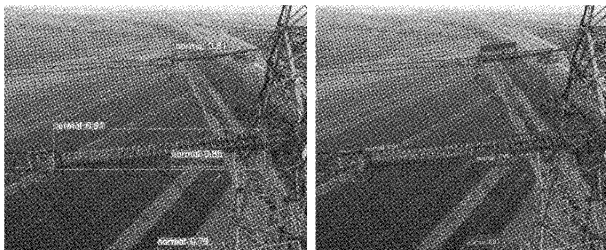


图7 模型检测对比

## 4 结论

本发明的优点在于与传统的针对输电线路绝缘子人工检测方法相比,本发明不仅能保证巡检人员安全,而且能大幅提升检测效率。而相对于其他基于深度学习的目

标检测算法,本发明最大的优势在于能够精确标注目标,并且满足算法移动端部署实时性的要求。本算法通过使用针对绝缘子自爆故障的旋转矩形框三标签标注法,有利于约束网络的训练方向和减少网络的收敛时间,并且提升网络的检测效果。通过引入*Ghost module* 模块大幅降低模型的网络参数,实现算法的轻量化,通过*Hardswish* 函数增强模型的数值稳定性。通过损失函数的改进,提升算法预测框回归效果。

尽管改进的算法有较好的表现效果,但依旧存在一些值得继续研究的问题,首先算法在高分辨率图像上的表现效果欠佳,其次虽然旋转矩形框的形状及中心点位置预测较为精准的,但是对角度的预测表现还有所不足,特别是角度为 $-\pi/2$ 的位置角度偏差可能很大,此外,本文只进行了静态图像测试,对于移动端视频检测的实时性要求还没有验证,实际检测中仍要考虑多种扰动因素,如无人机的抖动、环境干扰等,这些都需要继续深入研究。

### 参考文献

- [1] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[C]. IEEE Conference on European Conference on Computer Vision, 2014:346-361.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. R-CNN: Rich feature hierarchies for accurate object detection and semantic segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014:580-587.
- [3] LIU W, ANGUEROY D, ERHAN D, et al. SSD: Single shot multibox detector [C]. 14th European Conference on Computer Vision, 2016:21-37.
- [4] REDMON J, FARHADI A. YOLOv3: An incremental

- improvement [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 1-4.
- [5] 杨焰飞, 曹阳. 改进 YOLOv3 的无人机拍摄图玻璃绝缘子检测[J]. 计算机工程与应用, 2022, 58(3): 259-265.
- [6] AZIMI S M, VIG E, BAHMANYAR R, et al. Towards multi-class object detection in unconstrained remote sensing imagery [C]. Asian Conference on Computer Vision, 2018: 150-165.
- [7] LIN T Y, ROSS G, HE K M. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2117-2125.
- [8] 王雪, 程换新, 骆晓玲, 等. 基于改进的 YOLOv5 网络的异常行为检测算法研究[J]. 电子测量技术, 2022, 45(16): 137-141.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [10] WANG C Y, LIAO H Y, YE H I H, et al. CSPNET: A new backbone that can enhance learning capability of CNN [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020: 390-391.
- [11] WANG W H, XIE E Z, SONG X G, et al. PAN: Efficient and accurate arbitrary shaped text detection with pixel aggregation network [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 8440-8449.
- [12] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [13] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2019.
- [14] ZHENG Z H, REN D W. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [15] ZHENG Y F, ENR W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021.

#### 作者简介

余银锴, 硕士, 主要研究方向为深度学习, 计算机目标检测等。

E-mail: 1985148745@qq.com