

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

DES-YOLO: 一种更精确的目标检测方法

郑华伟, 王飞, 高建邦

引用本文:

郑华伟, 王飞, 高建邦. DES-YOLO: 一种更精确的目标检测方法[J]. *光电工程*, 2024, 51(11): 240212.

Zheng H W, Wang F, Gao J B. DES-YOLO: a more accurate object detection method[J]. *Opto-Electron Eng*, 2024, 51(11): 240212.

<https://doi.org/10.12086/oe.2024.240212>

收稿日期: 2024-09-09; 修改日期: 2024-10-11; 录用日期: 2024-10-11

相关论文

基于ZYNQ的轻量化YOLOv5声呐图像目标检测算法及实现

赵冬冬, 谢墩翰, 陈朋, 梁荣华, 沈伊, 郭新新

光电工程 2024, 51(1): 230284 doi: 10.12086/oe.2024.230284

融合视觉中心机制和并行补丁感知的遥感图像检测算法

梁礼明, 陈康泉, 王成斌, 冯耀, 龙鹏威

光电工程 2024, 51(7): 240099 doi: 10.12086/oe.2024.240099

改进YOLOv7的无人机视角下复杂环境目标检测算法

张润梅, 肖钰霏, 贾振楠, 陈中, 陈梓华, 袁彬, 曹炜威, 宋妮妮

光电工程 2024, 51(5): 240051 doi: 10.12086/oe.2024.240051

PAW-YOLOv7: 河道微小漂浮物检测算法

栾庆磊, 常昕昱, 吴叶, 邓从龙, 史艳琼, 陈梓华

光电工程 2024, 51(4): 240025 doi: 10.12086/oe.2024.240025

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



OE_Journal



Website



DOI: 10.12086/oe.2024.240212

CSTR: 32245.14.oe.2024.240212

DES-YOLO: 一种更精确的目标检测方法



郑华伟, 王飞*, 高建邦

西安石油大学电子工程学院, 陕西 西安 710065

摘要: 针对图像中背景复杂、目标小、分布密集等问题, 提出了一种改进的 DES-YOLO 方法。通过引入可变形注意力模块 (DAM), 网络可动态关注关键区域, 提高物体识别和定位精度; 采用高效交并比 (EIoU) 损失函数, 减少低质量样本影响, 增强泛化能力和检测精度; 在网络头部加入一层 160 pixel×160 pixel 的浅层特征图, 加强小目标特征提取; 并使用分步训练策略提升模型性能。实验结果表明, 该模型在遥感数据集上的 mAP@50 提升了 1.4%, 在纺织数据集上提升了 1.7%, 验证了 DES-YOLO 的广泛适用性与有效性。

关键词: 目标检测; 可形变注意力; EIoU; 浅层特征; 分步训练; DES-YOLO

中图分类号: TP391

文献标志码: A

郑华伟, 王飞, 高建邦. DES-YOLO: 一种更精确的目标检测方法 [J]. 光电工程, 2024, 51(11): 240212

Zheng H W, Wang F, Gao J B. DES-YOLO: a more accurate object detection method[J]. *Opto-Electron Eng*, 2024, 51(11): 240212

DES-YOLO: a more accurate object detection method

Zheng Huawei, Wang Fei*, Gao Jianbang

School of Electronic Engineering, Xi'an Shiyou University, Xi'an, Shaanxi 710065, China

Abstract: To address the challenges of complex backgrounds, small targets, and dense distributions in images, an improved method called DES-YOLO is proposed. By introducing the deformable attention module (DAM), the network can dynamically focus on key regions, improving object recognition and localization accuracy. The efficient intersection over union (EIoU) loss function is employed to reduce the impact of low-quality samples, enhancing the model's generalization ability and detection accuracy. A shallow feature map layer of 160 pixel×160 pixel is added to the network head to strengthen small target feature extraction. A stepwise training strategy is also adopted to further improve model performance. Experimental results show that the mAP@50 of the model increased by 1.4% on the remote sensing dataset and by 1.7% on the textile dataset, demonstrating the broad applicability and effectiveness of DES-YOLO.

Keywords: object detection; deformable attention; EIoU; shallow features; stepwise training strategy; DES-YOLO

收稿日期: 2024-09-09; 修回日期: 2024-10-11; 录用日期: 2024-10-11

基金项目: 西安石油大学研究生创新与实践能力的培养计划立项项目 (YCS23214252)

*通信作者: 王飞, 200102@xsyu.edu.cn。

版权所有©2024 中国科学院光电技术研究所

1 引言

目标检测作为计算机视觉中的一个重要研究方向,旨在识别和定位图像中的特定目标。它在城市规划、农业监测、环境保护和军事侦察等领域有广泛应用^[1]。但由于目标外观和姿态的多样性,以及成像时的光照、遮挡等因素,目标检测面临着诸多挑战,尤其是在小目标检测时更为明显。背景复杂、数据不平衡等问题通常会导致模型对小目标的检测精度降低。

近年来,小目标检测技术取得了显著进步,主要得益于深度学习的兴起和计算硬件性能的提升^[2]。传统方法通常受限于特征表示和检测精度,而深度学习能够自动学习高级特征,并在大规模数据集上进行训练,从而显著提升性能。现阶段目标检测算法大致分为两类:单阶段算法(如 YOLO、SSD)和两阶段算法(如 R-CNN 系列)。单阶段算法速度快,适用于实时应用,但对小目标检测效果不佳;两阶段算法精度高,但计算复杂度较大。这两种方法在近年来的发展中各有优势和局限,适用于不同的应用场景。

YOLOv5 作为 YOLO 系列发展较为完善的单阶段目标检测算法,具备快速处理能力。然而,其在小目标检测上的表现仍有不足。例如,由于锚框尺寸与小目标差异较大,导致预测框设置不合理。研究者们为此提出了动态生成锚框、K-means 聚类等改进方法,以提高小目标检测的准确性^[3]。

随着智能监控、自动驾驶和航空航天等领域的快速发展,小目标检测的需求也不断增长。尤其在遥感目标检测中,深度学习的应用显著提升了检测精度,推动了小目标检测技术的进步^[4]。相关研究提出了多种方法,如融合多级特征、基于迁移学习的检测模型等,成功提升了小目标检测的精度和速度^[5-6]。纺织工业中的瑕疵检测也亟需技术创新,尽管自动化生产水平不断提升,但微小瑕疵检测依然是一个难题,因为它们可能在图像中仅占据极少像素^[7]。传统的计算机视觉方法,如基于小波滤波器或形态分析的瑕疵检测^[8-9],往往由于特征表示不足和应用范围有限,导致检测效果不佳。近年来,深度学习技术的引入为纺织瑕疵检测提供了更高效的解决方案^[10]。结合特征金字塔网络、可变形卷积网络和距离 IoU 损失函数等技术,显著提升了检测的精度和速度^[11-12]。

本文提出了一种改进的目标检测模型 DES-YOLO(模型名称中的“D”代表可形变注意力模块(deformable attention module),“E”取自 EIoU 损失函

数,而“S”则源自小目标检测层(small object detection layer);“YOLO”则表示该模型基于目标检测领域广泛使用的 YOLO 架构),以提升小目标检测性能。具体改进包括:

1) 在主干网络中引入可变形注意力模块(DAM),动态调整注意力焦点,自适应地集中在图像的关键区域,从而提高目标检测的精度。

2) 采用 EIoU 损失函数,减少低质量样本的影响,提高模型的泛化能力和检测精度。

3) 在网络头部增加一个带有 160 pixel×160 pixel 浅层特征图的检测层,增强网络对浅层特征的提取能力。

4) 提出一种分步训练策略,优先训练小目标检测层,从而提升整体模型的训练效率。

此外,单一数据集的训练和测试往往存在容错率低的问题,限制了模型的泛化能力。因此,本文选用了包括遥感和纺织瑕疵数据集在内的多种数据集,验证改进算法在不同场景下的有效性和适用性。

本文结构安排如下:第二节综述目标检测方法的相关研究;第三节详细介绍模型改进的具体方法;第四节对实验结果进行分析与验证;第五节通过消融实验和对比实验证明改进模型的有效性;最后,第六节总结研究成果并展望了未来的研究方向。

2 相关工作

2.1 注意力机制

相较于神经网络学习全部的输入信息,注意力机制的引入使卷积神经网络能够自动学习并选择性地关注输入中的重要信息,极大地提高了对小目标的检测性能。Li 等人提出在 YOLOv3-ST 模型的残差块结构中分别嵌入 SENet 和 CBAM 两种不同的注意机制,大幅提高了对海草等小目标的检测精度^[13]。Cao 等人将坐标注意力添加到主干网络中,以增强网络对安全帽目标细节的学习^[14]。Zhang 等人引入一种将空间注意力和通道注意力串联的双注意力机制模块(CBAM)以获取更好的车道线特征表示,提高车道线分割精度^[15]。然而,这些方法在处理不规则形状的小目标时仍存在局限性,尤其是在不同背景下形状多样的目标检测方面。

2.2 小目标检测层

YOLOv5 网络模型包含三个目标检测层,主要侧

重于中大型目标的检测。然而, 对于小目标的检测, 尽管采用了多尺度检测机制, 但通常仅有一个尺度能够有效识别出小目标。Luo 等人在 YOLOv5 中增加了小目标检测层, 提高了对小火焰的检测效果^[16]。Li 等人在 YOLOv5s 算法的三个检测层上添加了一个具有四倍下采样特征的附加检测层, 该检测层集成了更多浅层特征, 减少了卷积和汇集操作后小目标的特征丢失问题^[17]。

然而, 添加小目标检测层不可避免地增加了网络参数, 进而降低模型的训练速度。如何高效地平衡训练速度与检测精度仍然是一个亟待解决的重要问题。采用更高效的特征融合策略对于提升小目标检测精度具有重要意义。Chen 等人设计了一种特征融合模块 SCAM, 该模块通过局部特征的自监督学习提高了细节信息的利用率, 并通过多尺度特征的有效融合, 显著提升了中大型目标的分类精度^[18]。Ge 等人使用 BiFPN 结构进行多尺度特征融合, 赋予浅层特征更高的权重, 更好地结合低层位置信息和高层语义信息, 加强浅层特征的利用率, 进一步提高网络的检测性能^[19]。Chen 等人将 YOLOv4-tiny 主干的高卷积特征层与颈部网络的特征层进行融合, 增加了 52 pixel×52 pixel 预测尺度的输出, 提高了小目标检测的精度^[20]。传统的特征融合方式存在一定的局限性, 由于未能充分融合浅层特征, 这对小目标的检测产生了显著的不利影响。

2.3 损失函数

损失函数作为模型训练和评估的核心要素, 对小目标检测精度具有至关重要的影响。Zhang 等人采用 EIoU_Loss 取代 CIoU_Loss, 解决了用等宽高比惩罚失败的问题^[21]。Pu 等人在回归损失的计算中引入了 NWD 损失, 以补偿 CIoU 损失对小目标位置差异的敏感性^[22]。Lin 等人在 RetinaNet 中提出 Focal Loss 损失函数, 通过聚焦硬样本训练平衡样本比例, 大幅提升单阶段检测器的精度^[23]。Wang 等人通过基于 FIoU 的损失函数, 显著提高了小目标检测的准确性^[24]。Luan 等人则采用 WIoU 损失函数替代 CIoU, 降低了网络模型的计算量, 提升了检测速度, 并增强了对低质量锚框的聚焦能力, 从而加快了模型的收敛速度^[25]。在公开数据集中普遍存在样本数据不均衡、质量低的问题, 损失函数的改进对解决这一问题具有积极意义。

随着深度学习的不断发展, 研究者在网络架构、

特征提取、特征融合、注意力机制以及训练策略等方面进行了大量改进与优化。这些努力共同促进了模型训练速度、收敛效率以及小目标检测精度的提升。

3 改进方法

3.1 基准算法

YOLOv5 作为 YOLO 系列中的基础模型, 凭借其卓越的速度和准确性, 超越了大多数现有的目标检测器。其网络架构由四个主要组件组成: 输入层、骨干网络、颈部模块和检测头。在输入层, 图像经过数据增强和预处理后被送入骨干网络。骨干网络的主要任务是从预处理后的图像中提取关键特征。接着, 颈部模块将提取的特征进行融合和处理, 生成大、中、小三种不同尺度的特征图。最终, 这些特征图被传递至检测头, 用于输出检测结果。

在 YOLOv5 中, SPPF 模块和 C3 模块扮演着重要角色。SPPF 模块通过引入多尺度的空间金字塔池化操作, 提升了网络对多尺度目标的感知能力, 尤其是在保持计算效率的前提下扩展了感受野。C3 模块则通过部分跨阶段连接, 减少了冗余的梯度信息流, 同时提高了特征提取的效率和网络的学习能力。原始网络架构如图 1 所示。

3.2 相关改进

3.2.1 可形变注意力模块

在不同应用背景下的图像中, 小目标检测是一项对敏感度要求极高的任务。DA (deformable attention) 机制是一种能够适应目标形状变化的先进注意力机制。通过学习偏移量和注意力权重, DA 机制能够有效地聚焦不规则形状的目标, 从而在一定程度上缓解小目标形状多样性带来的识别难题^[26]。DA 工作流程如图 2 所示。

引入注意力机制的主要优势包括:

- 1) 强化关键特征: DAM 能够在局部区域内聚焦注意力, 从而强化可能代表微小目标的关键特征。
- 2) 适应性形状注意力: DAM 能够更好地适应多种不规则形状, 提高检测精度。
- 3) 上下文信息: 通过结合周围区域的上下文信息, DAM 帮助模型区分待检测的小目标与背景纹理。

3.2.2 小目标检测层

YOLOv5 在小目标检测效果不佳的一个原因在于小目标样本尺寸较小, 而 YOLOv5 的下采样倍数较

大, 深层次特征图难以有效捕捉小目标的特征信息。为了解决这一问题, 我们在拼接浅层和深层特征图后, 增加了一个专门用于小目标检测的层。原始 YOLOv5 模型包含三个检测层, 分别对应三组预定义的锚点值,

负责从网络的颈部提取多个聚合特征图, 进而推断目标边界框并进行分类。为了提升小目标检测能力, 我们进一步对模型进行上采样操作, 并将其与主干网络中的特征融合, 添加回归和分类卷积块专门处理小目

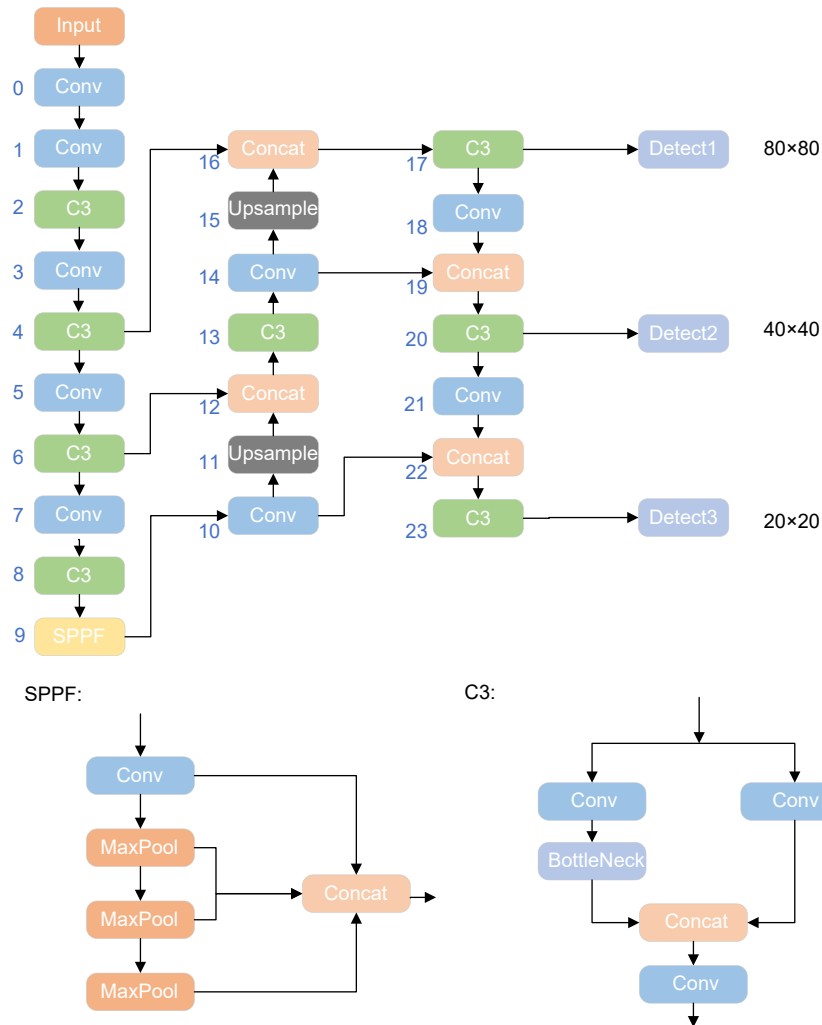


图 1 YOLOv5 网络架构
Fig. 1 YOLOv5 network structure

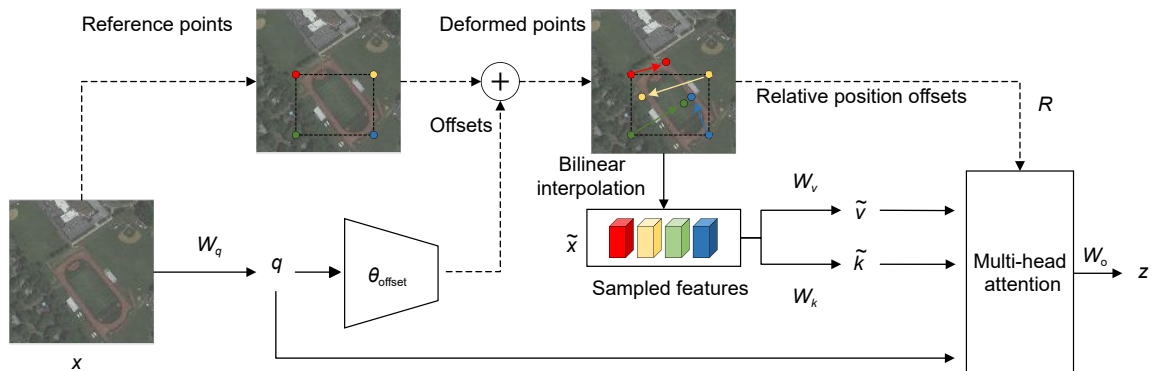


图 2 可形变注意力模块
Fig. 2 Deformable attention module

标检测。

在改进后的 YOLOv5 模型中, 输入图像尺寸为 640 pixel×640 pixel。原始模型中的三个检测层对应的特征图尺寸分别为 80 pixel×80 pixel、40 pixel×40 pixel 和 20 pixel×20 pixel, 用于检测尺寸为 8 pixel×8 pixel、16 pixel×16 pixel 和 32 pixel×32 pixel 的目标。本文提出的改进网络结构如图 3 所示, 增加了一个新的小目标检测层, 对应的特征图尺寸为 160 pixel×160 pixel, 适用于检测尺寸为 4 pixel×4 pixel 或更小的目标。通过提高对细粒度特征的分辨率, 改进后的模型显著增强了小目标的检测能力。尽管计算量有所增加, 导致检测速度略微下降, 但小目标的检测精度得到显著提升。

3.2.3 损失函数

EIoU 是在 CIoU 的惩罚项基础上将预测框和真实框的纵横比的影响因子拆开, 分别计算预测框和真实框的长和宽, 来解决 CIoU 存在的问题。EIoU 包括三个部分: IoU 损失、距离损失和高宽损失 (重叠面积、中心点举例和高宽比)。高宽损失直接最小化了预测目标边界框和真实边界框的高度和宽度的差异, 使其

有更快的收敛速度和更好的定位结果。计算公式如式 (1) 所示。

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

其中:

L_{EIoU} 表示 EIoU 总损失;

L_{IoU} 表示 IoU 损失, $L_{IoU} = 1 - IoU$, IoU 用于衡量预测框与真实框的重叠区域与二者并集的比例;

L_{dis} 表示距离损失, $L_{dis} = \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2}$, \mathbf{b} 表示预测框中心, \mathbf{b}^{gt} 表示真实框中心, ρ 表示欧氏距离, c 表示包含预测框和真实框的最小外接框对角线的长度;

L_{asp} 表示宽高损失, $L_{asp} = \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2}$, w 表示预测框的宽度, w^{gt} 表示真实框的宽度, h 表示预测框的高度, h^{gt} 表示真实框的高度, C_w 表示包含预测框和真实框的最小外接框的宽度, C_h 表示包含预测框和真实框的最小外接框的高度。

3.2.4 训练策略

提高小目标检测精度的常用方法是输出浅层特征并进行特征融合, 即通过增加一个小目标检测层来实

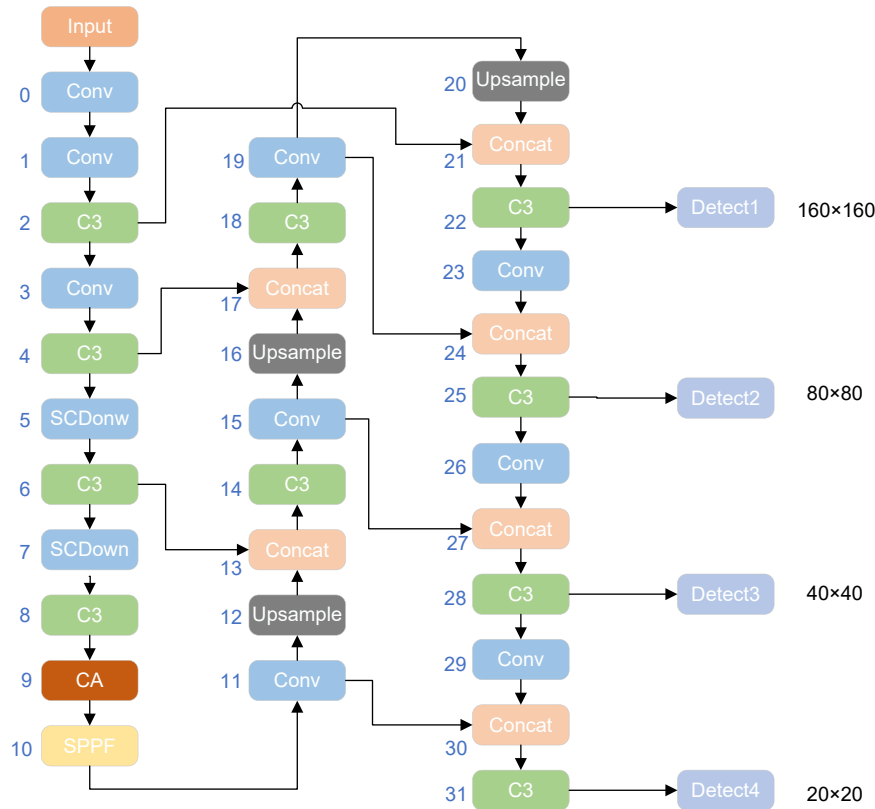


图 3 改进的网络架构
Fig. 3 Improved network structure

现。然而，这会增加网络的深度和参数量，进而降低训练速度和模型的收敛效率。为了解决这一问题，本文提出了一种分步训练策略，以在训练速度与检测精度之间取得更好的平衡，并优化小目标检测层的训练效果。

首先，我们冻结了改进 YOLOv5 的后两个检测层，使模型专注于前两个检测层的训练，尤其是小目标检测层。在这一阶段训练结束后，保存最佳权重。接下来，将网络的所有层解冻，进行联合训练，进一步提升大目标的检测能力，并对小目标检测参数进行微调。整个训练周期设为 300 个周期，第一阶段为 200 个周期，第二阶段为 100 个周期。

实验结果验证了该分步训练策略的有效性，不仅提高了小目标的检测精度，还加快了模型的收敛速度，使其能够更高效、准确地检测小目标。

4 实验结果及分析

为了验证 DES-YOLO 模型的性能，我们在两种

数据集上通过添加不同的注意力机制、使用不同的损失函数以及增加小目标检测层等对比验证模型的性能。

4.1 实验平台及数据集

4.1.1 实验平台

实验使用的开发平台为 Windows 11，具体平台配置如表 1 所示。

4.1.2 数据集

1) 遥感目标检测数据集

本文采用 NWPU VHR-10 数据集，该数据集由西北工业大学发布，是一个用于空间物体检测的 10 级地理遥感数据集，其拥有 650 张包含目标的图像和 150 张背景图像，共计 800 张图像，将其按比例分为 640 张训练集和 160 张测试集。目标种类包括飞机、舰船、油罐、棒球场、网球场、篮球场、田径场、港口、桥梁和汽车共计 10 个类别^[27-29]。NWPU VHR-10 数据集部分图像如图 4 所示。

2) 纺织物瑕疵检测数据集

表 1 实验环境与配置

Table 1 Experimental environment and configuration

Type	Configuration
GPU	NVIDIA GeForce RTX4090
CPU	13th Gen Intel(R) Core(TM) i7-13620H
CUDA	11.7
Deep learning framework	Pytorch
Python	3.12



图 4 NWPU VHR-10 数据集部分图像

Fig. 4 Images of part of the NWPU VHR-10 dataset

本文采用阿里天池布匹瑕疵检测数据集, 天池平台通过现场采集布匹图像, 制作并发布大规模的高质量布匹疵点数据集, 同时提供了精细的标注以满足算法要求。布匹数据包含无疵点图片、有疵点图片和瑕疵的标注数据, 并且每张图片包含一个或多个瑕疵。本文选取天池的其中一个测试集, 共 2389 张布匹瑕疵图片以及 2389 个标注文件, 并将其按比例分为 1896 张训练集和 493 张测试集。包括破洞、水渍、油渍、粗经、结头和磨痕等 30 个瑕疵类别^[30]。布匹瑕疵数据集部分图像如图 5 所示。

4.2 评价指标

本文采用精确率 (Precision)、召回率 (Recall)、平均精度 (mAP)、参数数量 (Parameters) 以及计算量 (GFLOPs) 等常见的性能指标来评估改进模型的效果。为了获得最终的实验结果, 实验首先利用预先划分的训练集和验证集对模型进行训练, 随后在测试集上对

训练完成的模型进行评估测试。

4.3 实验结果分析

为了验证改进模型的每项增强部分的有效性, 我们在两个不同的数据集上分析了模型的性能。

1) 不同注意机制的实验结果比较

为了便于对不同的注意机制进行综合评估, 并证实 DAM 在小目标检测领域的优势, 我们在 YOLOv5 模型中加入了多种注意机制, 包括 CA 和 CBAM。随后, 将这些改进的模型在两种数据集上进行训练。相应模型在训练过程中的结果对比见表 2 和表 3。

如表 2 所示, 在遥感数据集上, 原始模型引入 CBAM 后, 精确率略微提升, 但召回率下降, 导致 mAP@0.5:0.95 显著下降。使用 CA 模块时, mAP@0.5 略有提升, 但 mAP@0.5:0.95 仅为 55.1%, 提升有限。相比之下, 引入 DAM 后的模型表现更加平衡, 特别是在 mAP@0.5:0.95 上达到 56.1%, 在精确率和召回



图 5 布匹瑕疵数据集部分图像

Fig. 5 Images of part of the fabric defect dataset

表 2 遥感目标检测数据集

Table 2 Remote sensing target detection data set

	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s	0.937	0.899	0.929	0.562
YOLOv5s+CBAM	0.939	0.884	0.923	0.511
YOLOv5s+CA	0.922	0.886	0.931	0.551
YOLOv5s+DA	0.945	0.898	0.942	0.561

表 3 纺织物瑕疵检测数据集

Table 3 Textile defect detection data set

	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s	0.350	0.322	0.276	0.118
YOLOv5s+CBAM	0.382	0.290	0.282	0.141
YOLOv5s+CA	0.237	0.296	0.219	0.086
YOLOv5s+DA	0.350	0.342	0.285	0.121

率上均有显著提升, 分别达到 94.5% 和 89.8%, 相较于原始模型, mAP@0.5 提高了 1.3%。

如表 3 所示, 在纺织物数据集上, 原始模型引入 CBAM 后, 精确率显著提升了 3.2%, 但召回率下降, mAP@0.5 略有提升, mAP@0.5:0.95 达到了 14.1%。使用 CA 模块时, 精确率、召回率和 mAP@0.5 均有下降。相比之下, 引入 DAM 后的模型表现出色, 召回率和 mAP@0.5 均有显著提升, 分别达到 34.2% 和 28.5%, 相较于原始模型分别提高了 2% 和 0.9%。

综上所述, 虽然不同改进方法在某些指标上有所提升, 但引入 DAM 的模型在整体检测性能上表现最为平衡, 在高 IoU 阈值下的检测性能显著改善, 同时保持了高精度和高召回率。

2) 不同损失函数的实验结果比较

为了验证不同的损失函数对模型性能的影响, 该部分在模型引入 DA 注意力模块的基础上进行实验。

如表 4 所示, 在遥感数据集上, 原始模型的损失函数替换为 EIoU 后, 精确率略有下降, 但召回率、mAP@0.5 和 mAP@0.5:0.95 均有显著提升, 分别提高了 1.6%、2% 和 1.2%, 并且效果优于 SIoU 和 WIoU。

如表 5 所示, 在纺织物数据集上, 原始模型的损失函数替换为 EIoU 后, 精确率显著提升, 达到了

39.2%, 召回率下降, 导致 mAP@0.5 略有下降, 但依然与最优结果几乎相持。与 SIoU 和 WIoU 相比, EIoU 表现更加均衡。

综上所述, 在两种数据集上, 通过对比 CIoU、EIoU、SIoU、WIoU 四种损失函数, 证实 EIoU 在本文采用的数据集上优于其他三种损失函数。与原模型相比, EIoU 在准确率、召回率和 mAP@0.5 三项评价指标中均表现较优。

3) 小目标检测层

为了验证小目标检测层对模型性能的影响, 该部分在模型引入 DAM、采用损失函数 EIoU 的基础上进行实验, 以探讨改进后的模型在不同的应用场景中对小目标检测的表现。表中 STD 代表小目标检测层。

如表 6 所示, 在遥感数据集上, 原始模型添加小目标检测层后, 参数量增加, mAP@0.5 下降了 0.1%, 但对以 Ship 为代表的小目标的检测精度提高了 0.7%。

如表 7 所示, 在纺织物数据集上, 原始模型添加小目标检测层后, 参数量增加, 精确率显著提高, 达到了 45.4%, 召回率略有下降, mAP@0.5 提高了 1.2%, 以 Knot head 为代表的小目标的检测精度提高了 0.8%。

表 4 遥感目标探测损失函数的效果比较

Table 4 Effect comparison of remote sensing target detection loss function

	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s+DA+CIoU	0.945	0.898	0.942	0.561
YOLOv5s+DA+EIoU	0.936	0.914	0.944	0.573
YOLOv5s+DA+SIoU	0.933	0.922	0.935	0.571
YOLOv5s+DA+WIoU	0.848	0.842	0.884	0.506

表 5 纺织品缺陷检测损失函数的效果比较

Table 5 Effect comparison of textile defect detection loss function

	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s+DA+CIoU	0.350	0.342	0.285	0.121
YOLOv5s+DA+EIoU	0.392	0.315	0.281	0.130
YOLOv5s+DA+SIoU	0.382	0.272	0.283	0.144
YOLOv5s+DA+WIoU	0.357	0.292	0.258	0.103

表 6 遥感目标检测

Table 6 Remote sensing target detection

	Params/M	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Ship
YOLOv5s+DA	8.1	0.936	0.914	0.944	0.573	0.973
YOLOv5s+DA+STD	8.7	0.963	0.903	0.943	0.613	0.98

综上所述, 虽然小目标检测层使网络参数量剧增, 但在不同应用场景中的小目标检测精度均有提高, 模型表现出良好的性能。

4.4 讨论

4.4.1 消融实验

在本文中, DES-YOLO 模型的改进主要包括: 将 DAM 集成到主干网络中, 用 EIoU 损失函数替换 CIoU, 添加小目标检测层等。为了验证这些方法对模型性能的影响, 本小节进行了消融实验。

1) 遥感目标检测

如表 8 所示, 添加各种改进点后, 模型的性能较原模型有所提高。DES-YOLO 网络的 mAP@0.5 比基线网络高约 1.4%。本文中模型参数的数量有所增加, 但其精度与其他部分相比均处于优势, 尤其是网络

的 mAP@0.5:0.95 提高了 5.1%, 预测准确率提高了 2.6%。

2) 纺织物瑕疵检测

如表 9 所示, 添加各种改进点后, YOLOv5s 模型的性能较原模型有所提高。DES-YOLO 网络的 mAP@0.5 比基线网络高约 1.7%。本文中模型参数的数量有所增加, 但其精度与其他部分相比均处于优势, 尤其是网络的 mAP@0.5:0.95 提高了 2.2%, 预测准确率提高了 10.4%。

4.4.2 对比实验

为了全面确定 DES-YOLO 模型的先进性和鲁棒性, 本研究在 NWPU VHR-10 数据集上与几种先进的目标检测模型进行了比较分析。对比实验结果如表 10 所示。

表 7 纺织物瑕疵检测

Table 7 Textile defect detection

	Params/M	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Knot head
YOLOv5s+DA	8.1	0.392	0.315	0.281	0.13	0.303
YOLOv5s+DA+STD	8.7	0.454	0.258	0.293	0.14	0.311

表 8 消融实验结果

Table 8 Results of ablation experiments

	Params/M	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s	7.0	0.937	0.899	0.929	0.562
YOLOv5s+DA	8.1	0.945	0.898	0.942	0.561
YOLOv5s+DA+EIoU	8.1	0.936	0.914	0.944	0.573
YOLOv5s+DA+EIoU+STD	8.7	0.963	0.903	0.943	0.613

表 9 对比实验结果

Table 9 Results of ablation experiments

	Params/M	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s	7.1	0.350	0.322	0.276	0.118
YOLOv5s+DA	8.1	0.350	0.342	0.285	0.121
YOLOv5s+DA+EIoU	8.1	0.392	0.315	0.281	0.130
YOLOv5s+DA+EIoU+STD	8.7	0.454	0.258	0.293	0.140

表 10 对比实验结果

Table 10 Results of comparison experiments

	Params/M	Precision	Recall	mAP@0.5	mAP@0.5:0.95	GFLOPs/G
Faster-RCNN	41.1	0.861	0.957	0.901	0.559	33.2
YOLOv3-tiny	8.6	0.952	0.860	0.929	0.544	12.9
YOLOv3	61.5	0.956	0.918	0.952	0.602	155.4
YOLOv5s	7.0	0.937	0.899	0.929	0.562	15.8
YOLOv5m	20.9	0.864	0.832	0.888	0.523	48.0
YOLOv7-tiny	6.0	0.786	0.631	0.768	0.386	13.3
YOLOv8s	11.1	0.906	0.867	0.932	0.601	28.5
DES-YOLO	8.7	0.963	0.903	0.943	0.613	27.9

如表 10 所示, 我们将 DES-YOLO 模型与其他先进模型在 Precision、Recall、mAP、参数量和 GFLOPs 等方面进行了全面比较。

在参数量和计算复杂度方面, DES-YOLO 的参数量为 8.7 M, 显著低于 Faster-RCNN (41.1 M)、YOLOv3 (61.5 M) 和 YOLOv5m (20.9 M), 仅略高于 YOLOv3-tiny (8.6 M) 和 YOLOv5s (7.0 M)。其计算量为 27.9 G, 虽然高于轻量化模型 YOLOv3-tiny (12.9 G) 和 YOLOv5s (15.8 G), 但低于 YOLOv8s (28.5 G) 等先进模型。

在精度方面, DES-YOLO 的 mAP@0.5 达到了 94.3%, 仅次于 YOLOv3 (95.2%), 而其 mAP@0.5:0.95 达到了 61.3%, 在所有对比模型中最高, 表明该模型在不同 IoU 阈值下表现出强大的稳定性和一致性。此外, DES-YOLO 在 Precision 上达到了 96.3%, 为所有模型中最佳, 展示了其卓越的目标识别性能。其 Recall 为 90.3%, 仅次于 Faster-RCNN (95.7%) 和 YOLOv3 (91.8%), 表明在目标检测任务中具有较高的检出率。

综上所述, DES-YOLO 不仅在 Precision、Recall 和 mAP 等指标上表现出色, 还在计算效率和模型复杂度之间实现了良好的平衡, 具备在实际应用中推广的潜力。

如图 6 所示, 尽管改进后的模型 DES-YOLO 参数量增加, 导致 GFLOPs 高于其他轻量化模型, 但在计算资源允许的范围内, 改进后的网络模型在小目标检测方面的 mAP@0.5 显著提高。相比于 YOLOv5s 和 YOLOv5m, DES-YOLO 的检测精度有了明显提升, 而 GFLOPs 的增加也在可接受的范围内。在适当增加计算

量的前提下, 改进后的模型展现了出色的检测性能。

4.5 效果分析

DES-YOLO 模型在处理遥感目标时表现出了卓越的性能, 简单的目标场景 (图 7(a)), 能够做到不漏检并且检测精度高; 密集的目标场景 (图 7(b)), 通过增加小目标检测层, 显著提高了多目标图像中小目标的检测能力; 部分被阴影遮挡的目标 (图 7(c)), 改进的特征融合策略和注意力机制, 使得模型能够更好地提取目标特征, 从而提高检测精度, 减少漏检和误检; 此外, 在密集目标与稀疏目标同时存在的图像中, DES-YOLO 同样表现出色。如图 7 所示, DES-YOLO 模型在这些不同场景下的检测性能都超过了目前先进的目标检测模型, 特别是在显著减少小目标的漏检和误检方面, 从视觉上展示了其出色的性能。

5 结论

本文提出了一种改进模型—DES-YOLO, 旨在提高图像中小目标的检测精度, 并对算法的适用性进行了详细分析。首先, 模型在网络中嵌入了 DA 注意力模块, 使其能够更好地适应目标形状的变化。在 NWPU VHR-10 数据集和纺织瑕疵数据集上的实验验证表明, 这一改进显著提升了小目标的检测精度。

在增加注意力机制的基础上, 本文还对比了 CIoU、EIoU、SIoU 和 WIoU 四种损失函数, 验证了 EIoU 在提升小目标检测精度和加快收敛速度方面的优越性。此外, 通过增加专门的小目标检测层, 有效提取了目标的浅层特征, 进一步增强了模型对微小目

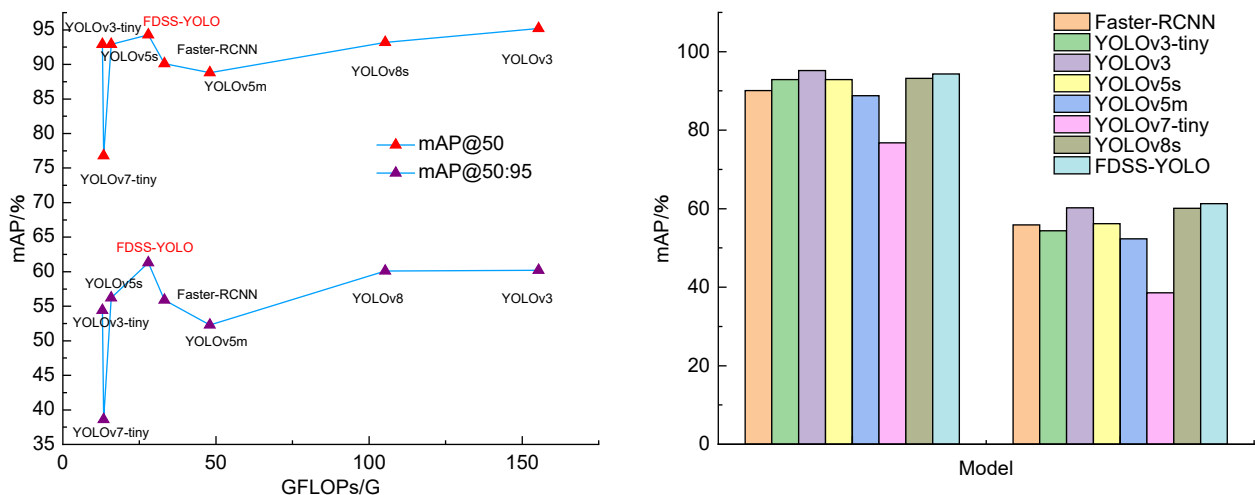


图 6 不同网络模型检测精度对比

Fig. 6 Comparison of detection accuracy of different network models

标的检测能力。

通过在两类数据集上的消融实验, 改进后的模型相比原始 YOLOv5s, 平均精度分别提升了 1.4% 和 1.7%, 充分验证了所提算法的有效性。尽管改进后的模型参数量有所增加, 导致训练速度略有下降, 但实验中采用的分步训练策略在一定程度上缓解了计算量

增加的影响。在小目标检测任务中, 改进模型在不同数据集上均取得了显著的进展。与其他类似算法的对比结果显示, 尽管 DES-YOLO 在某些指标上略逊一筹, 但在计算资源允许的范围内, 其检测精度有了显著提升, 并在计算复杂度与模型规模之间实现了良好的平衡。

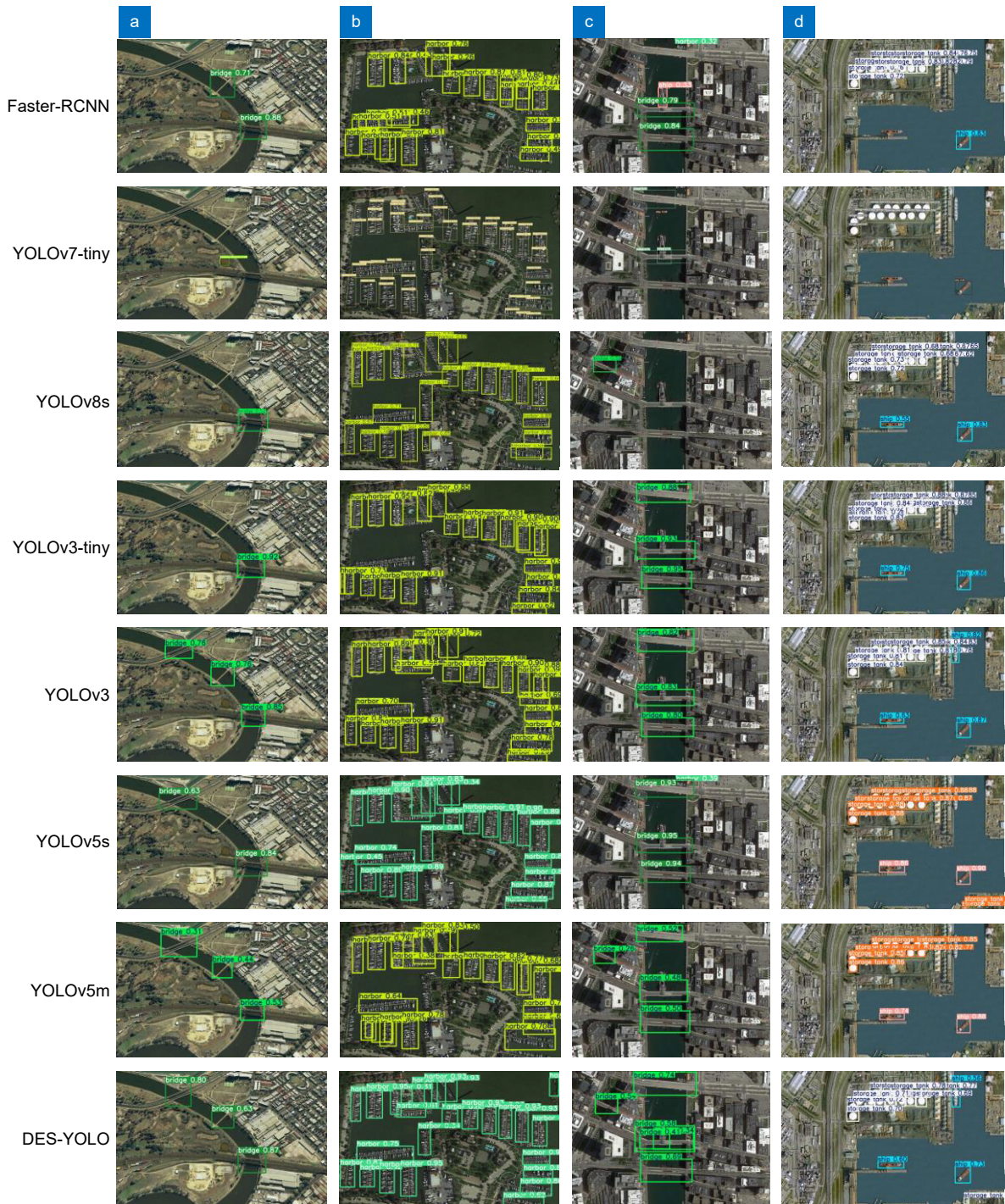


图 7 不同模型的检测效果对比
Fig. 7 Comparison of detection effect of different models

下一步工作将集中于对这些改进部分进行进一步优化, 继续提升小目标检测的精度; 在更多数据集上进行测试, 以涵盖更广泛的检测类别和场景; 同时对模型进行轻量化处理, 进一步提升其训练速度和实际应用性能。

参考文献

- [1] Zhang Y T, Huang D Q, Wang D W, et al. Review on research and application of deep learning-based target detection algorithms[J]. *Comput Eng Appl*, 2023, 59(18): 1–13. 张阳婷, 黄德启, 王东伟, 等. 基于深度学习的目标检测算法研究与应用综述[J]. *计算机工程与应用*, 2023, 59(18): 1–13.
- [2] Tong K, Wu Y Q. Research advances on deep learning based small object detection benchmarks[J]. *Acta Electron Sin*, 2024, 52(3): 1016–1040. 童康, 吴一全. 基于深度学习的小目标检测基准研究进展[J]. *电子学报*, 2024, 52(3): 1016–1040.
- [3] Jiang T, Mu X D, Wei X, et al. Research progress of single-stage small target detection based on deep learning[C]//2022 4th International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Hamburg, Germany, 2022: 893–898. <https://doi.org/10.1109/AIAM57466.2022.00180>.
- [4] Fu H, Fan X T, Yan Z Z, et al. Progress of object detection in remote sensing images based on deep learning[J]. *Remote Sens Technol Appl*, 2022, 37(2): 290–305. 付涵, 范湘涛, 严珍珍, 等. 基于深度学习的遥感图像目标检测技术研究进展[J]. *遥感技术与应用*, 2022, 37(2): 290–305.
- [5] Wang X L, Ban Y, Guo H M, et al. Deep learning model for target detection in remote sensing images fusing multilevel features[C]//IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019: 250–253. <https://doi.org/10.1109/IGARSS.2019.8898759>.
- [6] Dong R C, Xu D Z, Zhao J, et al. Sig-NMS-based faster R-CNN combining transfer learning for small target detection in VHR optical remote sensing imagery[J]. *IEEE Trans Geosci Remote Sens*, 2019, 57(11): 8534–8545.
- [7] Mehta A, Jain R. An analysis of fabric defect detection techniques for textile industry quality control[C]//2023 World Conference on Communication & Computing (WCONF), Raipur, India, 2023: 1–5. <https://doi.org/10.1109/WCONF58270.2023.10235154>.
- [8] Karlekar V V, Biradar M S, Bhangale K B. Fabric defect detection using wavelet filter[C]//2015 International Conference on Computing Communication Control and Automation, Pune, India, 2015: 712–715. <https://doi.org/10.1109/ICCUBEA.2015.145>.
- [9] Alimohamadi H, Ahmadyard A, Shojaee E. Defect detection in textiles using morphological analysis of optimal Gabor wavelet filter response[C]//2009 International Conference on Computer and Automation Engineering, Bangkok, Thailand, 2009: 26–30. <https://doi.org/10.1109/ICCAE.2009.43>.
- [10] Cheng H Q, Xiong J P, Chen J W. Research progress of fabric defect detection[J]. *Comput Era*, 2023, (11): 16–21. 程汉权, 熊继平, 陈经纬. 布匹瑕疵检测算法研究进展[J]. *计算机时代*, 2023, (11): 16–21.
- [11] Zhou H, Jang B, Chen Y X, et al. Exploring faster RCNN for fabric defect detection[C]//2020 Third International Conference on Artificial Intelligence for Industries (AI4I), Irvine, CA, USA, 2020: 52–55. <https://doi.org/10.1109/AI4I49448.2020.00018>.
- [12] 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 (CVPR), Honolulu, HI, USA, 2017: 936–944. <https://doi.org/10.1109/CVPR.2017.106>
- [13] Li J F, Zhu Y W, Chen M X, et al. Research on underwater small target detection algorithm based on improved YOLOv3[C]//2022 16th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 2022: 76–80. <https://doi.org/10.1109/ICSP56322.2022.9965317>.
- [14] Cao K Y, Cui X, Piao J C. Smaller target detection algorithms based on YOLOv5 in safety helmet wearing detection[C]//2022 4th International Conference on Robotics and Computer Vision (ICRCV), Wuhan, China, 2022: 154–158. <https://doi.org/10.1109/ICRCV55858.2022.9953233>.
- [15] Zhang C, Huang Y P, Guo Z Y, et al. Real-time lane detection method based on semantic segmentation[J]. *Opto-Electron Eng*, 2022, 49(5): 210378. 张冲, 黄影平, 郭志阳, 等. 基于语义分割的实时车道线检测方法[J]. *光电工程*, 2022, 49(5): 210378.
- [16] Luo M M, Huang J H, Sun X Y, et al. Small target forest fire recognition method based on deep learning[C]//2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 2023: 593–597. <https://doi.org/10.1109/ICIBA56860.2023.10165608>.
- [17] Li R Z, Chen Y J, Sun C Y, et al. Improved algorithm for small target detection of traffic signs on YOLOv5s[C]//2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), Guangzhou, China, 2023: 339–344. <https://doi.org/10.1109/ICHCI58871.2023.10278065>.
- [18] Chen X, Peng D L, Gu Y. Real-time object detection for UAV images based on improved YOLOv5s[J]. *Opto-Electron Eng*, 2022, 49(3): 210372. 陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. *光电工程*, 2022, 49(3): 210372.
- [19] Ge R, Mao Y L, Li S, et al. Research on ship small target detection in SAR image based on improved YOLO-v7[C]//2023 International Applied Computational Electromagnetics Society Symposium (ACES-China), Hangzhou, China, 2023: 1–3. <https://doi.org/10.23919/ACES-China60289.2023.10249265>.
- [20] Chen Z G, Liu G X, Fan S W. Research on target detection algorithm based on improved YOLO[C]//2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), Xi'an, China, 2022: 485–489. <https://doi.org/10.1109/ICICML57342.2022.10009683>.
- [21] Zhang H Y, Deng L X, Bi L Y, et al. Small object detection algorithm based on improved yolov5[C]//2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT), Jilin, China, 2023: 280–283. <https://doi.org/10.1109/ICCECT57938.2023.10141436>.
- [22] Pu J T, Zhang H Y, Yuan M D, et al. ACN-YOLO: an algorithm for small target detection in aerial images[C]//2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), Haikou, China, 2023: 158–163. <https://doi.org/10.1109/PRAI59366.2023.10331968>.
- [23] Lin T. Focal loss for dense object detection[Z]. arXiv:1708.02002, 2017. <https://doi.org/10.48550/arXiv.1708.02002>
- [24] Wang F L, Su J Y. Based on the improved YOLOV3 small

- target detection algorithm[C]//2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2021: 2155–2159.
<https://doi.org/10.1109/IMCEC51613.2021.9482076>.
- [25] Luan Q L, Chang X Y, Wu Y, et al. PAW-YOLOv7: algorithm for detection of tiny floating objects in river channels[J]. *Opto-Electron Eng*, 2024, 51(4): 240025.
栾庆磊, 常昕昱, 吴叶, 等. PAW-YOLOv7: 河道微小漂浮物检测算法[J]. *光电工程*, 2024, 51(4): 240025.
- [26] Xia Z F, Pan X R, Song S J, et al. Vision transformer with deformable attention[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 2022: 4784–4793.
<https://doi.org/10.1109/CVPR52688.2022.00475>.
- [27] Cheng G, Han J W, Zhou P C, et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors[J]. *ISPRS J Photogramm Remote Sens*, 2014, 98: 119–132.
- [28] Cheng G, Han J W. A survey on object detection in optical remote sensing images[J]. *ISPRS J Photogramm Remote Sens*, 2016, 117: 11–28.
- [29] Cheng G, Zhou P C, Han J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Trans Geosci Remote Sens*, 2016, 54(12): 7405–7415.
- [30] Tianchi. Smart diagnosis of cloth flaw dataset[EB/OL]. 2020[2024-12-10].
<https://tianchi.aliyun.com/dataset/dataDetail?dataId=79336>.
天池. 布匹瑕疵检测数据集[EB/OL]. 2020[2024-12-10].
<https://tianchi.aliyun.com/dataset/dataDetail?dataId=79336>.

作者简介



郑华伟 (1999-), 男, 硕士研究生, 主要从事图像处理, 目标检测研究。

E-mail: 22211030375@stumail.xsyu.edu.cn



高建邦 (1990-), 男, 博士, 助理教授, 主要研究方向为图像处理、信号处理、故障诊断。

E-mail: gjbang2008@126.com



【通信作者】王飞 (1985-), 男, 博士, 讲师, 从事图像处理、视频分析、信号处理和各种嵌入式设备相关算法的软件开发。目前主要研究方向为图像处理、信号分析与算法优化。

E-mail: 200102@xsyu.edu.cn



扫描二维码, 获取PDF全文

DES-YOLO: a more accurate object detection method

Zheng Huawei, Wang Fei*, Gao Jianbang



Remote sensing target detection

Overview: In image analysis, detecting objects accurately remains a significant challenge due to the complexity of backgrounds, the small size of targets, and their dense distribution. To address these issues, we propose an advanced detection method named DES-YOLO. This method incorporates several innovative techniques to enhance the performance of object detection in remote sensing imagery. Firstly, we introduce a deformable attention module (DAM), which allows the network to dynamically adjust its focus on crucial areas of the image. This module enables the network to better recognize and localize objects by concentrating on significant regions and ignoring irrelevant background noise. Secondly, we implement the efficient intersection over union (EIoU) loss function, designed to mitigate the influence of low-quality samples. This loss function improves the generalization ability and detection accuracy of the model, ensuring more precise object localization. Furthermore, we augment the network head with an additional shallow feature map layer of 160 pixel×160 pixel. This enhancement specifically targets extracting features from small objects, often challenging to detect in remote-sensing images. By capturing more detailed information, this layer significantly boosts the detection capability for small-sized targets. Additionally, we employ a stepwise training strategy to refine the model's performance progressively. This training approach helps stabilise the learning process and improves the robustness of the model, leading to superior detection outcomes. Our experimental results are compelling. The improved DES-YOLO model demonstrates a 1.4% increase in the mean average precision (mAP@0.5) on a standard remote sensing dataset. To further validate the model's effectiveness, we conducted extended experiments on a textile dataset, where the model achieved an impressive mAP@0.5 increase of 1.7%. These results not only highlight the improvements brought by our method but also confirm its versatility and applicability to various types of datasets. In conclusion, DES-YOLO represents a significant advancement in object detection, offering enhanced accuracy and reliability. Integrating the deformable attention module, EIoU loss function, shallow feature enhancement, and stepwise training collectively contribute to its superior performance. Our research demonstrates the potential of DES-YOLO to set a new benchmark in object detection, paving the way for future developments and applications.

Zheng H W, Wang F, Gao J B. DES-YOLO: a more accurate object detection method[J]. *Opto-Electron Eng*, 2024, 51(11): 240212; DOI: [10.12086/oe.2024.240212](https://doi.org/10.12086/oe.2024.240212)

Foundation item: Project supported by the Innovation and Practical Ability Cultivation Program for Postgraduates of Xi'an Shiyou University (YCS23214252)

School of Electronic Engineering, Xi'an Shiyou University, Xi'an, Shaanxi 710065, China

* E-mail: 200102@xsyu.edu.cn