

刘畅. 基于 ACE-YOLOv10 的无人机船舶检测算法[J]. 智能计算机与应用, 2025, 15(3): 152-157. DOI: 10. 20169/j. issn. 2095-2163. 24090202

基于 ACE-YOLOv10 的无人机船舶检测算法

刘 畅

(西安石油大学 计算机学院, 西安 710065)

摘 要: 无人机由于其成本低廉、敏捷、灵活并且可以搭载高分辨率摄像头和传感器,正在成为各个行业不可或缺的工具。在检测环境中,无人机可以适应复杂地形以及恶劣气候,收集到大气、土壤等各项珍贵数据。在工业探索领域,无人机可以代替人工对于危险地域的探查,降低事故的发生率。在海上救援方面,无人机可以快速定位到遇险的船舶,提升救援人员的救援效率。然而,无人机在此情况下的目标检测任务仍然面临不小挑战,如目标被障碍物遮挡、太阳光照的变化等。为了提高无人机在海上救援中的船舶检测性能,本文提出了一种基于改进 YOLOv10 (You Only Look Once, YOLO) 的目标检测算法,加入了 CA 注意力机制,并采取 $EIoU$ 损失函数。实验结果表明,相对于原始的 YOLOv10 模型,实验数据集在 ACE-YOLOv10 模型中获得了更好的结果。

关键词: 无人机; YOLOv10; 目标检测; CA 注意力机制; $EIoU$ 损失函数; 海上救援

中图分类号: TP751

文献标志码: A

文章编号: 2095-2163(2025)03-0152-06

Drone ship detection algorithm based on ACE-YOLOv10

LIU Chang

(College of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

Abstract: Due to its low cost, agility, flexibility, and the ability to carry high-resolution cameras and sensors, drones are increasingly becoming indispensable tools in various industries. In environmental detection, drones can adapt to complex terrains and harsh climates, and collect valuable data such as the atmosphere and soil. In the field of industrial exploration, drones can replace manual exploration of dangerous areas and reduce the incidence of accidents. In terms of maritime rescue, drones can be utilized to swiftly locate ships in distress, enhancing the rescue efficiency of rescue teams. However, the target detection tasks for drones in such scenarios still face considerable challenges, such as targets being obscured by obstacles, variations in sunlight, etc. To enhance the detection capabilities of drones in dealing with complex situations during maritime rescues, this paper proposes a target detection algorithm based on the improved YOLOv10 (You Only Look Once, YOLO), adding the CA attention mechanism and adopting the $EIoU$ loss function. Experimental results show that compared with the original YOLOv10 model, the experimental dataset obtains better results in the ACE-YOLOv10 model.

Key words: Unmanned Aerial Vehicle; YOLOv10; object detection; CA attention mechanism; $EIoU$ loss function; maritime rescue

0 引 言

目标检测^[1-4]是深度学习中机器视觉领域的一个核心任务,重点是在图像或是视频中准确无误地定位识别多个不同的对象。该技术分为 2 个必要的阶段:首先通过深度学习中的模型,如卷积神经网络 (Convolutional Neural Networks, CNN) 提取图像的特征,确定目标的位置;接着对这些位置进行分类,区分目标的类别。随着 YOLO、快速区域卷积神经网络

络^[5] (Region-based Convolutional Neural Network, R-CNN)、单次多框检测器 (Single Shot MultiBox Detector, SSD) 等算法的进步,目标检测对不同对象在定位的精度和分类的准确度上均有着明显的提高,并且在机器人导航、医疗影像分析、智能监控等领域都有广泛应用。

目前,很多研究人员对视频图像的人员检测算法进行了研究。如:李霄峰^[6] 为了降低模型的误报率,融合了目标分支和运动分支。张永强等学者^[7]

为了提高对人员检测的准确率和检测速度,采用 ResNet50 替换了原始的 CSPDarknet53 改进 YOLOv4。师后勤等学者^[8] 为了提高人员在低光照下的检测效果,将 Retinex 图像增强的算法和注意力机制同时融入到 YOLOv3 中。梁京涛等学者^[9] 为了提升网络的特征融合能力,在主干网络和颈部网络中添加了注意力机制。孙日明等学者^[10] 尝试融合不同维度差异性和高维数字序列整体性,提高跌倒检测的准确率。为了提高原 YOLOv7 模型中 PAnet 在特征融合方面的性能,赵江平等学者^[11] 加入了 BiFPN 结构,提出的算法模型轻量且高效,可以更加准确的利用和协调来自不同尺度的特征信息。Zhang 等学者^[12] 融合了主干网络所提取的语义特征信息和细粒度,提出了跨层多尺度的特征融合网络。该网络提高了相邻特征层之间的空间位置信息。

为了提高检测的精度,张铭泉等学者^[13] 在 Faster-RCNN 的基础上设计了一种能够弥补变电站漏油缺陷的小目标检测算法。该模型取消了网络中的下采样操作的一部分,并用小卷积核替换大卷积核,但模型的体积过大,对于计算机的硬件要求比较高。Hao 等学者^[14] 改进特征提取结构,加入了小目标检测层来提高多尺度下目标检测的性能,但仍然存在着一些延迟。对于提高模型目标检测的性能,还需要做进一步的研究工作。

1 相关模型简介

1.1 YOLOv8 模型

由骨干网络 (Backbone)、颈部网络 (Neck)、预测头 (Head) 等关键部分组成的 YOLOv8 在 2023 年由 Joseph Redmon 团队提出^[4]。其中,在骨干网络中结合 EfficientNet 的架构理念,使用一种轻量级的网络结构 EfficientNetV2,降低了计算量的同时也保证了准确率。为了提高模型对于不同尺度对象的检测水平,研究人员在 YOLOv8 中添加了更有效的特征融合机制,如改进后的路径聚合网络 (Path Aggregation Network, PAnet) 版本,使模型可以在不同层级之间进行特征融合。

骨干网络部分主要用于提取图像中的底层特征,将有利于目标检测。颈部网络处在骨干网络和检测头中间,负责对骨干网络所提取的特征进行更进一步的融合。不同尺度的特征融合能够提供更加全面的信息,从而有利于模型对不同的形状和大小的对象进行检测。检测头也是 YOLOv8 模型的重要

部分,一般用于生成目标检测时的边框和实现对象类别的预测。在检测头部分,把原始模型的耦合头替换成了解耦头的结构,把分类任务和回归任务分隔开。同时,YOLOv8 用 Anchor-free 替换掉 Anchor-Based,简化了计算量,提升了泛化能力,使得模型拥有更强的适应性。最后,检测头的输出会和真实标签做对比,再使用损失函数计算预测的误差,从而有助于对模型进行后续不断的优化升级。

YOLOv8 的原始网络结构如图 1 所示。其中,数据在输入到模型之前,需要经过归一化、缩放等一系列的预处理步骤,以方便数据的后续加工分析。

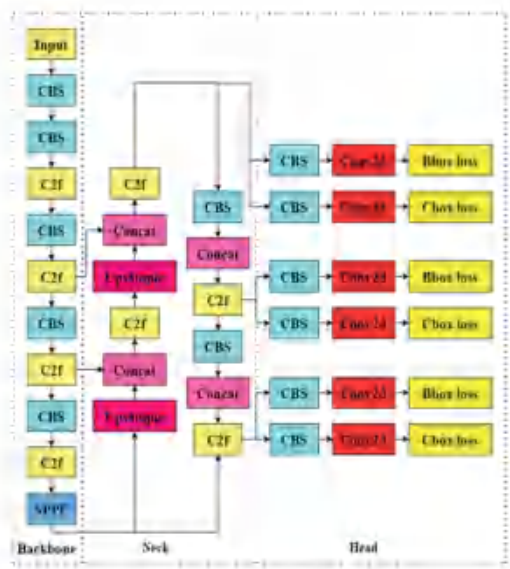


图 1 YOLOv8 网络结构

Fig. 1 The network architecture of YOLOv8

1.2 YOLOv10 模型

YOLOv10^[15-16] 在 2024 年由清华大学的研究人员开发,是基于 YOLOv8 的新版本。YOLOv10 在实时的目标检测方面性能得以显著提升,分析可知其主要特点如下。

(1) 整体效率-精度驱动设计: YOLOv10 优化了原始模型中的各个模块,在效率和精度方面减少了计算时的冗余,从而提升了参数的利用效率。

(2) 无 NMS 训练: YOLOv10 使用了双重分配策略对数据进行无非极大值抑制 (Non - Maximum Suppression, NMS) 训练。这种训练策略结合了一对一与一对多的标签分配,使推理过程摒弃了对 NMS 的需求。

(3) 性能和效率: YOLOv10 在精度和速度上都比之前的版本有所优化。目前, YOLOv10 有着十多种型号。对于 COCO 数据集,在保持了相似的精确

度下, YOLOv10-S 的计算速度比较 RT-DETR-R18 快了 1.8 倍, YOLOv10-B 比 YOLOv9-C 在相同条件下减少了 46% 的推理延迟以及 25% 的参数, 优化程度较为显著。

(4) 架构增强: YOLOv10 采取了较为紧凑的倒置模块紧凑反转块 (Compact Inverted Block, CIB) 来提升特征提取能力, 同时也能实现计算成本最小化。YOLOv10 还融合了空间-通道解耦降采样, 这样不仅能够保留更全面的信息, 还能提升降采样的效率。

虽然 YOLOv10 在实际应用中展现了诸多优点, 但也要看到: 一方面, 随着模型特征图的下采样, 特征图的空间分辨率也逐渐降低, 对提取特征不利; 另一方面, 对于较小的目标, 由于输入时的特征图分辨率过高, 可能无法对目标进行有效的检测。而对于较大的目标, 特征分辨率又较低, 也无法获得足够多的数据信息。YOLOv10 的原始网络结构如图 2 所示。

小的卷积核产生初始的采样目标。

(2) 灵活的参数数量: 固定的尺寸和形状是传统的卷积核所具有的特点, 如 2×2 或是 4×4 的正方形网格。而 AKConv 可以允许卷积核拥有任意数量的参数, 使得卷积核的形状不再受限于标准的正方形, 而是能够更加贴合任务的要求和目标本身的特征, 采取更加灵活多变的形状, 提高效率。

(3) 动态调整卷积核大小: 除了能够自适应采样的形状, AKConv 使用坐标生成算法来确定卷积核的初始位置, 使初始位置能够根据图像中的目标和特征动态变化, 不再在原始模型中固定不变。AKConv 还可以由目标的特点来改变卷积核的采样位置, 以适应图像里目标形状和大小的不断变化, 适应性更强。

YOLOv10 改进模型 ACE-YOLOv10 的结构如图 3 所示。

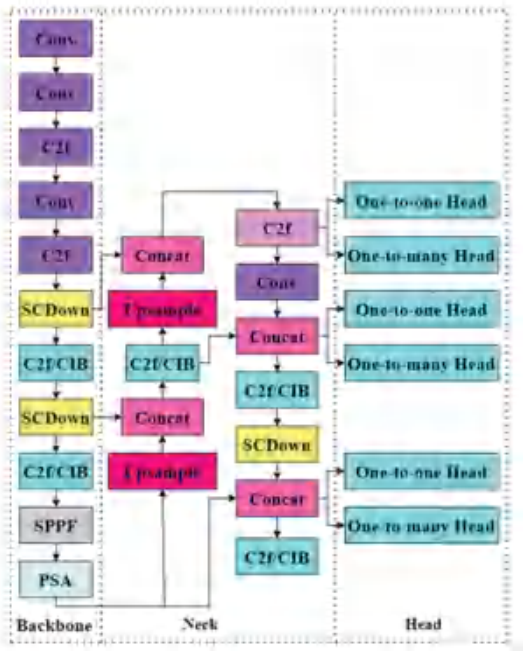


图 2 YOLOv10 网络结构

Fig. 2 The network architecture of YOLOv10

1.2.1 AKConv 模块

为了解决上述问题, 自适应核卷积^[17] (Adaptive Convolution, AKConv) 模块被引入 YOLOv10 当中, 该模块的主要原理如下。

(1) 自适应采样的形状: 当模型在分析不同类别的目标和图像时, AKConv 的卷积核可以自动的进行调整, 从而使得算法模型可以为各种形状和大

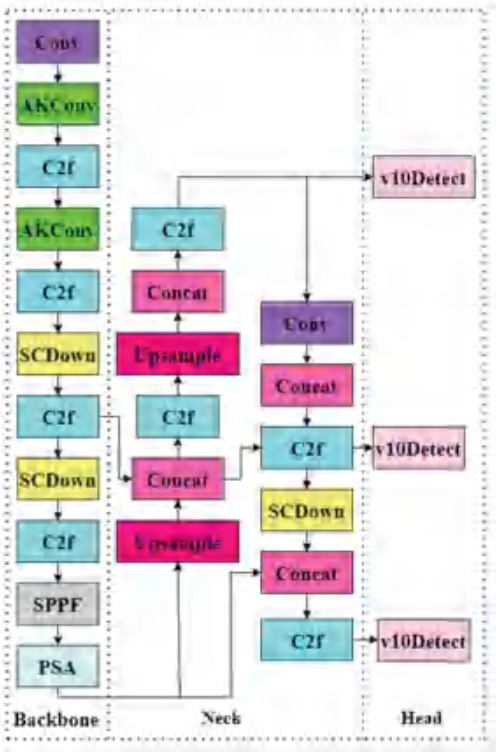


图 3 ACE-YOLOv10 网络结构

Fig. 3 The network architecture of ACE-YOLOv10

1.2.2 CA 注意力机制

坐标注意力^[18] (Coordinate Attention, CA) 机制可以提取远距离的通道和空间信息, 有助于模型对小型目标的辨别能力, 更快地识别和定位目标物体。具体优点如下。

(1) 提高定位精度: CA 注意力机制引入了图像的空间坐标信息, 使得模型可以更准确地区分和识

别目标,提升了模型处理复杂场景下的性能。

(2)提升特征融合:CA 注意力机制可以使得模型同时融合不同层次的特征信息。融合不同尺度带有特征的图像后,模型可以拥有更好的性能。

(3)减少背景干扰:通常情况下,图像中噪声信息大量存在,对目标检测工作产生干扰。CA 注意力机制能够聚焦与目标相关的特征,从而有利于减轻背景中噪声对目标检测结果的影响。

(4)增强模型鲁棒性:CA 注意力机制通过关注目标区域,使得模型在目标检测时,遇到光照变化和遮挡等情况,鲁棒性有很大的提高。

(5)优化特征图表示:CA 注意力机制对特征图^[19]加权之后,加强了与目标相关的特征,减弱了与目标无关的特征。这种优化可以提升特征图的质量,使得检测准确性得到进一步提高。

(6)加速目标检测:尽管引入 CA 注意力机制会增加一定的计算复杂度,但总体检测速度及精确度均得到很大的提升。

1.2.3 $EIoU$ 损失函数

改进的损失函数使目标检测的性能得到一定提升。对比完全交并比(Complete Intersection over Union, $CIoU$)损失函数,本文采用 $EIoU$ ^[20] 损失函数,此函数能够充分考虑真实框与预测框之间的中心点距离、重叠面积和纵横比等情况。与普通的 IoU 相比, $EIoU$ 可以更灵活地调整预测框去匹配真实框。由于纵横比是一个相对的比值,很难保证样本难易度的平衡,会存在一些不确定性。为了解决上述的问题,在原始 $CIoU$ 函数的基础上添加了新的惩罚项,得到 $EIoU$ 损失函数,其公式如下:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \quad (1)$$

其中, h^c 和 w^c 分别表示覆盖预测框以及真实框的最小外接矩形的高度与宽度, $\rho(b, b^{gt})$ 表示真实框与预测框的中心点之间的欧氏距离。

$EIoU$ 的新惩罚项在原有的 $CIoU$ 基础上进行了改进,计算预测框和真实框的高度和宽度损失,损失函数能够减小高与宽之间的差异性,从而提升新模型的收敛速度。重叠损失、宽高损失和中心距离损失三部分共同组成了损失函数。而中心距离损失和重叠损失与 $CIoU$ 中的损失相同。

2 实验结果与分析

2.1 实验环境与数据集

本文实验在一台装有 Windows10 操作系统的计算机上进行,编程软件使用 PyCharm 专业版,PyTorch 作为深度学习框架。GPU 为 NVIDIA CeForce RTX 3080ti。训练轮数 200 轮次,图像输入采用尺寸为 440×400 左右,动量大小为 0.937,初始学习率为 0.01, Batch Size 为 4,权重衰减系数为 0.000 5,采用从头开始训练的方式,对原始模型和改进模型进行对比。

为了验证本文算法在无人机拍摄数据集集中的检测性能,实验采用飞桨 AI Studio 官网下载的 2022 年船海数据智能应用创新大赛中水面水下典型目标识别的数据集,并从其中选取了 2 000 张船舶图片作为本次实验的数据集,以 6:2:2 的比例分隔训练集、验证集和测试集,用于测试改进模型的各项指标。

2.2 评价指标

实验评价指标对模型的效果和后续的改进都有着很大的影响。文中的实验结果评价标准如下:平均精度(Average Precision, AP)、平均精度均值(mean Average Precision, mAP)、精确率($Precision, P$)、召回率($Recall, R$)、模型参数量(Model Parameters, MP)、模型复杂度(Giga Floating point Operations Per second, $GFLOPs$)。各种评价标准的具体计算方法如下:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$AP = \int_0^1 P(R) dR \quad (4)$$

$$mAP = \frac{\sum_{m=1}^M AP}{M} \quad (5)$$

其中, TP 表示模型能够正确识别的正样本数; FP 表示模型把负样本错误识别成正样本数; FN 表示模型把正样本错误识别成负样本数。 $P-R$ 曲线下方的面积记为 AP , 代表了分类器性能的好坏。 mAP 则为 AP 的均值,可以用于整体数据的评估。

2.3 结果分析

为了更好地展现出改进模型的有效性,图 4 展示了 3 组改进前后的检测对比效果图。

由图 4 可以清晰看出,经过 200 轮的训练后,改进模型 ACE-YOLO 相较于原始模型,在所用数据集上展现了更好的检测识别能力,漏检、误检的情况也

更少。新模型可以更加有效地帮助救援人员在海洋中搜索到遇险船舶,也可以节省搜索时间,提前做好救援准备,增加遇险人员的生还几率,保护人民群众的生命财产安全。

消融实验通过把消融前后的模型性能进行对比,可以判断出哪些模块对于模型的性能起着关键的作用。消融实验还可以展示出新模型对于不同参数变化后的敏感程度,如果移除了某个参数后,整个模型性能则有大幅的下降,说明新模型对该参数比较敏感,该参数对模型的行为有着较大的影响,应该

慎重修改。此外,消融实验还可以用来评估新模型的鲁棒性,一个鲁棒的模型在某些部分受到噪声影响时,也能够有一定的适应能力。因此,通过逐渐移除或是更换模型中的某些组件,研究人员就能够观察到这些组件变化对于整体性能的影响。如:在自然语言的处理中,消融实验能够帮助研究人员了解各阶段的特征提取对于最终检测精度的贡献。总之,消融实验作为一种必不可少的工具,不仅能够提升研究人员的模型理解能力,也能让后续的改进模型更加高效,更加鲁棒。

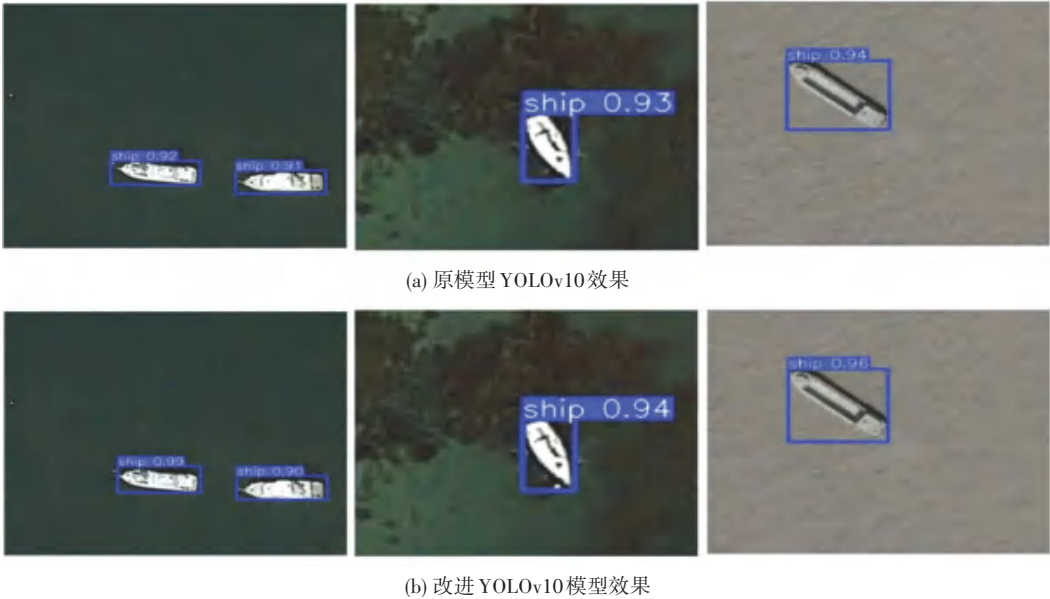


图 4 效果对比图

Fig. 4 Comparison diagram of the effect

消融实验结果见表 1。由表 1 可知,使用 AKConv 模块后模型的计算规模和参数的数量有着明显的下降。如在 mAP 、 P 、 R 三个指标上都降低了 2.2%、0.1%、0.5%,因为 AKConv 模块可以通过近似卷积操作来减少模型的计算规模和参数数量。而在模型中引入 CA 注意力机制与 $EIoU$ 损失函数后,新模型 ACE-YOLOv10 比原始的 YOLOv10 模型在目标检测性能方面有一定的提升,从而可以更好地应用到海上救援任务的船舶检测工作中。

表 1 消融实验结果

Table 1 The results of the ablation experiment							
ACConv	CA	$EIoU$	$P/\%$	$R/\%$	$mAP50/\%$	$MP/MGFLOPs/G$	
-	-	-	95.68	95.54	96.85	11.15	28.64
✓	-	-	95.57	95.03	96.53	10.13	28.62
✓	✓	-	97.10	97.27	98.14	10.32	28.66
✓	✓	✓	98.08	98.19	98.43	10.32	28.66

3 结束语

本文提出的 ACE-YOLO 模型,通过引入 AKConv 模块以及 CA 注意力机制,将原始 $CIoU$ 损失函数替换为 $EIoU$ 损失函数,优化了模型对小型目标的检测能力。与原模型相比,新模型表现更好,能够良好地适应复杂情况下的目标检测。在实验的结果上,比起最原始的模型,ACE-YOLO 模型也表现出更好的效果,在 P 、 R 两个指标上分别有 2.5%、2.8%的提升。

虽然改进的 ACE-YOLO 模型具有较高的目标检测性能,但在计算量以及低延时性方面还亟待完善,本文提出的新模型在验证的数据集方面还是存在一定的局限性。因此,后续的研究工作将继续关注模型的性能提升和计算量减少,同时扩充数据集以提高泛化性。

参考文献

[1] ZOU Zhengxia, CHEN Keyan, SHI Zhenwei, et al. Object detection in 20 years: A survey[J]. Proceedings of the IEEE, 2023, 111(3): 257-276.

[2] WU Xiongwei, SAHOO D, HOI C. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020, 396: 39-64.

[3] TAN Mingxing, PANG Ruoming, LE Q. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2020: 10781-10790.

[4] 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. Piscataway, NJ:IEEE, 2016: 779-788.

[5] 白堂博, 杨建伟, 许贵阳, 等. 基于 Faster R-CNN 的铁路扣件定位方法研究[J]. 铁道科学与工程学报, 2021, 18(2): 502-508.

[6] 李霄峰. 基于视觉的高铁周界入侵检测算法研究[D]. 北京: 北京交通大学, 2020.

[7] 张永强, 李胜男, 张子强, 等. R-YOLO 轨道人员目标检测模型[J]. 河北科技大学学报, 2023, 44(6): 580-588.

[8] 师后勤, 谢辉, 张梦钰, 等. 一种基于注意力机制的低光照下行人检测算法[J]. 物联网技术, 2023, 13(2): 27-29.

[9] 梁京涛, 赵聪, 马志刚. 多源遥感技术在地质灾害早期识别应用中的问题探讨: 以西南山区为例[J]. 中国地质调查, 2022, 9(4): 92-101.

[10] 孙日明, 郭虎, 邹丽, 等. 融合决策下的数字序列跌倒检测方法[J]. 系统仿真学报, 2023, 35(9): 2045-2053.

[11] 赵江平, 刘星星, 张想卓. 基于改进 YOLOv5s 的外脚手架隐患图像识别技术[J]. 中国安全科学学报, 2023, 33(12): 60-66.

[12] ZHANG Qinglong, YANG Yubin. SA-Net: Shuffle attention for deep convolutional neural networks [C]// 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2021: 2235-2239.

[13] 张铭泉, 邢福德, 刘冬. 基于改进 Faster R-CNN 的变电站设备外部缺陷检测[J]. 智能系统学报, 2023, 19(2): 290-298.

[14] HAO Z, WANG Q. Improved algorithm for small target detection in UAV aerial images based on YOLOv7 [J]. Software Guide, 2024, 23(1): 167-172.

[15] WANG Ao, CHEN Hui, LIU Lihao, et al. YOLOv10: Real-time end-to-end object detection [J]. arXiv preprint arXiv, 2405.14458, 2024.

[16] SAPKOTA R, QURESHI R, CALERO M, et al. YOLOv10 to its genesis: A decadal and comprehensive review of the you only look onceseries[J]. arXiv preprint arXiv, 2406.19407, 2024.

[17] ZHANG Xin, SONG Yingze, SONG Tingting, et al. AKConv: Convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters[J]. arXiv preprint arXiv, 2311.11587, 2023.

[18] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72.

[19] 郭璠, 张泳祥, 唐璵, 等. YOLOv3-A: 基于注意力机制的交通标志检测网络[J]. 通信学报, 2021, 42(1): 87-99.

[20] YANG Zuoming, WANG Xianlun, LI Jianguang. EIoU: An improved vehicle detection algorithm based on vehiclenet neural network[J]. Journal of Physics: Conference Series, 2021, 1924(1): 012001.