

DOI:10.19651/j.cnki.emt.2210050

# 基于改进 YOLOv5s 绝缘子缺陷检测技术研究\*

肖荣俊<sup>1</sup> 潘睿志<sup>2</sup> 李超<sup>2</sup> 黄纪刚<sup>3</sup>

(1. 成都工业学院数字孪生实验室 成都 610031; 2. 成都理工大学机电工程学院 成都 610059;  
3. 四川大学机械工程学院 成都 610065)

**摘要:** 采用无人机对输电线路进行智能巡检作业已成为行业主流,绝缘子缺陷检测是智能巡检作业中的关键环节。针对复杂环境中绝缘子缺陷检测精度低的问题,本文提出一种基于改进 YOLOv5s 绝缘子缺陷检测算法。首先,对现有数据集利用随机矩形遮挡、水平翻转、随机像素置零、添加随机像素等操作进行数据增强,并利用 K-means 算法对数据集进行聚类分析,得到最佳锚框尺寸,有效提高模型的泛化能力和定位精度;其次,在 YOLOv5s 的主干网络的末端和最后 3 个不同规模的卷积网络后加入 GAM 注意力模块,使模型可以在更大的网络上进行注意,来解决无效特征对识别精度的影响;最后,在特征金字塔结构 FPN 的基础上,引入自适应特征融合 ASFF 模块,来增强网络的特征提取能力。实验结果表明,改进后 YOLOv5s 模型的精确率和 mAP0.5 相比于原 YOLOv5s 网络分别提高了 2.4% 和 2.2%。

**关键词:** 绝缘子;YOLOv5s;GAM;目标检测;ASFF

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## Research on defect detection technology based on improved YOLOv5s insulator

Xiao Canjun<sup>1</sup> Pan Ruizhi<sup>2</sup> Li Chao<sup>2</sup> Huang Jigang<sup>3</sup>

(1. Digital Twin Laboratory, Chengdu Technological University, Chengdu 610031, China;  
2. College of Mechanical and Electrical Engineering, Chengdu University of Technology, Chengdu 610059, China;  
3. School of Mechanical Engineering, Sichuan University, Chengdu 610065, China)

**Abstract:** The use of unmanned aerial vehicles for intelligent inspection of transmission lines has become the mainstream of the industry. Insulator defect detection is a key link in intelligent inspection operations. Aiming at the problem of low accuracy of insulator defect detection in complex environment, this paper proposes an improved YOLOv5s insulator defect detection algorithm. Firstly, the existing data sets are enhanced by random rectangle occlusion, horizontal flip, random pixel zeroing and adding random pixels, and the K-means algorithm is used to cluster the data sets to obtain the optimal anchor frame size, which effectively improves the generalization ability and positioning accuracy of the model. Secondly, GAM attention module is added to the end of the main network of YOLOv5s and the last three convolution networks of different scales, so that the model can be noticed on a larger network to solve the influence of invalid features on the recognition accuracy. Finally, based on the feature pyramid structure FPN, the adaptive feature fusion ASFF module is introduced to enhance the feature extraction ability of the network. The experimental results show that the accuracy and mAP0.5 of the improved YOLOv5s model are 2.4% and 2.2% higher than those of the original YOLOv5s network, respectively.

**Keywords:** insulator;YOLOv5s;GAM;target detection;ASFF

## 0 引言

我国是目前世界上人口最多的国家,为了满足各地区

的基础设施与能源供给,我国建设了大量的输电线路,维持高压输电线路和输电系统的正常运行,至关重要。绝缘子作为一种绝缘控件,主要作用是用来固定导体,其在

收稿日期:2022-05-22

\* 基金项目:四川省科技计划项目(2022YFG0326)资助

架空输电线路中起到重要作用。由于绝缘子长期暴露在户外,受到恶劣天气、机械负荷等因素的破坏,成为输电线路中损坏率最高的器件之一<sup>[1]</sup>。因此对于绝缘子及缺陷部分检测有重大的工程意义<sup>[2]</sup>。目前,我国大部分地区的电力线路巡检的方式还比较落后,仍然采取人工巡检方式为主<sup>[3]</sup>。巡检人员使用望远镜对高压输电线路上的器件进行观察,这种方式的精准度非常低下,当在环境复杂的山区巡检时,巡检人员需要翻山越岭对高压输电线路进行观察,不仅费时费力,甚至可能会危及巡检人员的安全。最重要的是无法全面和及时的掌握绝缘子等设备的状态。近年来,随着人工智能领域中深度学习的兴起,电力行业开始使用无人机对高压输电线路进行巡检,从而取代巡查员,以更高效、更安全地方式完成巡逻作业。这对于电力行业是一大步的跨越<sup>[4]</sup>。

目前针对绝缘子的目标检测主要有如下两个途径:传统方法:石丽<sup>[5]</sup>提出采用不同的边缘检测方法检测绝缘子图像的边缘,并利用边缘方向直方图对绝缘子进行识别。这种方法检测精度低、鲁棒性差,并且检测速度很慢,达不到在复杂环境下对绝缘子检测精度的要求。张晶晶等<sup>[6]</sup>提出一种感知聚类平行线段的方法。该算法克服了基于颜色检测绝缘子的不足,但存在步骤繁琐等问题。

深度学习检测方法:主要分为两大类,分别为单阶段检测器和两阶段检测器。两阶段检测器的主要思想是通过两步操作来实现目标检测,首先是生成可能包含目标物体的候选区域,然后针对每个候选区域,进行分类和校准。两阶段检测器代表模型主要有 Faster R-CNN<sup>[7]</sup>、Fast R-CNN<sup>[8]</sup>、R-CNN<sup>[9]</sup>等。单阶段检测器是指仅需要一次特征提取就可以实现目标检测。单阶段检测器的典型模型有 YOLO (you only look once)<sup>[10-12]</sup> 系列、SSD (single shot multibox detector)<sup>[13-14]</sup> 系列等。

利用深度学习的目标检测方法:纪超等<sup>[15]</sup>提出联合显著区域和 Fast-CNN 网络用于绝缘子特征检测研究,虽然该方法相较原 Fast-CNN 网络识别精度和速度都有提高,但与达到快速检测的要求仍存在一定的差距。Liu 等<sup>[16]</sup>提出 MTI-YOLO 网络,使用了多尺度特征检测头、多尺度特征融合结构和空间金字塔池模型,提升了模型的精准度,但是速度略有降低,并且只检测了正常绝缘子,并没有针对绝缘子缺陷进行检测。Zhao 等<sup>[17]</sup>提出了改进区域卷积神经网络的 RCNN 和 NMS 的锚生成方法,改进后的网络能够有效地识别和检测绝缘子,有效地提高了检测速度,但检测精度较低。

针对复杂环境中绝缘子检测精度低的问题,本文提出基于改进 YOLOv5s 绝缘子缺陷检测算法,该算法架构以 YOLOv5s 网络为基础进行改进,首先对现有数据集进行数据增强,并利用 K-means 算法对训练数据进行聚类分析,有效提高模型的泛化能力和定位精度,其次融合 GAM<sup>[18]</sup> 注意力机制和 ASF<sup>[19]</sup> 自适应特征融合机制,提高网络的性能。改进后的网络在复杂环境下对绝缘子缺陷有更好的检测效果。

### 1 YOLOv5s 模型

YOLOv5 网络是一个持续更新并且开源的神经网络,其中 YOLOv5-v5.0 网络共有 4 个版本,分别为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。从 YOLOv5s 到 YOLOv5x 的网络深度与宽度依次增加。表 1 为 YOLOv5 的各个版本在 COCO 数据集<sup>[20]</sup>上各项指标的对比,YOLOv5s 在保证较高的精度的同时,具有非常快的检测速度,并且模型参数量较少,容易在嵌入式终端部署,因此本文选择 YOLOv5s 作为基础网络模型。

表 1 YOLOv5 各版本指标对比

网络模型	图片大小	mAP <sup>val</sup> 0.5	mAP <sup>val</sup> 0.5;0.95	检测速度 v100/ms	参数/M
YOLOv5s	640	55.4	36.7	2.0	7.3
YOLOv5m	640	63.1	44.5	2.7	21.4
YOLOv5l	640	66.9	48.2	3.8	47.0
YOLOv5x	640	68.8	50.4	6.1	87.7

YOLOv5s 模型主要由 Backbone、Head 两个部分组成,其网络结构如图 1 所示。Backbone 是 YOLOv5s 网络中的特征提取模块,主要包括 Focus、C3、SPP 模块。Head 部分主要包含提取融合特征的颈部 Neck 和 Detect 模块。

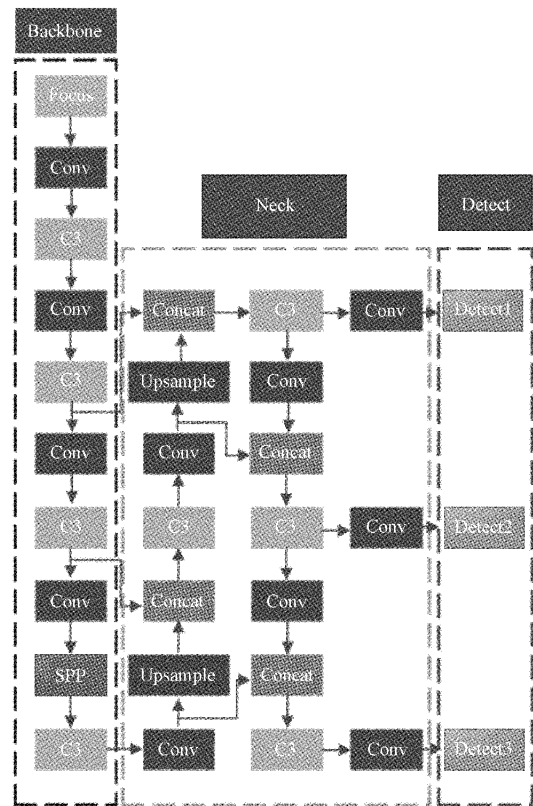


图 1 YOLOv5s 网络结构

Focus 模块:这是 YOLOv5s 中一个独创结构,主要的作用是减少计算量和参数量,从而提升速度。主要思想是把输入数据切分为 4 份,每份数据都等同于经过 2 倍下采样得到的,然后在 channel 维度上进行拼接,最后进行卷积操作。Focus 操作如图 2 所示。

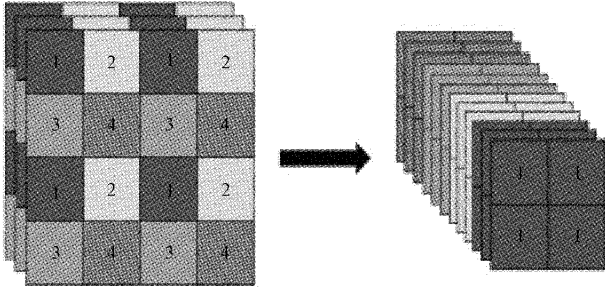


图 2 Focus 操作

C3 模块:在 YOLOv5s 中将 BottleneckCSP 模块替换为 C3 模块,替换之后不仅精简了网络结构,并且减少了计算量和降低了模型的推理时间。C3 模块与 BottleneckCSP 模块相比,不同点是经历过残差输出后的卷积模块被去除,并且 Concat 后的 BN 层与激活层也被去除。C3 模块如图 3 所示。

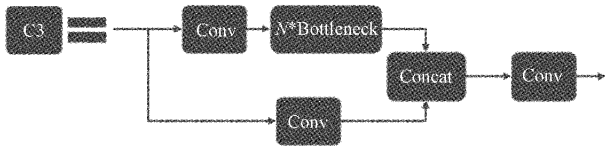


图 3 C3 模块

效仿何凯明提出的空间金字塔池化 (spatial pyramid pooling<sup>[21]</sup>, SPP),其主要思想是将输入的特征层并行通过多个不同大小的最大池化下采样层,再将得到的输出与原输入进行拼接。SPP 模块最大的作用就是增大了感受野,提取出最重要的上下文特征,解决了多尺度问题。SPP 模块如图 4 所示。



图 4 SPP 模块

Head 模块是由特征金字塔结构 (feature pyramid networks, FPN)<sup>[22]</sup>与路径聚合网络结构 (path aggregation network, PAN)<sup>[23]</sup>相结合组成。FPN 是一个自上而下的特征金字塔,其主要思想是通过 CNN 网络层次结构高效的进行强特征计算,再使用上采样的方式将高层与底层的特征信息进行融合,从而获得较强的语义特征,提高在目标检测上面的性能表现。而 PAN 结构则是在 FPN 的后面增加一个自下而上的金字塔。通 FPN+PAN 两者相结合的方法,实现了不同的主干层对不同的检测层的特征聚合,有效地解决了多尺度问题。

## 2 YOLOv5s 改进

### 2.1 GAM

由于本文使用数据集的图像背景多为山丘和树林,环境复杂多变,并且绝缘子缺陷目标仅占据图像的很小一部分,为了提高模型对绝缘子及缺陷部分的特征表达能力,本文在 YOLOv5s 的主干网络末端与最后 3 个不同规模的卷积网络后加入 GAM 注意力模块。

GAM 注意力模块是在 2021 年提出的一种通过减少信息弥散和放大大局交互表示来提升神经网络性能的全局注意力机制。GAM 全局注意力模块能够提高网络模型对绝缘子及缺陷部分的关注度,减少复杂背景的干扰,提升模型检测精度。GAM 全局注意力模块采用 (convolutional block attention module, CBAM)<sup>[24]</sup>中的顺序通道,首先由通道注意力特征图 (channel attention, CA) 对原特征图进行校正,再由空间注意力特征图 (spatial attention, SA) 对校正结果再进行校正,得到最终的特征图。其 GAM 模块结构如图 5 所示。

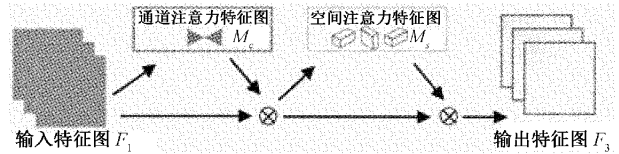


图 5 GAM 模块

图 6 为通道注意力机制结构图。首先将输入特征图  $F_1$  的大小设定为  $C \times W \times H$ ,其中  $H$  和  $W$  表示输入特征的长和宽, $C$  表示通道数,对输入的图像使用 3D 排列操作来保留三维信息,再将输出通过 1 个 2 层 (multi layer perception, MLP) 来放大跨维通道,再对输出使用逆向 3D 排列操作,将得到结果通过 sigmoid 函数激活,最后得到新特征图  $M_c(F_1)$ 。

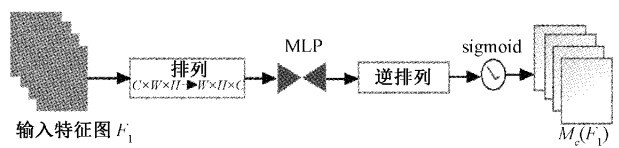


图 6 通道注意力机制结构图

图 7 是空间注意力机制的结构图,假定输入特征图  $F_2$  的大小为  $C \times H \times W$ ,首先将输入特征经过两个  $7 \times 7$  大小的卷积层进行空间融合,再将输出经过 sigmoid 函数激活,最后得到新特征图  $M_s(F_2)$ 。

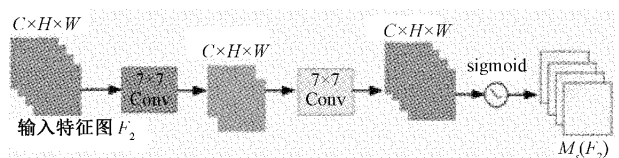


图 7 空间注意力机制结构图



为了验证 GAM 注意力模块的有效性,对复杂环境下绝缘子进行热力图实验,从图 8 可以看出未使用 GAM 注意力模块的网络,未能对绝缘子形成有效聚焦。当加入注意力模块后,对检测目标有了一定的侧重,一定程度改善了对检测区域的关注,有效提升了复杂环境下对待检目标的聚焦能力。

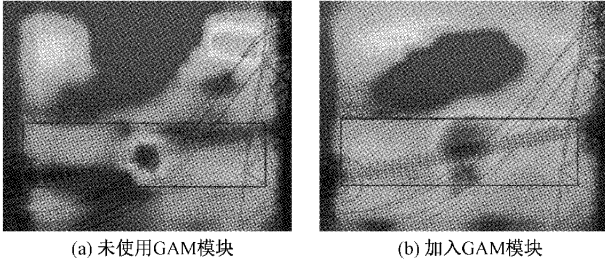


图 8 热力图对比试验

### 2.2 ASFF

为了提高对绝缘子及缺陷部分的特征提取能力,并且有效抑制复杂背景中无效特征对检测的干扰,需要加强对多尺度特征的融合利用,目前多尺度特征金字塔是解决目标检测中多尺度变化的常用手段。在原 YOLOv5s 算法中的 FPN 结构仅将不同特征层调整为统一尺寸后再进行累加,但是不同特征尺度之间不一致将导致融合特征图噪声变大,使得效果变差。

ASFF 模块会通过自适应方法来调整不同特征层之间的融合比例,在空间上过滤冲突信息以抑制梯度反向传播时不一致,从而改善特征比例不变性,并且减少推理开销,改善传统多尺度特征融合中图像空间信息冲突的问题。ASFF 模块可以被应用在任何 FPN 形式的网络中,并且它的融合方式的过程是可微分的,非常适用于反向传播。本文在 YOLOv5s 网络多尺度特征提取的基础上,引入自适应特征融合 ASFF 模块,能够融合不同尺度感受野下绝缘子及缺陷部分目标的特征信息,有效改善了目标特征信息丢失的问题。ASFF 的结构设计图如图 9 所示。

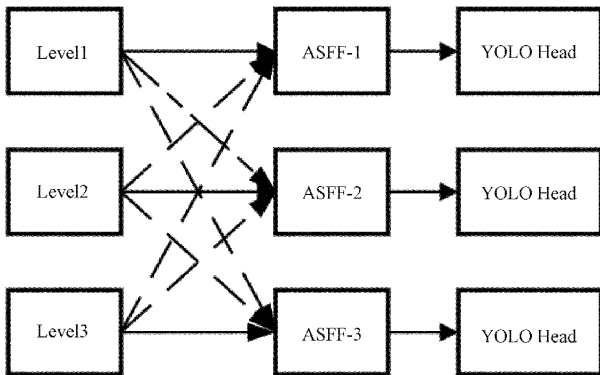


图 9 ASFF 的结构设计图

YOLOv5s 网络中的 Neck 部分输出为 Level1、Level2 和 Level3 的特征图,以 ASFF-1 为例,融合后的 ASFF-1 输

入来自于 Level1、Level2、Level3 的特征与可学习权重  $\alpha^1$ 、 $\beta^2$ 、 $\gamma^3$  权重相乘后再进行相加后的结果。如式(1)所示。

$$y_{ij}^l = \alpha_{ij}^l \cdot x_{ij}^{1 \rightarrow l} + \beta_{ij}^l \cdot x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot x_{ij}^{3 \rightarrow l} \quad (1)$$

其中,  $y_{ij}^l$  代表通道之间输出特征映射  $y^l$  的第  $(i, j)$  个向量,  $\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l$  代表 3 个不同级别到  $l$  层特征图的可学习权重,  $x_{ij}^{1 \rightarrow l}, x_{ij}^{2 \rightarrow l}, x_{ij}^{3 \rightarrow l}$  代表某个位置特征图的输出。

由于是采用相加的方式,所以需要保证 Level1 ~ Level3 的通道数与特征大小一致。通过对不同层级使用下采样或上采样方式来调整尺寸,然后 Level1、Level2 与 Level3 特征图经过一个  $1 \times 1$  卷积得到权重参数  $\alpha, \beta, \gamma$ , 再将权重参数进行拼接通过 softmax 函数归一化,将输入映射到  $[0, 1]$  范围内,且总和为 1。其中  $\alpha_{ij}^l$  公式如下:

$$\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1 \quad (2)$$

$$\alpha_{ij}^l = \frac{e^{\lambda_{\alpha_{ij}}^l}}{e^{\lambda_{\alpha_{ij}}^l} + e^{\lambda_{\beta_{ij}}^l} + e^{\lambda_{\gamma_{ij}}^l}} \quad (3)$$

YOLOv5s 网络在目标检测、无人驾驶、物流分类等领域都有良好的表现,其重要原因之一就是其主干网络有强大的特征表达能力,YOLOv5s 在设计时主要用于通用目标检测,但本文使用的绝缘子数据集的图像背景较为复杂,针对这种复杂场景的辨别能力还存在不足。此外,本文中使用的绝缘子训练数据集与通用目标检测训练集数据规模相差比较大,仅用未改进的 YOLOv5s 网络难以有效检测目标,针对上述问题,在网络中融合 GAM 注意力机制和 ASFF 自适应特征融合机制,强化网络在复杂环境下的特征表达能力,提升网络在复杂环境中检测绝缘子及缺陷的检测精度。

改进后 YOLOv5s 网络整体结构如表 2 所示,其中输入列中的 -1 表示为来自于上一层输出,张量信息列代表该模块的输入通道数、输出通道数、卷积核大小、步长等参数信息。模块列表示该模块的名称。

### 2.3 基于 K-means 聚类的初始候选框

K-means 是一种迭代求解的聚类算法,当应用于较大数据集的时候,具有良好的伸缩性。由于算法复杂程度低,实现起来也比较简单,聚类效果也很好,因此 K-means 聚类算法应用非常广泛。由于原 YOLOv5s 网络默认锚定框主要是针对 COCO 数据集聚类得到的,因此对于本文需要识别的对象并不适用。因此利用 K-means 算法来确定锚定框的最优尺寸,对于提升模型精度有很好的效果。本文采用的标记框到类簇的中心的距离  $D$  如下:

$$D = 1 - IOU(box, centroid) \quad (4)$$

其中,  $IOU$  为交并比运算,  $box$  为标记框,  $centroid$  为类簇中心。聚类之后得到的 9 个先验框的尺度,作为本文网络的先验框的尺度。聚类之后得到的 9 个先验框的尺度如表 3 所示。

其中 K-means 算法具体的步骤如下:

- 1) 随机设置类簇的个数  $K$ 。

表2 改进后YOLOv5s整体网络结构图

序号	输入	模块	张量信息
0	-1	Focus	[3,32,3]
1	-1	Conv	[32,64,3,2]
2	-1	C3	[64,64,1]
3	-1	Conv	[64,128,3,2]
4	-1	C3	[128,128,3]
5	-1	Conv	[128,256,3,2]
6	-1	C3	[256,256,3]
7	-1	Conv	[256,512,3,2]
8	-1	SPP	[512,512,[5,9,13]]
9	-1	C3	[512,512,1,False]
10	-1	GAM	[512,512]
11	-1	Conv	[512,256,1,1]
12	-1	Up	[None,2,nearest]
13	[-1,6]	Concat	[1]
14	-1	C3	[512,256,1,False]
15	-1	Conv	[256,128,1,1]
16	-1	Up	[None,2,nearest]
17	[-1,4]	Concat	[1]
18	-1	C3	[256,128,1,False]
19	-1	GAM	[128,128]
20	-1	Conv	[128,128,3,2]
21	[-1,15]	Concat	[1]
22	-1	C3	[256,256,1,False]
23	-1	GAM	[256,256]
24	-1	Conv	[256,256,3,2]
25	[-1,11]	Concat	[1]
26	-1	C3	[512,512,1,False]
27	-1	GAM	[512,512]

表3 修改前后的先验框的长宽

原始的先验框	K-means 聚类后的先验框
(10, 13)	(37, 26)
(16, 30)	(134, 24)
(33, 23)	(164, 39)
(30, 61)	(146, 108)
(62, 45)	(303, 52)
(59, 119)	(448, 61)
(116, 90)	(436, 92)
(156, 198)	(453, 131)
(373, 326)	(447, 206)

2)在数据集中随机选取  $K$  个默认宽高比作为类簇的初始质心。

3)计算所有样本与类簇中心的距离,再将所有样本划分到离它最近的类簇中。

4)更新类簇的中心,计算每个类簇中所有样本的均值,并将其作为新的聚类中心。

5)重复上面两步直到类簇中心不在变化。

## 2.4 损失函数

YOLOv5s 的损失函数包括 3 部分:分类损失(classification loss)、定位损失(localization loss)、置信度损失(confidence loss)。总的损失函数的大小为这 3 部分损失结果的总和。目标检测回归损失函数有很多种。从最开始的 IOU 到 GIOU<sup>[25]</sup> 再到 DIOU 最后发展到 CIOU<sup>[26]</sup>, CIOU 是在 DIOU 在基础上进行改进而提出的。CIOU 相比于 DIOU 增加了检测框尺度的 Loss,这样会使预测框更加符合真实框。因此本文最后选择了 CIOU\_loss 作为网络的损失函数。CIOU 损失函数计算式如下:

$$CIoU = IoU - \left( \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \right) \quad (5)$$

$$L_{CIOU} = 1 - CIOU \quad (6)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

其中,  $v$  表示度量长宽比的一致性;  $\rho^2(b, b^{gt})$  表示两个中心点之间的欧氏距离;  $\alpha$  表示一个权重参数;  $c$  表示最小包围框的对角线距离,  $w^{gt}, h^{gt}$  表示真实框的宽度与高度;  $w, h$  表示预测框的宽度与高度。

## 3 实验与分析

### 3.1 数据准备

本文的实验数据集来自中国电力线路绝缘子数据集,由无人机拍摄正常绝缘子图像共 600 张,缺陷绝缘子 248 张,图像大小为  $1\ 152 \times 864$  pixels,由于数据集太少,避免样本不足导致出现过拟合现象,进而影响绝缘子及缺陷部分的检测效果。故采用数据增强工具对原数据集进行扩充,通过随机矩形遮挡、水平翻转、随机像素置零、随机裁剪与填充等方式将图像扩充至 3 740 余张,有效地增大训练集的规模,从而提高了模型的泛化能力,本文随机将训练集、验证集、测试集的比例划分为 8 : 1 : 1。部分图像数据增强效果如图 10 所示。

本文使用 LabelImage 目标检测标注工具对所有正常绝缘子与缺陷绝缘子进行标注,并将正常绝缘子标注为 insulator,缺陷位置标注为 defect,具体的操作如图 11 所示,通过上述工作,完成数据集的制作。

### 3.2 实验环境

本文实验的操作系统为 Window10,并采用 Pytorch 框架。具体的实验硬件环境:CPU 型号为 i7-10870H,显卡型为 GeForce RTX 3060 Laptop GPU,显存大小为 6 G,内存大小为 16 G。本文提出改进后的 YOLOv5s 模型训练的超参数设置如表 4 所示。

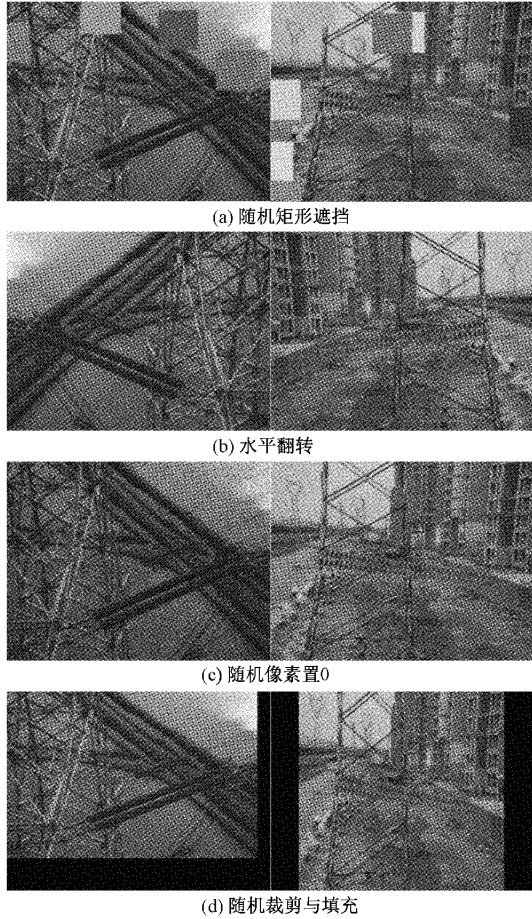


图 10 部分图像增强实例

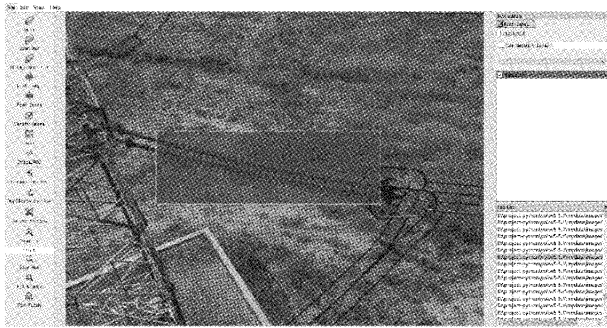


图 11 数据集标注

表 4 模型超参数设置

动量	初始学习率	迭代次数	批量大小	权重衰减
0.937	0.01	300	16	0.000 5

3.3 模型评估

为了全面客观评价本文提出的改进后 YOLOv5s 模型的性能,使用精确率(Precious)、召回率(recall)、平均精度(mAP)等指标来衡量,具体公式如下所示。

精确率的计算式为:

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

召回率的计算式为:

$$recall = \frac{TP}{TP + FN} \tag{10}$$

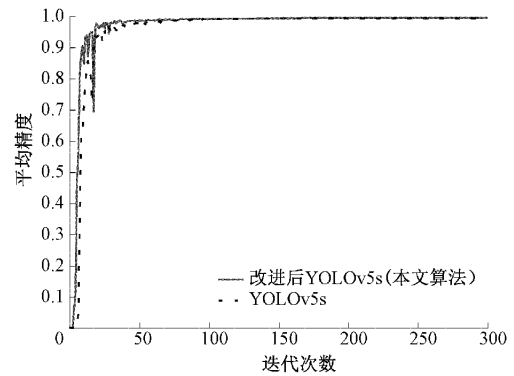
平均精度的计算式为:

$$AP = \int_0^1 P(r)dr \tag{11}$$

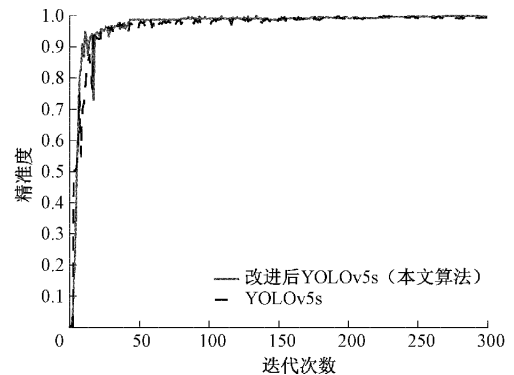
其中, TP 表示被模型预测为正类的正样本数量, FP 表示被模型预测为负类的正样本数量, FN 表示被模型预测为正类的负样本数量。

3.4 实验结果分析

为了验证本文算法的优势,利用 Tensorboard 可视化工具可以看到,改进后 YOLOv5s 网络与原 YOLOv5s 网络训练过程的 mAP、精准度变化情况对比如图 12 所示。改进后的网络在 30 轮左右开始逐渐收敛,而原 YOLOv5s 网络在 60 轮左右开始逐渐收敛,在有限的时间下改进后的网络可以取得更好的效果。



(a) 300次迭代后mAP的变化



(b) 300次迭代后精准度的变化

图 12 模型改进前后性能指标变化

消融实验一般用于复杂的神经网络中,用于探究网络中的特定模块是否对模型产生了正面影响或负面影响,对于网络结构的设计具有重要指导意义。本文在原 YOLOv5s 的基础上,在主干网络的末端与最后 3 个不同规模的卷积网络后加入 GAM 注意力模块,提高对提取特征



中重要区域的关注度。并且引入ASFF模块,融合不同尺度感受野下被检测目标的特征信息。为了验证改进后网络的正面影响,设置了消融实验,将改进后的YOLOv5s模型中的改进机制逐个添加并进行训练。验证模型的改进机制是否具有正面影响。消融实验具体结果如表5所示。

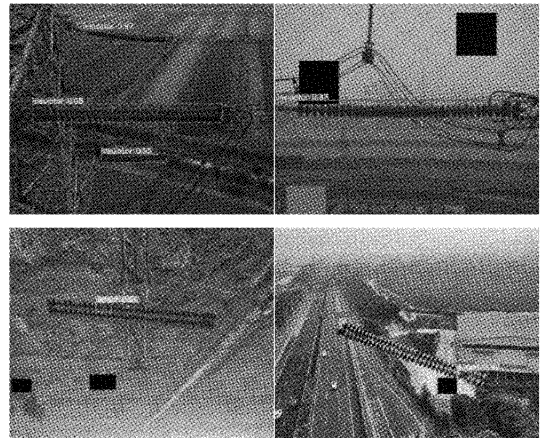
表5 消融实验结果

ID	ASFF	GAM	K-means	mAP0.5/%	Precious/%
1	×	×	×	0.813	0.939
2	×	×	√	0.815	0.961
3	×	√	×	0.829	0.959
4	√	×	×	0.826	0.921
5	√	√	×	0.828	0.934
6	√	√	√	0.837	0.961

表5中第2行为原始YOLOv5s网络的检测结果。通过表5可以看出,当对数据集使用K-means算法聚类后,提高了定位精度,相比于原YOLOv5s算法Precious提高了2.2%。当向原网络中加入ASFF模块后mAP提高了1.13%,当向原网络中加入GAM模块后Precious相比于原YOLOv5网络提高了2%,表明GAM模块、K-means算法聚类与ASFF模块对于网络提升检测精度均具有正面影响。

最后向网络中同时融合GAM注意力机制和ASFF自适应特征融合机制,并使用K-means算法将数据集聚类后,有效提高了模型的检测精度,相比于原YOLOv5s网络mAP0.5提高了2.4%,精准率提高了2.2%。本文算法综合了以上各模块的优点,检测精度可以达到96.1%,mAP可以达到83.7%,在复杂环境下对缺陷绝缘子及正常绝缘子均有较好的识别效果。

为了验证改进后算法的优势,选取了部分正常绝缘子和缺陷绝缘子在复杂环境下进行测试验证。由图13(a)、(b)对比得知,在第1组实验和第3组实验可以看出,改进后的模型检测效果有所提升。在第2组实验和第4组实验可以



(b) 本文算法检测结果

图13 不同场景下检测效果对比

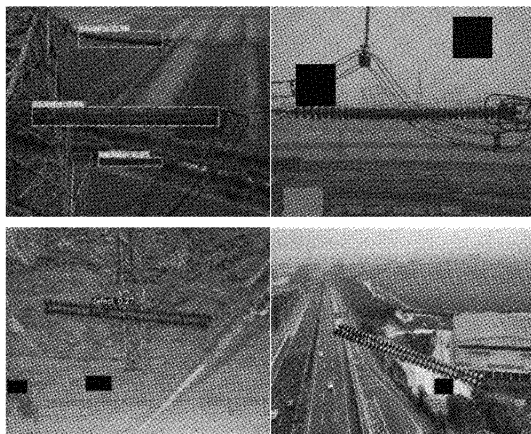
看出,由于待检测目标的背景较为复杂,原YOLOv5s算法出现漏检现象,而本文算法采用了GAM注意力模型和自适应特征融合ASFF模块,显著提高了模型特征提取能力,进而可以准确检测出目标,有效的降低了漏检率。因此本文提出的算法是一种可行的绝缘子缺陷检测算法。

## 4 结 论

针对复杂环境下绝缘子识别精度低问题。本文提出了一种改进的YOLOv5s的目标检测模型,该模型以YOLOv5s网络为基础进行改进,首先对现有数据集进行数据增强,并利用K-means算法对训练数据进行聚类分析,有效提高模型的泛化能力和定位精度,并且为了增加特征图的表达能力,在网络中融合GAM注意力机制和ASFF自适应特征融合机制,最后为了减少目标框的定位损失使用了CIoU损失函数,通过以上的操作,使得绝缘子的精确率、mAP0.5相比于原网络提升了2.4%、2.2%。精确率达到96.1%,mAP达到83.7%。未来,进一步优化网络结构,提升检测精度,并且降低漏检率。

## 参考文献

- [1] 陈庆,闫斌,叶润,等. 航拍绝缘子卷积神经网络检测及自爆识别研究[J]. 电子测量与仪器学报, 2017, 31(6): 942-953.
- [2] 赵振兵,王乐. 一种航拍绝缘子串图像自动定位方法[J]. 仪器仪表学报, 2014, 35(3): 558-565.
- [3] 董召杰. 基于YOLOv3的电力线关键部件实时检测[J]. 电子测量技术, 2019, 42(23): 173-178.
- [4] 张培培,王昭,王菲. 基于深度学习的图像目标检测算法研究[J]. 国外电子测量技术, 2020, 39(8): 34-39.
- [5] 石丽. 输电线路绝缘子缺陷图像检测方法[D]. 北京: 华北电力大学, 2013.
- [6] 张晶晶,韩军,赵亚博,等. 形状感知的绝缘子识别与缺陷诊断[J]. 中国图象图形学报, 2014, 19(8): 1194-1201.



(a) 原YOLOv5s算法检测结果

- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. International Conference on Neural Information Processing Systems, Istanbul: Springer, 2015: 91-99.
- [8] GIRSHICK R. Fast R-CNN [C]. IEEE International Conference on Computer Vision, Santiago: ICCV, 2015: 1440-1448.
- [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] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii, USA: IEEE, 2017: 6517-6525.
- [11] REDMON J, FARHADIA Z. YOLOv3: An incremental improvement [J]. Arxiv Preprint, 2018, ArXiv:1804.02767.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [13] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector [J]. ArXiv Preprint, 2017, ArXiv:1701.06659.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision. Amsterdam, the Netherlands, DE: Springer, 2016: 21-37.
- [15] 纪超, 黄新波, 曹雯, 等. 改进的 Fast-CNN 模型在绝缘子特征检测中的研究 [J]. 计算机与现代化, 2019(4): 59-64.
- [16] LIU C, WU Y, LIU J, et al. MTI-YOLO: A lightweight and real-time deep neural network for insulator detection in complex aerial images [J]. Energies, 2021, 14(5): 1426.
- [17] ZHAO Z, ZHEN Z, ZHANG L, et al. Insulator detection method in inspection image based on improved Faster R-CNN [J]. Energies, 2019, 12(7): 1204.
- [18] LIU Y, SHAO Z, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions [J]. ArXiv Preprint, 2021, ArXiv: 2112.05561.
- [19] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection [J]. ArXiv Preprint, 2019, ArXiv: 1911.09516.
- [20] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C]. European Conference on Computer Vision. Cham: Springer, 2014: 740-755.
- [21] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Trans Pattern Anal Mach Intell, 2015, 37(9): 1904-1916.
- [22] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE Computer Society, 2017, DOI: org/10.48550/arXiv.1612.03144.
- [23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [24] WOO S, PARK J, LEE J Y, et al. Cham: Convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [25] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [26] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

### 作者简介

肖黎俊, 博士, 讲师, 主要研究方向为数字孪生技术、图像处理技术。

E-mail: canjunxiao@cdu.edu.cn

潘睿志, 硕士研究生, 主要研究方向为图像处理技术。

E-mail: 524453268@qq.com

李超(通信作者), 博士, 副研究员, 主要研究方向为特种机器人与视觉检测技术。

E-mail: lichao@163.com