

DOI:10.19651/j.cnki.emt.2415519

基于 YOLOv5s 的轻量化架空输电线路鸟巢检测网络^{*}

徐业东¹ 蔡亚恒¹ 李 严¹ 刘学雷¹ 曹英丽^{1,2}

(1. 沈阳农业大学信息与电气工程学院 沈阳 110161; 2. 辽宁省农业信息化工程技术中心 沈阳 110161)

摘要: 架空输电线路上的鸟巢侵占会对铁塔上的电力设备造成安全隐患,间接可能影响整个电力系统的稳定运行。针对目前架空输电线路鸟巢检测模型在复杂场景以及小目标场景下检测精度不高,检测效率低,模型复杂等问题。本研究提出一种基于 YOLOv5s 框架的轻量化架空输电线路鸟巢检测网络。首先在主干部分采用 Fasternet 重构 YOLOv5s 特征提取网络,降低模型复杂度,提高运行速度;然后在特征融合网络部分嵌入 ConvMixer 层,ConvMixer 层的结构设计有助于在特征信息中更好的捕捉空间和通道的关系,提升模型对于小目标的检测能力;最后在特征融合网络部分引入 ODConv 模块,令送入检测头的特征图包含更多有效特征,提高模型对复杂场景和小目标的检测性能。实验结果表明,本文与基线模型 YOLOv5s 相比,计算量和模型体积分别减少了 86% 和 72%,平均精度均值达到 96.4%,检测速度达到 104.2 帧/s,验证了本文改进模型的有效性和可行性。

关键词: 轻量化;架空输电线路;YOLOv5s;鸟巢;目标检测

中图分类号: TP391.4;TM75 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Lightweight overhead transmission line bird's nest detection network based on YOLOv5s

Xu Yedong¹ Cai Yaheng¹ Li Yan¹ Liu Xuele¹ Cao Yingli^{1,2}

(1. College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110161, China;

2. Agricultural Informatization Engineering Technology Center of Liaoning Province, Shenyang 110161, China)

Abstract: Bird nest encroachment on overhead transmission lines can cause safety hazards to the power equipment on the towers, which may indirectly affect the stable operation of the whole power system. Aiming at the current overhead transmission line bird's nest detection model in the complex scene as well as the small target scene detection accuracy is not high, the detection efficiency is low, the model is complex and other problems. This study proposes a lightweight overhead transmission line bird's nest detection network based on YOLOv5s framework. Firstly, the YOLOv5s feature extraction network is reconstructed by Fasternet in the backbone part, which reduces the model complexity and improves the operation speed; then the ConvMixer layer is embedded in the feature fusion network part, and the structural design of the ConvMixer layer helps to better capture the relationship between space and channel in the feature information, which improves the model's detection ability for small targets; finally, the ODConv module is introduced in the feature fusion network part, so that the feature map sent to the detection head contains more effective features to improve the detection performance of the model for complex scenes and small targets. The experimental results show that compared with the baseline model YOLOv5s, the computational amount and model volume are reduced by 86% and 72%, the average accuracy reaches 96.4%, and the detection speed reaches 104.2 frames/s, which verifies the effectiveness and feasibility of the improved model in this paper.

Keywords: lightweight; overhead transmission lines; YOLOv5s; a bird's nest; object detection

0 引 言

随着自然环境的日益改善,鸟类数量增多,由鸟类活动

引起的电网事故时有发生^[1]。常见由鸟害导致的电网故障包括鸟类直接触电短路、鸟类筑巢活动导致的短路^[2]以及鸟类排泄物覆盖绝缘器件引起的闪络^[3-4]等,其中,鸟类的

收稿日期:2024-02-22

^{*} 基金项目:国家自然科学基金(61903264)项目资助

筑巢活动会对架空输电线路的安全运行造成重大威胁^[5]。因此,对架空输电线路进行定期巡视检查,及时发现并拆除鸟巢是非常有必要的。近年来,基于无人机技术相关的巡检方式逐渐取代了传统的人工巡检方法,并广泛应用于架空输电线路的巡检工作^[6-7]。通过对无人机航拍图像进行目标检测,可以有效地识别输电线路上的各种缺陷^[8-9]。

针对架空输电线路图像中鸟巢的识别和定位,国内外相关研究已经开始,传统图像处理方法一般基于特征提取和分类模型。段旺旺等^[10]提出了一种基于关键区域 HOG 特征的铁路接触网鸟巢识别,通过对疑似鸟巢区域进行特征提取,有很高的识别率。祝振敏^[11]利用 Sobel 水平边缘检测算子得到图像边缘,概率 Hough 变换直线检测方法对图像进行角度校正,最后对图像进行 Otsu 二值化处理。Wu 等^[12]利用条纹方向直方图和条纹长度直方图来描述鸟巢,并应用进接触网的鸟巢检测。此类基于传统图像处理方法可以在特定的条件下对鸟巢进行识别和定位,但对数据依赖性不强,在泛用性上存在一定的局限,其本质原因是传统图像处理方法对复杂场景中鸟巢可区分特征的提取和描述能力不足。

近年来,随着深度学习与计算机视觉技术的不断发展,其中以卷积神经网络为基础的 FasterR-CNN^[13]、YOLO 系列^[14]、RetinaNet^[15]等目标检测算法成为研究热点,并在无人机输电线路目标检测中获得了一定的成果。王纪武等^[16]研究基于 FasterR-CNN 的多尺度高压架空输电线路鸟巢识别方法,完善特征提取和特征增强模块。刘国文等^[17]通过改进 RetinaNet 模型,增加 P2 特征层,来扩充网络的感受野范围,提高目标较小的鸟巢检测能力。丁建等^[18]提出了一种适用于鸟巢检测的双尺度 YOLOv3 网络,其特点在于提升检测算法的精度和抗噪声能力,但缺少对复杂背景下泛用性的测试。赵霖等^[19]在 YOLOv5 主干网络中引入 CBAM 注意力模块,自适应特征融合机制替代原始特征金字塔结构,并使用 Mish 函数作为激活函数,有效提升模型对鸟巢目标的检测精度。张焕龙等^[20]将设计的特征平衡网络和特征增强模块引入 YOLOv5 的特征融合网络部分,检测精度和 FPS 分别提高了 4.19% 和 2.1 ms。赵立杰等^[21]利用剪枝方法对改进后的 YOLOv5s 网络模型进一步稀疏化训练,减轻模型体积的同时保持较高的检测精度。肖黎俊等^[22]使用 K-means 算法对数据集进行聚类分析,结合注意力模块和自适应特征融合模块增强网络的特征提取能力,改进后的网络模型在精度上获得较好的提升。郑婷婷^[23]通过增加小目标检测层,将改进后的特征融合网络和注意力机制引入 YOLOv5s 模型中,检测精度提升了 5.6%。上述基于卷积神经网络的检测方法在模型体积,检测精度和推理速度等方面做了很多工作,并取得了较好得到检测性能。但多数模型参数规模较大,并没有更好的平衡模型体积和识别精度,因此轻量化且具有高精确识别能力的目标检测模型仍是当前重点研究对象。

综上所述,尽管有关架空输电线路鸟巢检测的研究在国内外已经展开,但目前的工作仍存在一些不足。主要问题在于视频图像尺度随无人机拍摄机位变化明显,在目标较小的场景以及复杂背景下难以保证对鸟巢的精确定位,检测性能下降;随着我国无人机巡检的普及与应用,为有效嵌入移动端,需保持较小的参数规模,以确保实时检测速度。针对以上问题,本文选用检测精度和速度较好的 YOLOv5s 作为基线模型,根据上述鸟巢检测中存在的问题对原网络模型进行改进,保证提高网络模型检测精度的同时降低模型参数体积,并与 YOLOv3、YOLOv7、YOLOv8 和 YOLOv9 等网络模型进行对比,以评估经过改进后的模型性能,为维护电网的安全运行和预防潜在的事故危险提供技术参考。

1 相关理论和技术

YOLOv5 是由 Ultralytics 公司于 2020 年发布的单阶段目标检测网络,在 YOLO 系列中具有较高的速度和性能,适用于实时检测的应用场景,作为 YOLO 系列的代表作,是目前比较成熟的目标检测模型之一^[24-25]。在 YOLOv5 的官方代码中根据模型深度倍数 (depth_multiple) 和层通道倍数 (width_multiple) 的大小分为 s、m、l 和 x,同时模型的复杂度也依次增加,本文选择了参数量和计算量较少的 YOLOv5s 作为基线模型。

YOLOv5 网络结构包括 4 个部分:输入端 Input、主干网络 Backbone、多尺度特征融合网络 Neck 和预测分类器 Head。如图 1 所示,YOLOv5s 整体网络框架,在输入端的图像预处理阶段采用 Mosaic 数据增强、自适应初始锚框计算以及图片缩放。其次,本文采用 YOLOv5-6.0 版本,该版本模型的主干网络主要由 CBS (Conv、batchnormalization、silu)、C3 (cross stage partial bottle)、SPPF (spatial pyramid pooling fast) 组成,通过对输入图像的特征提取,得到蕴含不同特征信息和尺寸的特征图。多尺度特征融合网络的网络结构采用路径聚合网络 (path aggregation network, PANet)^[26] 结构,该结构对高层级特征图中的语义信息以及低层级特征图中的位置信息进行融合,从而获得更丰富的目标特征。预测分类器由 3 个检测头组成,使用 1×1 卷积来代替全连接层进行预测,每个检测头分别对应小、中、大 (80×80 , 40×40 , 20×20) 3 个不同尺度的特征图进行目标检测,最后输出的预测结果对应着目标框坐标、置信度以及类别信息。

针对上述鸟巢目标检测领域存在的问题,本文提出了基于改进 YOLOv5s 的轻量化鸟巢检测网络,其网络框架如图 2 所示。首先在主干部分采用 Fasternet 重构 YOLOv5s 特征提取网络,大幅度降低模型参数数量和计算复杂度。其次是在特征融合网络部分集成相应数量 ConvMixer 层,ConvMixer 层的结构设计有助于在特征信息中更好的捕捉空间和通道的关系,提升模型对于小目标

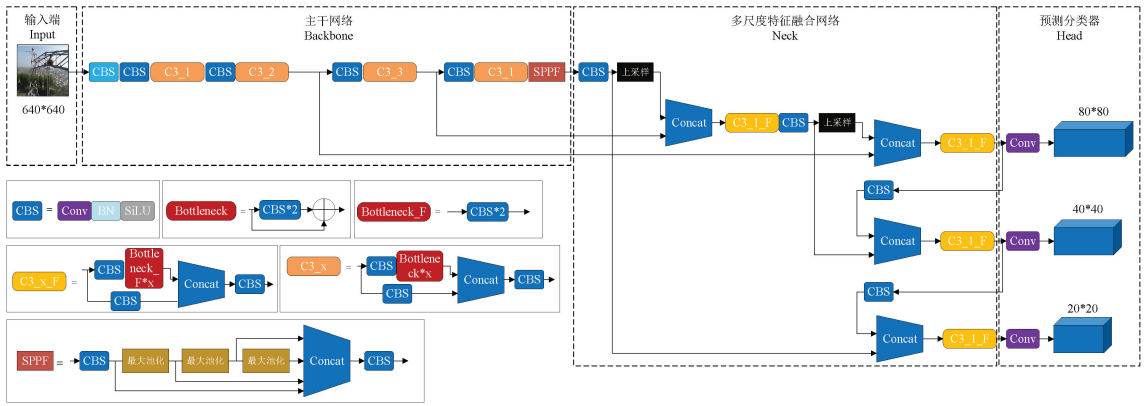


图 1 整体网络框架

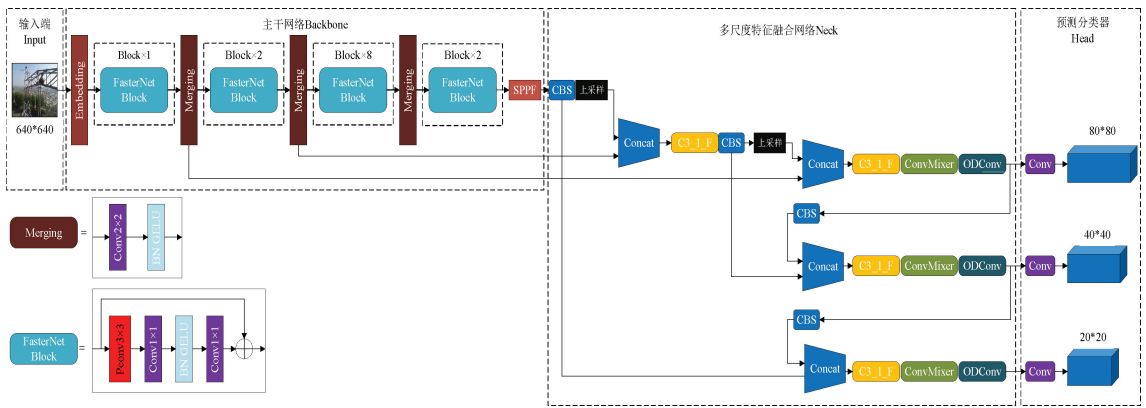


图 2 改进后得整体网络框架

的检测能力。最后在特征融合网络部分的 ConvMixer 层后引入 ODConv 模块,融合 ODConv 模块后的目标检测网络在特征聚合时能够充分的捕捉上下文信息,使得送入检测头的特征图包含更多有效特征,从而提升网络模型对于复杂背景以及小目标的检测精度。

2 网络模型的改进

2.1 FasterNet 轻量化主干网络

与使用昂贵的计算设备以提高神经网络检测速度相比,目前研究更趋向于开发经济高效的快速神经网络。在当前轻量化目标检测任务中通常使用 MobileNet、ShuffleNet、GhostNet 等卷积神经网络,这些网络利用深度可分离卷积(depthwise separable convolution, DWConv)^[27]或分组卷积(grouped convolution, GConv)^[28],以实现轻量化检测任务的目的。深度可分离卷积将完整的卷积运算分成两个步骤实现,这种分解方式可以在一定程度上减少参数数量和计算量,提高模型的检测效率。其中模型在逐点卷积阶段尝试整合不同通道信息,但由于深度卷积阶段每个通道上都是独立操作,这就会导致不同通道之间的信息交互受到限制,进而导致模型学习到的特征信息不够丰富。分组卷积将输入通道分为多个组,每个组内的通道共享权重,各组之间独立计算,这种方式可以有效地减少模

型复杂度,带来计算效率上的提升,但这也导致跨组的信息交互受限,模型无法充分利用整体信息进行特征学习,这对于识别航拍架空输电线路鸟巢图像并不利。

在 FasterNet 神经网络中,香港科技大学运用了部分卷积(partial convolution, PConv)^[29]操作,旨在减少计算冗余和内存访问,其主要原理如图 3 所示。

由图 3 可知,PConv 只需要对输入的一部分通道上应用 Conv 来进行空间特征提取,并保持其余通道不变,在处理连续的通道规则访问时,将第一个或最后一个连续通道视为整个特征图的代表进行计算,有限提取特征的同时减少了计算量和内存的访问,计算公式如式(1)所示。

$$h \times w \times k^2 \times c_p^2 \quad (1)$$

式中: h, w 为特征图的宽和高, k 为卷积核的大小, c_p 是常规卷积作用的通道数。PConv 的内存访问情况如式(2)所示。

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (2)$$

式中: h, w 为特征图的宽和高, k 为卷积核的大小, c_p 为常规卷积作用的通道数。

2.2 ConvMixer

在计算机视觉任务中,卷积神经网络一直是深度学习系统常用的主导结构,但近年来,基于 Transformer 模型的体系结构,例如视觉 Transformer (vision transformer, ViT)^[30]

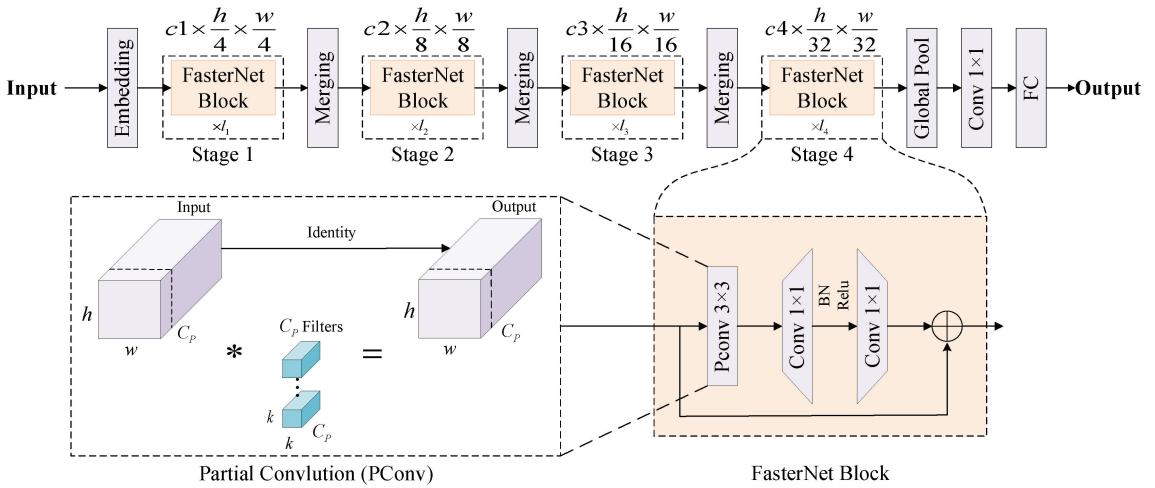


图3 FasterNet 架构和 PConv 工作原理

架构在此类任务中表现出良好的性能,通常优于经典卷积体系结构。ConvMixer^[31]是最近计算机视觉应用中出现的一种新型架构,它最初应用于验证 ViT 架构性能是否归因于补丁的使用,其主要思想是采用分离空间和通道维度的混合来提升性能。尽管它的结构简单,但实验表明该架构相比于 ViT 架构更高效,同时占用更少的计算时间和资源空间。ConvMixer 架构如图 4 所示,ConvMixer Layer 层由深度卷积 (depthwise convolution) 和逐点卷积 (pointwise convolution) 以及残差连接组成,在每一个卷积操作之后,会接一个激活函数 GELU 和批量归一化 BatchNorm,而残差连接则用于跨层的直接连接来传递信息,帮助模型减轻梯度消失问题,加快模型的快速收敛。

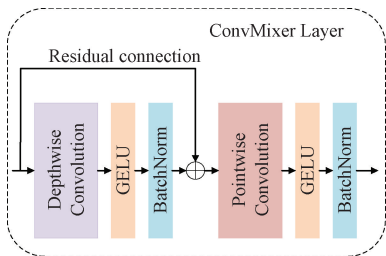


图4 ConvMixer 架构图

针对架空输电线路鸟巢检测问题,在本文改进模型的特征融合网络 C3 模块后集成 ConvMixer 架构增强小目标检测性能。ConvMixer 的设计有助于在特征图像中更好的捕捉空间和通道之间的关系,而 ConvMixer 在通道维度上的特征融合,促进了特征之间的交互和信息传递,从而提高模型对小目标图像特征的理解能力。

2.3 全维动态卷积

基于无人机航拍鸟巢图像存在小目标,复杂背景,推理速度慢等问题,本文模型引入 Intel 实验室姚安邦团队提出的全维动态卷积 (omni-dimensional dynamic convolution, ODConv)^[32],用以对基线模型的特征融合网络部分进行改

进,实现尽可能的提升网络性能的同时不增加网络参数量以及计算量。

全维动态卷积采用一种新颖的多维注意力机制,在所有卷积层沿着卷积核的 4 个维度进行 (卷积内核数、卷积核空间大小、输入信道数、输出信道数) 以并行的策略进行注意力加权,动态地对不同类型的关注点来分配不同权重,提升卷积核的自适应性能,以此来提升卷积神经网络的特征提取能力^[33],其结构图如图 5 所示。相较于传统卷积,全维动态卷积具有更强的灵活性和表达能力。

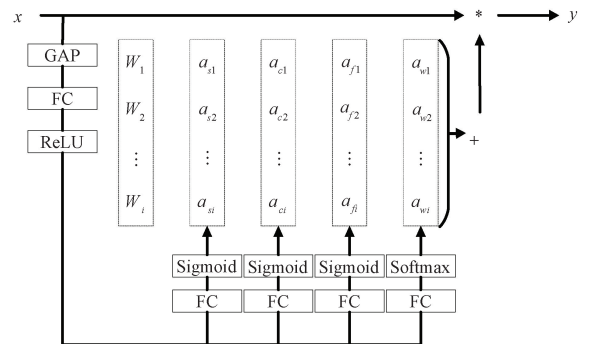


图5 ODConv 模块结构图

由图 5 可知,将输入 x 通过通道级全局平均池化 (global average pooling, GAP) 操作压缩为长度为 c_{in} 的特征向量,然后经过全连接层 (fully connected, FC) 和 ReLU 激活函数进行进一步处理。其中全连接层会将压缩后的特征向量映射到较低维空间。4 个分支头中每一个分支都有一个输出大小为 $k \times k, c_{in} \times 1, c_{out} \times 1, n \times 1$ 的全连接层,第 1~3 分支经 Sigmoid 激活函数生成归一化注意力标量 a_{si}, a_{ci}, a_{fi} ; 第 4 个分支经 Softmax 激活函数生成归一化注意力标量 a_{wi} 。其中 w_i 表示卷积核, a_{si} 是空间位置分配的关注标量, a_{ci} 是输入通道分配的关注标量, a_{fi} 是输出通道的关注标量, a_{wi} 是整个卷积核的注意力标量。通过全维动态卷积增强卷积神经网络的基本卷积运算的特

征提取能力,可有效弥补目标信息量不足的缺陷。

2.4 改进模型整体网络结构

改进后算法网络模型架构及模块参数如表 1 所示。

表格中的来源代表本层来自哪一层的输出, -1 代表来自上一层的输出;参数设置包括本层所用模块的输入通道数、输出通道数、卷积核尺寸及步长等信息。

表 1 改进网络整体结构

序号	来源	参数量	模块名称	参数设置
0	-1	1 200	PatchEmbed_FasterNet	[3, 24, 4, 4]
1	-1	2 724	BasicStage	[24, 1]
2	-1	3 920	PatchMerging_FasterNet	[24, 40, 2, 2]
3	-1	7 460	BasicStage	[40, 1]
4	-1	12 960	PatchMerging_FasterNet	[40, 80, 2, 2]
5	-1	88 560	BasicStage	[80, 1]
6	-1	51 520	PatchMerging_FasterNet	[80, 160, 2, 2]
7	-1	117 440	BasicStage	[160, 1]
8	-1	64 480	SPPF	[160, 160, 5]
9	-1	20 736	Conv	[160, 128, 1, 1]
10	-1	0	Upsample	[None, 2]
11	[-1,5]	0	Concat	[1]
12	-1	64 192	C3	[208, 128, 1, False]
13	-1	8 320	Conv	[128, 64, 1, 1]
14	-1	0	Upsample	[None, 2]
15	[-1,3]	0	Concat	[1]
16	-1	16 224	C3	[104, 64, 1, False]
17	-1	9 664	ConvMix	[64, 64]
18	-1	40 377	ODConv	[64, 64, 3, 1, 1]
19	-1	36 992	Conv	[64, 64, 3, 2]
20	[-1,13]	0	Concat	[1]
21	-1	53 952	C3	[128, 128, 1, False]
22	-1	27 520	ConvMix	[128,128]
23	-1	154 297	ODConv	[128, 128, 3, 1, 1]
24	-1	147 712	Conv	[128, 128, 3, 2]
25	[-1,9]	0	Concat	[1]
26	-1	214 400	C3	[256, 256, 1, False]
27	-1	87 808	ConvMix	[256, 256]
28	-1	603321	ODConv	[256, 256, 3, 1, 1]

3 实验结果与数据分析

3.1 数据集构建与处理

本实验数据集由 3 部分组成:第 1 部分采用大疆 Phantom 4 RTK 型无人机采集的架空输电线路图像数据,对采集到的图像进行人工筛选保留其中有效图像后,共得到架空输电线路鸟巢图像 590 张;第 2 部分在百度飞浆公开的 200 张输电线路鸟巢数据集中选取 132 张小目标鸟巢图像;第 3 部分为公开网络收集到的 243 张架空输电线路鸟巢图像。本文构建的数据集包括复杂背景、小目标鸟巢、遮挡等多种场景,不同场景下的鸟巢图像如图 6 所示。

为避免数据集图像过少以及多样性不足,影响其模型训练效果和泛化性能^[34],本文对数据集采用仿射变换和添加噪声的方式进行数据扩充,扩充后的部分数据集图像如图 7 所示,经过扩充及处理后的数据集共 4 543 幅图像。使用 LabelImg 图像标注软件对架空输电线路上的鸟巢目标进行标记,并将图像数据与标签文件按照 YOLO 数据集格式进行转换并整合,整合后的数据集按 82%/12%的比例划分为训练集和测试集。

由图 8 长宽分布图可知,本文构建的鸟巢数据集目标的长宽分布在[0.0~0.2]范围内颜色较深,说明本文鸟巢目标数量在此区域数量最多,且多数为小目标特征。

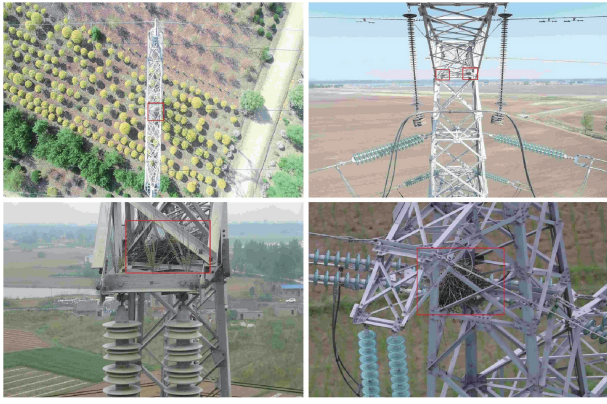


图6 不同场景下的鸟巢图像



图7 扩充后的部分数据集图像

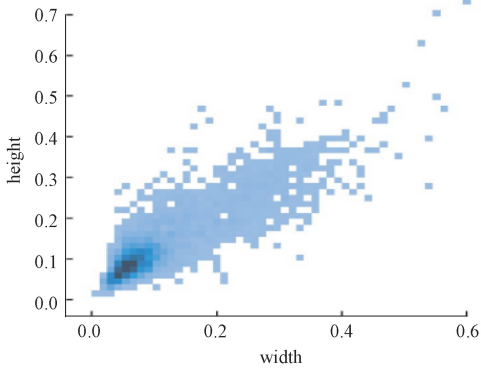


图8 长宽分布图

3.2 实验环境及参数设置

本实验租用云服务器,该服务器平台处理器型号为 24 vCPU AMD EPYC 7642 48-Core Processor,使用 NVIDIA RTX3090 显卡。显存为 24 GB,深度学习框架选择 Pytorch,版本为 1.9.0,CUDA 版本为 11.1,Python 版本为 3.7。训练过程中,批量大小设置为 16,进程 works 设置为 16,训练轮数设置为 300 轮,后续不同算法对比试验中也将保持统一变量。实验参数设置如表 2 所示。

表2 实验参数设置

训练参数	参数设置
初始学习率(lr0)	0.01
动量参数(momentum)	0.937
权重衰减参数(weight_decay)	0.000 5
模型预热动量(warmup_momentum)	0.8
预热批数(warmup_epochs)	3.0

3.3 评价指标

为通过检测结果来评价改进后 YOLOv5s 模型的性能,利用准确率(precision,P)、召回率(recall,R)、平均精度均值(mean average precision,mAP)作为检测性能的评价指标;采用参数量(parameters)、计算量(computation)、每秒传输帧数(frames per second,FPS)、模型体积作为检测效率的评价指标。

3.4 实验结果分析

为测试轻量化主干网络 Fasternet 与 ConvMixer 架构相结合对模型性能的影响,本文将 ConvMixer 架构分别应用进 YOLOv5s-Fasternet 网络的不同层级中,并从 mAP、参数量、计算量 3 个方面来进行测试评估,实验结果如表 3 所示。

在表 3 中,实验 2 表示在 YOLOv5s-Fasternet 模型主干网络的第 7 层 BasicStage 模块后应用此架构;实验 3 表示在主干网络的第 9 层 SPPF 特征金字塔模块后应用此架构;实验 4 表示在颈部网络的第 13 层 C3 特征提取模块后应用此架构;实验 5 表示在颈部网络的第 13、18、22、26 层 4 个 C3 特征提取模块后应用此模块;实验 6 表示在颈部网络的第 17、21、25 层 3 个 C3 特征提取模块后应用此模块。

表3 ConvMixer 架构替换位置对比结果

实验	模型	mAP/%	Parameters/M	Computation/G
1	YOLOv5-Fasternet	95.1	1.046	2.2
2	ConvMixer_backbone1	94.0	1.086	2.2
3	ConvMixer_backbone2	95.0	1.086	2.2
4	ConvMixer_neck1	94.1	1.074	2.3
5	ConvMixer_neck2	95.0	1.207	2.6
6	ConvMixer_neck3	95.9	1.043	2.2

由表 3 可知,将 ConvMixer 架构应用进颈部网络的第 17、21、25 层后,较为明显地提升模型的检测精度,且最小限度的增加参数数量和计算量。

为对比嵌入不同位置的 ODConv 全维动态卷积模块

对结合 Fasternet 与 ConvMixer 改进模型的影响,本文将 ODConv 模块分别应用进 YOLOv5-Fasternet + ConvMixer 网络的不同层级中,并从 mAP、参数量、计算量 3 个方面进行测试对比,实验结果如表 4 所示。

表 4 ODConv 模块替换位置试验结果

实验	模型	mAP/%	Parameters/M	Computation/G
1	YOLOv5-Fasternet+ConvMixer	95.9	1.043	2.2
2	ODConv__backbone1	95.4	1.426	2.5
3	ODConv__backbone2	95.5	1.426	2.5
4	ODConv_neck1	96.0	1.325	2.5
5	ODConv_neck2	96.1	2.132	2.5
6	ODConv_neck3	95.5	1.194	2.2
7	ODConv_neck4	96.3	1.841	2.2
8	ODConv_neck5	96.4	1.841	2.2

在表 4 中,实验 2 表示在 YOLOv5-Fasternet + ConvMixer 模型主干网络的第 7 层 BasicStage 模块后应用此模块;实验 3 表示在主干网络的第 9 层 SPPF 特征金字塔模块后应用此模块;实验 4 表示在颈部网络的第 14 层 C3 特征提取模块后应用此模块;实验 5 表示在颈部网络的第 13、19、24、29 层 C3 特征提取模块以及 ConvMixer 架构后应用此模块;实验 6 表示在颈部网络的第 9、13、18、22 层替换 Conv 卷积模块为 ODConv 全维动态卷积模块;实验 7 表示在颈部网络的第 17、21、25 层 3 个 C3 特征提取模块后

应用此模块;实验 8 表示在颈部网络的第 18、23、28 层 3 个 ConvMixer 架构后应用此模块。分析表 4 可知,实验 8 在平均精度均值上较其他实验提升的最多达到 96.4%,且参数量仅增长 0.798 M,最终确认在颈部网络的第 18、23、28 层应用此模块。

为验证本文所提改进策略的有效性,本文以 YOLOv5s 作为基线模型,进行多种不同改进策略之间的消融实验:1)FasterNet 轻量化主干网络;2)ConvMixer 架构;3)ODConv 模块。消融对比实验结果如表 5 所示。

表 5 消融对比实验结果

实验	FasterNet	ConvMixer	ODConv	P/%	R/%	mAP/%	Parameters/M	Computation/G	模型体积/MB
1				93.9	92.8	95.5	7.012	15.8	14.4
2	✓			95.8	92.1	95.1	1.046	2.2	2.4
3		✓		95.0	91.9	95.7	7.434	16.6	15.3
4			✓	95.2	91.5	95.9	10.181	16.0	20.8
5	✓	✓	✓	97.2	93.0	96.4	1.841	2.2	4.0

由表 5 可知,与基线模型相比,实验 2 在主干位置引用 FasterNet 轻量化主干网络后计算量、参数量以及模型大小均大幅减少,mAP 下降 0.4%,验证本文提出的轻量化主干网络能够极大的压缩模型体积,并保持一定的检测精度;实验 3 引入 ConvMixer 架构后,mAP 提升了 0.2%,这是因为 ConvMixer 架构包含全局注意力机制和多层感知器,能够更好的捕捉图像中全局和局部特征,提高模型对于小目标局部特征的学习能力;实验 4 在颈部嵌入 ODConv 模块,模型的 mAP 值提高了 0.4%,这是因为 ODConv 能够精准定位鸟巢局部颜色特征,提高改进模型对复杂背景下鸟巢的检测精度;在实验 5 中本文提出的优化模型 mAP 值提高 0.9%,精确率上升 3.3%,召回率上升 0.2%,参数量下降 74%,计算量下降 86%,模型体积下降 72%。可以看出本文模型的参数量、计算量以及模型体

积大幅下降,并且检测精度有一定的涨幅。综上所述,本文模型满足轻量化模型部署的要求,并保持较高的检测精度。mAP 值曲线变化对比图如图 9 所示。

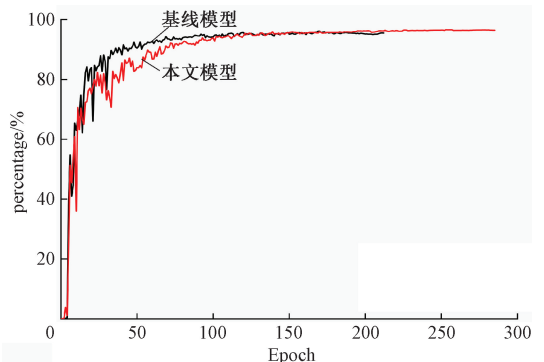


图 9 mAP 值曲线变化对比图

在2021年9月5日以后YOLOv5更新了早停机制,当模型在训练过程中性能未获得提升,自动提前终止训练。由图9可知,基线模型训练至212轮触发早停机制;本文模型在训练过程中mAP值稳定提升直至300轮。两个实验模型mAP值最终都趋于收敛,但在训练后期本文模

型在mAP值上的提升更为明显且稳定。

为进一步验证本文改进策略的有效性,将消融实验模型在测试集上进行可视化性能比对,选取复杂背景、小目标鸟巢、遮挡3种场景的航拍图像来展示其改进性能,实验结果如图10所示。

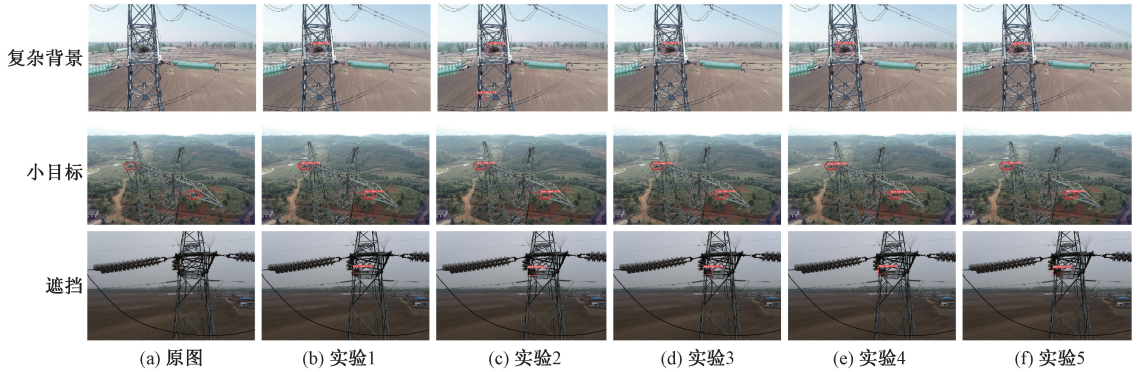


图10 消融实验模型结果对比

由图10(b)可知,基线模型对小目标鸟巢和复杂背景的特征提取能力较弱,置信度相较于本文改进后的模型偏低;由图10(c)可知,实验2将YOLOv5s的CSP-Darknet主干特征提取网络替换为FasterNet轻量化主干网络后,存在较为明显的误检现象以及精度损失;由图10(d)可知,实验3在YOLOv5s特征融合网络中引用ConvMixer架构增强多尺度小目标辨识能力,有效提取图像中小目标的局部特征,获得较好的检测能力;由图10(e)可知,实验4中的ODConv动态卷积模块细化鸟巢的细粒度特征,有效的提取复杂背景以及遮挡场景下的鸟巢颜色特征,提高了模

型的检测效果;由图10(f)可知,本文改进的整体模型相比于基线模型具有较好的检测效果,在复杂背景、小目标以及遮挡场景中均表现出良好的检测精度,且轻量化主干特征提取网络后的误检问题有较大的改善,以上数据结果以及可视化性能对比实验均说明本文改进模型有效性。

为了更深入评估本文改进模型的检测性能,将当前主流目标检测算法YOLOv3、YOLOv5s、YOLOv7、YOLOv8和YOLOv9等模型进行比较测试。所有模型均在同一台训练设备下采用控制变量原则进行实验,不同目标检测模型性能实验结果对比如表6所示。

表6 不同目标检测模型性能实验结果对比

模型	mAP/%	Parameters/M	Computation/G	FPS(帧/s)
YOLOv3	95.0	61.497	154.5	48.5
YOLOv5s	95.5	7.012	15.8	108.7
YOLOv5s-Ghost+CA+C3CBAM ^[35]	95.7	5.017	10.4	66.7
YOLOv7	95.6	36.481	103.2	53.7
YOLOv8	99.3	3.009	8.2	105.5
YOLOv9	95.4	60.756	264.9	62.5
YOLOv5s-FasterNet+ConvMixer+ODConv	96.4	1.841	2.2	104.2

由表6可知,本文提出基于YOLOv5s的轻量化鸟巢检测网络模型的平均精度均值mAP达到96.4%,相较于YOLOv3、YOLOv5s、YOLOv7和YOLOv9分别提高1.4%、0.9%、0.8%、1%,低于YOLOv8模型2.9%;虽然在精度上不及YOLOv8,但本文提出的改进模型参数数量和计算量最小,参数量仅为1.841M,与上述模型相比,分别减小了59.656M、5.171M、34.64M、1.168M、58.915M;计算量为2.2GFLOPs,与上述模型相比分别减少了99%、86%、98%、73%、99%。从模型推理速度上来看,

YOLOv5s的推理速度最优达到108.7帧/s,本文提出的改进模型虽然不及YOLOv5s,但仅低4.5帧/s,在参数规模和推理速度上取得较好的平衡,满足工程应用的需求。将本文模型与张学立等^[35]所提的轻量化改进方案进行比较,可以看出本文模型在各项评价指标中都表现出了更好的性能,进一步验证本文改进模型的可靠性。为了更加直观地展示本文改进模型对架空输电线路鸟巢的关注度,选择YOLOv3、YOLOv7、YOLOv9和本文改进的YOLOv5s模型进行比较,不同目标检测模型实验结果如图11所示。

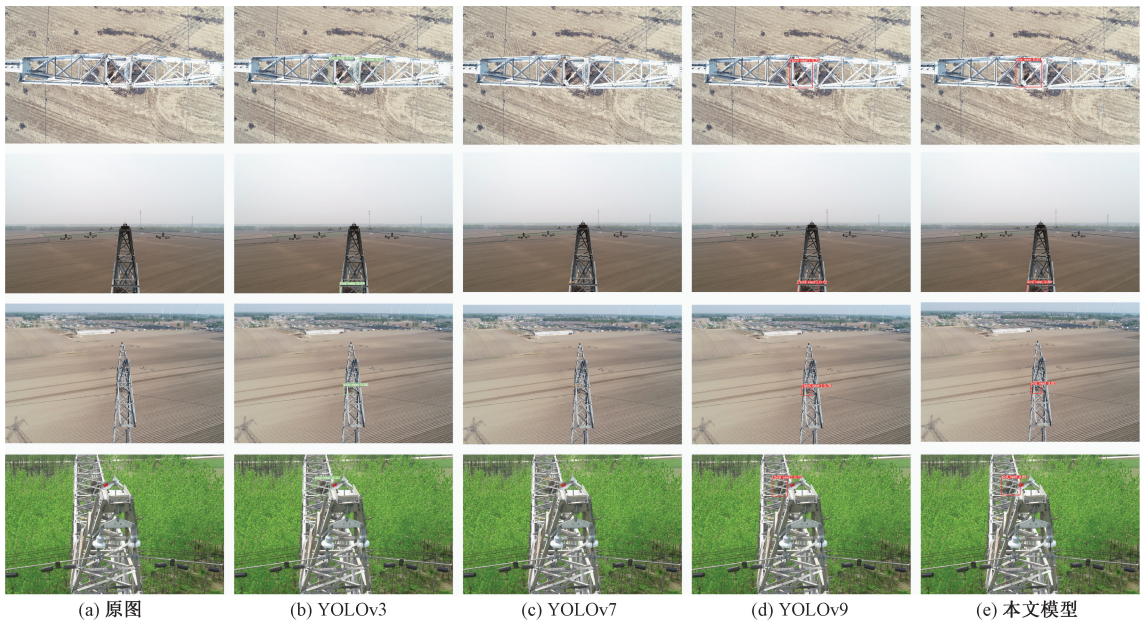


图 11 不同目标检测模型实验结果对比图

由图 11(b)可知, YOLOv3 网络模型对鸟巢目标的检测过程中出现误检情况, 置信度较低, 相比于其他模型不能准确完成对本文鸟巢目标的识别定位; 由图 11(c)和(d)可知, YOLOv7 网络模型和 YOLOv9 网络模型能够识别多种场景下的鸟巢目标, 没有出现误检和漏检情况, 具有较好的识别效果; 由图 11(e)可知, 本文模型可以实现对鸟巢目标的精准识别, 在置信度上高于 YOLOv3、YOLOv7 以及 YOLOv9 网络模型。根据不同目标检测模型性能实验数据以及实验对比图像可以看出, 本文提出的改进模型在参数数量以及计算复杂度上在主流算法模型中最低; 平均精度均值达到 96.4%, 能够准确识别出鸟巢目标, 并保持较高的置信度; 推理速度达到 104.2 帧/s, 满足对航拍架空输电线路鸟巢图像进行实时检测的需求。

4 结 论

本文针对架空输电线路鸟巢目标检测任务的需求, 提出一种基于 YOLOv5s 的轻量化检测网络, 本文主要结论如下: 通过重构 YOLOv5s 主干网络, 引入轻量化 FasterNet 网络作为主干特征提取网络, 有效降低了模型的计算复杂度和参数数量, 并能保持一定的检测精度。针对应用 FasterNet 网络导致的精度下降以及模型对小目标检测精度低的问题, 在颈部集成 ConvMixer 层, 保证参数数量以及模型体积的情况下, 增强模型对小目标鸟巢的检测能力。融合 ODConv 模块, 使得特征融合网络能够充分的捕捉上下文信息, 帮助网络更好的理解整体图像语境, 抑制复杂背景对目标的干扰, 提升模型对复杂背景以及小目标的检测性能。本文实验结果表明, 本文提出的改进模型平均精度均值 mAP 为 96.4%, 参数量为 1.841 M, 计算量

为 2.2 GFLOPs, 模型体积为 4.0 MB, 检测速度为 104.2 帧/s。本文改进模型相比于 YOLOv5s 虽然检测速度略慢 4.5 帧/s, 但大幅降低模型体积, 能更好的平衡模型的检测精度和参数规模, 可以有效对航拍鸟巢图像进行检测。本文所作研究有一些不足, 经过实验测试可知, 本文改进模型可以对航拍架空输电线路鸟巢图像进行精确定位, 但对于一些背景复杂或目标被遮挡严重的图像识别精度不高, 仍有一定的提升空间。未来研究在保证网络参数规模和推理速度的情况下, 进一步提升在复杂背景和遮挡场景任务中的检测表现。

参考文献

- [1] 李帆, 李阳林, 张宇, 等. 架空输电线路涉鸟故障分析与防范[J]. 中国电力, 2019, 52(10): 92-99.
- [2] 王胜辉, 郭奉天, 董兴浩, 等. 交流电压下架空输电线路巢材类异物放电特性研究[J]. 华北电力大学学报(自然科学版), 2022, 49(5): 24-32.
- [3] 黄宇辰, 房俊龙. 架空输电线路用防鸟粪绝缘护套配置方法的研究[J]. 绝缘材料, 2023, 56(3): 115-121.
- [4] 张宇, 况燕军, 陈铭业, 等. 架空输电线路绝缘子在典型鸟粪污染条件下的闪络特性[J]. 高电压技术, 2018, 44(6): 1800-1805.
- [5] 王少华, 叶自强. 架空输电线路鸟害故障及其防治技术措施[J]. 高压电器, 2011, 47(2): 61-67.
- [6] 隋宇, 宁平凡, 牛萍娟, 等. 面向架空输电线路的挂载无人机电力巡检技术研究综述[J]. 电网技术, 2021, 45(9): 3636-3648.
- [7] 麻卫峰, 王成, 王金亮, 等. 基于激光点云的高压输电线

- 覆冰厚度反演[J]. 电力系统保护与控制, 2021, 49(4): 89-95.
- [8] 赵霖, 王素珍, 邵明伟, 等. 基于改进YOLOv5的输电线路鸟巢缺陷检测方法[J]. 电子测量技术, 2023, 46(3): 157-165.
- [9] 肖繁俊, 潘睿志, 李超, 等. 基于改进YOLOv5s绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.
- [10] 段旺旺, 唐鹏, 金炜东, 等. 基于关键区域HOG特征的铁路接触网鸟巢检测[J]. 中国铁路, 2015, (8): 73-77.
- [11] 祝振敏, 谢亮凯. 基于相对位置不变性的接触网鸟巢识别检测[J]. 铁道科学与工程学报, 2018, 15(4): 1043-1049.
- [12] WU X, YUAN P, PENG Q, et al. Detection of bird nests in overhead catenary system images for high-speed rail [J]. Pattern Recognition, 2016, 51(C): 242-254.
- [13] REN S, HE K, 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.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [15] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE Press, 2017: 2980-2988.
- [16] 王纪武, 罗海保, 鱼鹏飞, 等. 基于Faster R-CNN的多尺度高压塔鸟巢检测[J]. 北京交通大学学报, 2019, 43(5): 37-43.
- [17] 刘国文, 张彩霞, 李斌, 等. 基于改进RetinaNet模型的接触网鸟巢检测[J]. 数据采集与处理, 2020, 35(3): 563-571.
- [18] 丁建, 黄陆明, 朱迪锋, 等. 面向高空塔架鸟巢检测的双尺度YOLOv3网络学习[J]. 西安理工大学学报, 2021, 37(2): 253-260.
- [19] 赵霖, 王素珍, 邵明伟, 等. 基于改进YOLOv5的输电线路鸟巢缺陷检测方法[J]. 电子测量技术, 2023, 46(3): 157-165.
- [20] 张焕龙, 齐企业, 张杰, 等. 基于改进YOLOv5的输电线路鸟巢检测方法研究[J]. 电力系统保护与控制, 2023, 51(2): 151-159.
- [21] 赵立杰, 袁昌彪, 黄明忠, 等. 基于通道剪枝的ACAM-YOLOv5s绝缘子缺陷检测[J]. 电子测量技术, 2023, 46(9): 108-116.
- [22] 肖繁俊, 潘睿志, 李超, 等. 基于改进YOLOv5s绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.
- [23] 郑婷婷, 周浩, 王秋忆. 基于改进YOLOv5的电力设备检测算法[J]. 电子测量技术, 2023, 46(4): 155-160.
- [24] ZHOU F B, ZHAO H L, NIE Z. Safety helmet detection based on YOLOv5 [C]. 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), January 22-24, 2021, Shenyang, China: 6-11.
- [25] FENG Z Q, GUO L, HUANG D R, et al. Electrical insulator defects detection method based on YOLOv5[C]. 2021 IEEE 10th Data Driven Control and Learning Systems Conference(DDCLS), May 14-16.
- [26] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA.
- [27] POUDEL R P K, BONDE U, LIWICKI S, et al. Contextnet: Exploring context and detail for semantic segmentation in real-time[J]. ArXiv Preprint, 2018, ArXiv:1805.04554.
- [28] CHEN J, LEI B, SONG Q, et al. A hierarchical graph network for 3d object detection on point clouds [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 392-401.
- [29] MA X, GUO F M, NIU W, et al. Pconv: The missing but desirable sparsity in dnn weight pruning for real-time execution on mobile devices [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5117-5124.
- [30] ALEXEY D, LUCAS B, ALEXANDER K, et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale [EB/OL]. (2021-06-03) [2023-11-20] <https://doi.org/10.48550/arXiv.2010.11929>.

- [31] TROCKMAN A, KOLTER J Z. Patches are all You need? [EB/OL]. 2022. ArXiv:2201.09792. <https://arxiv.org/abs/2201.09792.pdf>.
- [32] LI C, ZHOU A, YAO A. Omni-dimensional dynamic convolution [J]. ArXiv Preprint, 2022, ArXiv:2209.07947.
- [33] CHEN Y, DAI X, LIU M, et al. Dynamic convolution: Attention over convolution kernels[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11030-11039.
- [34] 张珂, 余锴镭, 石怀涛. 改进 YOLOv5 的绝缘子旋转目标检测[J]. 电子测量技术, 2023, 46(15):52-58.
- [35] 张学立, 贾新春, 王美刚, 等. 安全帽与反光衣的轻量化检测:改进 YOLOv5s 的算法[J]. 计算机工程与应用, 2024, 60(1):104-109.

作者简介

徐业东, 硕士研究生, 主要研究方向为图像处理与目标检测。

E-mail:1326090734@qq.com

蔡亚恒, 硕士研究生, 主要研究方向为采摘机器人的关键技术研究。

E-mail:1652335999@qq.com

李严, 硕士研究生, 主要研究方向为电缆隧道巡检机器人位姿估计算法研究。

E-mail:849371898@qq.com

刘学雷, 硕士研究生, 主要研究方向为电力需求响应。

E-mail:2673145020@qq.com

曹英丽(通信作者), 教授, 博士生导师, 主要研究方向为农业信息化研究。

E-mail:caoyingli@163.com