

DOI:10.19651/j.cnki.emt.2417451

基于改进 YOLOv8 的无人机红外目标检测算法^{*}张瑞芳^{1,2} 刘占占² 程小辉³ 赵虹^{1,2}

(1. 广西高校先进制造与自动化技术重点实验室(桂林理工大学) 桂林 541006; 2. 桂林理工大学机械与控制工程学院 桂林 541006; 3. 桂林理工大学信息科学与工程学院 桂林 541006)

摘要: 针对无人机航拍红外图像中因为噪声干扰、光照波动和复杂背景带来的目标检测困难的问题,提出了一种基于 YOLOv8 的无人机红外目标检测算法 SDE-YOLOv8。首先,引入了 YOLOv10 中的 SCDowm 模块让每个尺度最大化地保留上下文的语义信息;其次,引入动态上采样器 DySample 来提升模型对于图像细节的敏感度;同时引入三重注意力机制改进 C2f,来强化模型对空间和通道维度之间关系的理解和复杂数据的处理能力;最后,设计了轻量级解耦头 Efficient_Head 模块,确保了检测精度的同时大幅度减少了模型参数。实验结果表明,改进后的算法 mAP50 达到 83.7%,较 YOLOv8n 提高了 4.2%,精确率提升了 1.2%,召回率提升了 3.8%,浮点运算次数下降了 2.5%,FPS 达到了 323.17 fps 的检测速度,充分说明改进算法整体性能优于其他主流算法,能更好的完成无人机红外目标检测任务。

关键词: 红外目标检测;YOLOv8;注意力机制;语义信息;无人机

中图分类号: TP391;TN219 **文献标识码:** A **国家标准学科分类代码:** 520.6

UAV infrared target detection algorithm based on improved YOLOv8

Zhang Ruifang^{1,2} Liu Zhazhan² Cheng Xiaohui³ Zhao Hong^{1,2}

(1. Key Laboratory of Advanced Manufacturing and Automation Technology (Guilin University of Technology), Education Department of Guangxi Zhuang Autonomous Region, Guilin 541006, China; 2. School of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541006, China; 3. School of Information Science and Engineering, Guilin University of Technology, Guilin 541006, China)

Abstract: In order to solve the problem of difficulty in target detection caused by noise interference, illumination fluctuation and complex background in UAV aerial infrared images, an infrared target detection model for UAV based on YOLOv8 was proposed. Firstly, the SCDowm module in YOLOv10 was introduced to maximize the preservation of contextual semantic information for each scale. Secondly, the dynamic upsampler DySample was introduced to improve the sensitivity of the model to image details. At the same time, the triplet attention mechanism is introduced to improve C2f to strengthen the model's understanding of the relationship between spatial and channel dimensions and the processing ability of complex data. Finally, a lightweight decoupling head Efficient_Head module is designed to ensure the detection accuracy and greatly reduce the model parameters. Experimental results show that the improved algorithm mAP50 reaches 83.7%, which is 4.2% higher than YOLOv8n, the accuracy is increased by 1.2%, the recall rate is increased by 3.8%, the number of floating point operations is reduced by 2.5%, and the FPS reaches the detection speed of 323.17 fps, which fully shows that the overall performance of the improved algorithm is better than that of other mainstream algorithms, and it can better complete the task of UAV infrared target detection.

Keywords: infrared target detection; YOLOv8; attention mechanisms; semantic information; drone

0 引言

随着深度学习检测技术的不断进步,无人机目标检测在计算机视觉领域中日益重要,尤其在军事侦察、交通巡查和地质勘探等多个应用场景中发挥着关键作用^[1]。目前,无人机目标检测主要运用在可见光图像上的处理。然而,

可见光图像常常受到天气、光照变化和背景因素等环境干扰,影响检测精度。尤其在夜间环境下,这些问题更为显著。低光照、能见度下降和背景杂乱等因素,使得无人机在夜间执行目标检测任务时面临更大的挑战^[2]。

然而,无人机红外目标检测得益于红外成像技术的优势,如卓越的抗干扰能力、广阔的搜索范围以及不受昼夜影

收稿日期:2024-11-26

^{*} 基金项目:国家自然科学基金(61662917)、广西科技计划重点研发项目(桂科 AB17195042)、广西中青年教师基础能力提升项目(2018KY0248,2020KY06026)、广西建筑新能源与节能重点实验室项目(桂科能 15-J-21-1)资助

响的持续工作特性,成为了弥补夜间任务空白的有效手段^[3]。因此为无人机红外目标检测专门设计网络结构显得尤为重要,具有显著的应用价值和实际意义^[4]。

近年来,许多研究者在无人机红外目标检测领域提出了创新算法。Cao 等^[5]设计了一种名为 YOLO-TSL 的红外目标检测算法,巧妙地将 Slim-Neck 架构和一次性聚合策略融合,使模型参数量减少了 0.5 MB,同时 mAP0.5 达到了 81.2%。Wang 等^[6]则推出 PHSI-RTDETR 的算法,通过引入 Slim Neck-SSFF 架构提升对复杂红外目标的识别,使其在密集目标检测方面超越了 RTDETR 和 YOLOv8,提高了 3.81% 和 2.67% 的精度,然而过大的参数量限制了它在低资源设备上的部署。Aibibu 等^[7]提出了一种高效 ERGW-Net 网络,通过结合 Res-Net、Inception-Net 和 YOLOv8 的优势专攻道路小目标检测,改进后的算法实现了 81.5% 的平均精度,然而,在识别红外图像特征差异较大的电动自行车时,仍存在漏检和误检的问题。Zhang 等^[8]提出了一种基于 Picodet 的改进红外小目标检测方法,采用轻量级 LNet 网络作为特征提取的骨干网络,这有助于解决弱小目标导致漏检和误检率高的问题,但是改进后的算法 mAP0.5 为 77.51%。

综上所述,上述方法在无人机红外目标检测领域提供了创新的思路。然而,现有的检测方法仍存在一些不足之处。首先,普通卷积在处理小目标时可能无法捕捉到足够的细微特征,可能导致关键细节丢失,识别效率下降^[9]。其次,对于复杂场景中的目标,上采样的策略可能强化背景噪音,使得小目标的边缘和微小特性变得模糊不清,这无疑加大了目标与背景间的辨识难度。同时,在小目标密集的环境中,现有的算法往往难以高效整合并突出小目标的核心特征,从而影响了对小目标的精确检测能力^[10]。此外,现

有的解耦技术在处理小目标位置精度方面比较出色,但是原模型的检测头参数较大。因此,为了提高无人机红外目标检测在更短的时间内实现较高的检测精度,本文提出了一种基于 YOLOv8 新型的检测方法,即 SDE-YOLOv8。

算法主要改进如下:

1)为了解决普通卷积难以充分捕捉到足够的细微特征,进而可能导致识别精度下降的问题,引入了 SCDwn 模块,这一改进确保了信息流的有效传递,最大程度地保留了细节信息的完整性。

2)针对密集目标和复杂背景带来的目标与背景区分困难的问题,使用了 DySample 动态上采样器,这种动态机制能够增强了对小目标的边缘和细节的感知能力,从而提升对小目标特征的精确识别能力。

3)考虑到红外图像特征变化大可能导致漏检的问题,重构了以 Triple Attention 为核心的 C2f 模块,进一步增强关键特征的融合,从而提高了对小目标检测的能力。

4)针对模型参数大的问题,设计了高效且轻量化的解耦头 Efficient_Head 模块,通过共享卷积实现检测头的轻量化,以保持检测精度的同时显著降低模型参数。

1 YOLOv8 目标检测算法

YOLO 是一种实时目标检测算法。YOLOv8 在保留了 YOLO 系列快速定位优点的同时,也对结构做了多项改进。YOLOv8 通过采用多尺度特征金字塔结构,捕获不同尺度的目标。YOLOv8 的模型结构主要由三部分组成:Backbone 部分、Neck 部分和 Head 部分。其中 Backbone 部分用于提取图像的细节和高级特征。Neck 部分用于获取不同尺度下的目标信息,以提高检测的准确性和鲁棒性。Head 部分用于生成目标检测的结果。YOLOv8 结构图如图 1 所示。

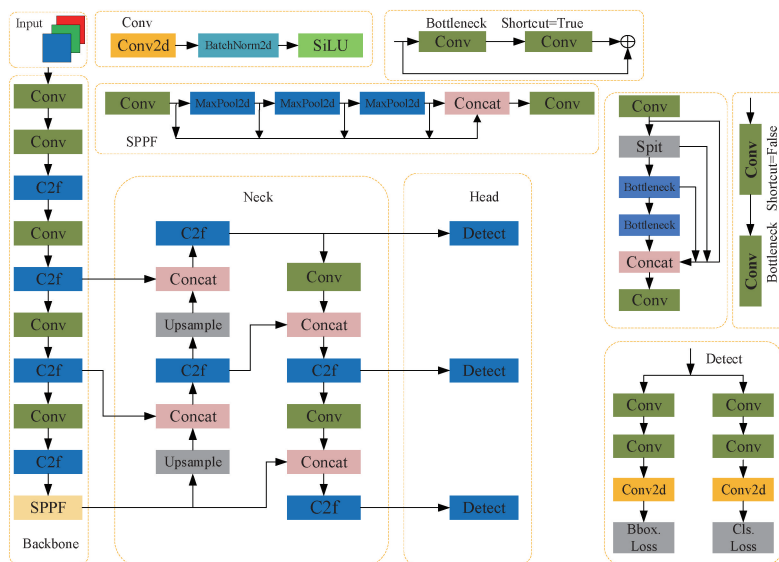


图 1 YOLOv8 结构

Fig. 1 Structure of YOLOv8

2 SDE-YOLOv8 网络架构

根据无人机红外目标检测存在的问题,本文基于 YOLOv8 提出了一种高效的 SDE-YOLOv8 算法。首先,该模型在骨干特征提取网络中采用 SCDwn 模块替换了原始普通的卷积,该模块主要将空间与通道进行解耦,利用深度卷积进行下采样,以减少了计算开销和参数数量,同时保持了性能。其次,在 neck 特征融合阶段,提出一个高效的动态上采样算子。该算子在不增加参数数量的基础上,通过寻找每个上采样点的正确语义聚类来进行采样过程,从而实现高效的上采样。通过重新设计了 C2f 模块,引进了 C2f_TripletAt 模块。该模块是利用 3 个分支结构实现通道与空间之间的交互,并以微不足道的计算开销编码通道和空间的信息。YOLOv8 虽然采用了解耦检测头,但也增加了计算资源的开销。因此,本文通过重新设计检测头的结构来减少参数量,并提高检测的精度和速度,并将这种优化结构定义为 Efficient_Head。SDE-YOLOv8 如图 2 所示。

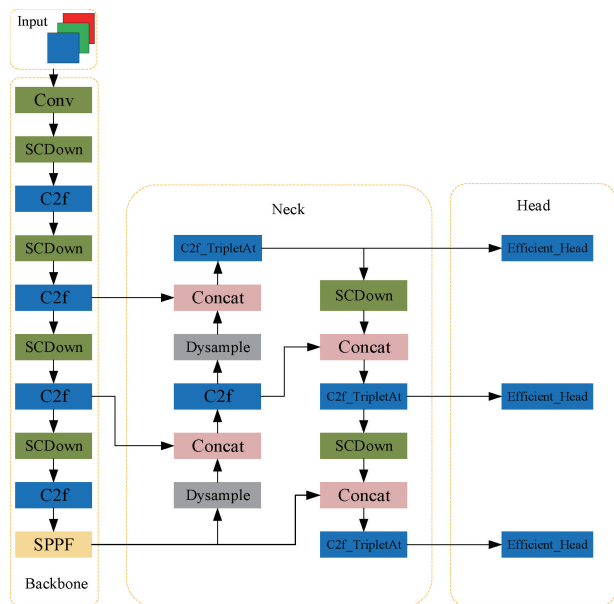


图 2 SDE-YOLOv8 结构

Fig. 2 Structure of SDE-YOLOv8

2.1 SCDwn 模块

在无人机红外航拍图像中,普通卷积的感知能力上存在局限性,难以有效地提取图像特征,特别是小目标在复杂的背景下容易出现漏检和误检,导致目标检测的准确率较低。基于上述原因,本文引进了 YOLOv10 提出的轻量级且高效的 SCDwn^[11]模块。该模块克服了普通卷积的缺点,并且能够加强复杂背景的特征提取。从而更有利于目标检测的应用场景。Conv 模块通常是卷积核为 3 的标准卷积,而 SCDwn 模块采用了带有步幅 1 的 1×1 的普通卷积与带有步幅 2 的 3×3 的深度卷积串联形成的。通

过解耦操作,点卷积用于信道维度的调制,深度卷积用于空间下采样,从而降低了计算成本和参数数量。在降低延迟的情况下,较多的图像信息被保留下来了,如图 3 所示。



图 3 SCDwn 结构

Fig. 3 Structure of SCDwn

2.2 DySample 模块

DySample 是一种轻量级且有效的动态上采样器,与传统基于内核的动态上采样器不同的是 DySample 采用了点采样的方法。由于无人机红外图像里有大量且密集的小目标物体,通常主干网络输出的为多尺度特征,需要将低分辨率特征上采样到高分辨率。因此,高效且轻量化的上采样器将有利于模型对目标的预测。近年来,学者们提出了 CARAFE、FADE 和 SAPA 等采样器来提高输出图像的质量^[12]。但是,这些采样方法会增大计算资源消耗,并且还会受到图像质量的约束。然而, DySample 不仅可以节约计算资源,而且也不需要高分辨率的图像质量。DySample 的设计如图 4 所示,采样点生成器如图 5 所示。

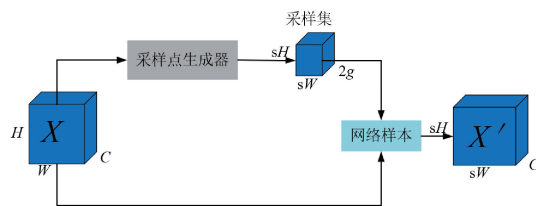


图 4 DySample 结构

Fig. 4 Structure of DySample

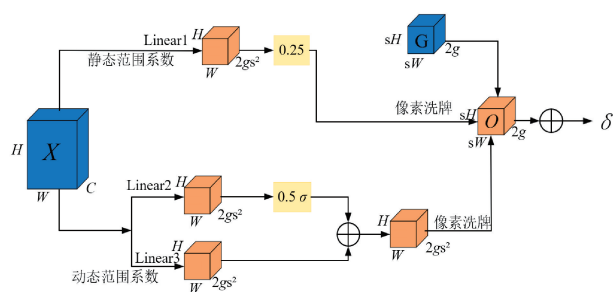


图 5 DySample 采样点发生器

Fig. 5 DySample point generator

在 DySample 结构图中, X 为输入特征, X' 为输出特征, s 为上采样比例因子, G 为原始表格, O 为生成的偏移量, σ 表示 Sigmoid 函数用于生成范围因子, δ 为采样函数。输入特征通过采样点生成器创建采样集,然后利用网络样本函数对输入特征进行重新采样,最后得到上采样特征。同一个采样点生成器生成的采样点具有相同的初始位置,因此, DySample 使用双线性插值技术将输入特征转换为每个采样点的偏移量 O ,以便采样点在输入特征图上

均匀分布,再通过像素洗牌技术结合固定的范围因子来重新调整偏移量 O ,然后与原始网格位置 G 相加得到采样集。如果不做处理,采样点偏移通常会重叠,从而导致预测伪影。除了线性层和像素洗牌,还引入了动态范围因子,用来调节偏移量 O 。偏移量 O 乘以范围因子 0.25 以减少重叠,从而解决了采样点的初始位置及偏移范围的问题。因此,本文将 DySample 与颈部相结合,以较低的计算成本实现高实用性和快速检测能力。

2.3 C2f_TripletAt 模块

红外图像通过向外辐射的热量获得的,因此其分辨率通常低于可见光图像。低分辨率可能导致目标信息的表达不够充分,因此在特征融合和信息传递过程中,低级特征的表达可能会受到抑制,尤其是在复杂环境和小目标的场景下,这种分辨率的不足显著的降低了目标检测的准确性。为了解决这一问题,本文引入了注意力机制^[13]。然而,传统的注意力机制主要关注通道维度的权重分配,可能导致空间信息的大量丧失。为了弥补这一缺陷,CBAM 模块应运而生,虽然同时考虑了空间和通道维度的信息,但该模块在计算通道注意力权重和空间注意力权重是分开进行的^[14]。为了进一步提升性能,本文提出了三重注意力机制,如图 6 所示。

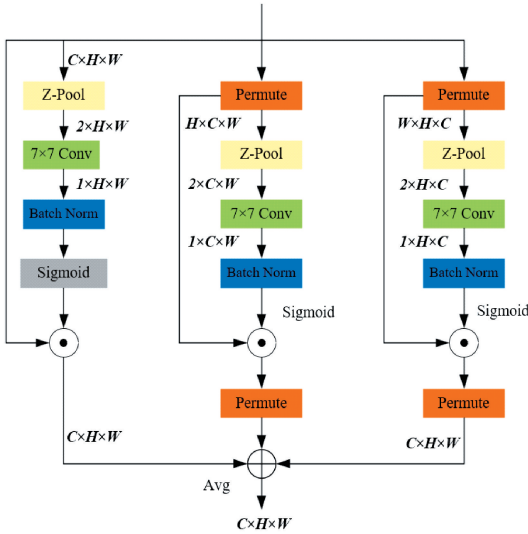


图 6 Triple Attention 结构

Fig. 6 Triple Attention structure

Z 池层的主要作用是将同维度上的平均池化和最大池化的特征结合。这一操作能够在保留张量丰富表达的同时,减少其深度,从而使后续计算更加高效和轻量化。公式如下:

$$Z\text{-Pool}(X) = [\text{MaxPool0d}(X), \text{AvgPool0d}(X)] \quad (1)$$

式中: $\text{MaxPool0d}(X)$ 表示最大池化, $\text{AvgPool0d}(X)$ 表示平均池化, $0d$ 是进行 Max 和 Average 池化操作的第 0 维。

第 1 个分支是空间注意力。输入张量通过 Z-Pool 层

后,通道数减少到 2,得到 \widehat{A}_1 。再通过卷积、批量归一化层,得到 $1 \times H \times W$ 的中间张量。该张量通过 Sigmoid 函数获得 $1 \times H \times W$ 的空间注意力权重。生成的张量在保存原来形状的同时也包含了空间注意力信息。公式如下:

$$A_{top} = \widehat{A}_1 \sigma(\phi_1(\widehat{A}_1)) \quad (2)$$

第 2 个分支是通道维度 C 和宽度维度 W 之间的交互。首先,输入张量 $C \times H \times W$ 沿 W 轴逆时针旋转 90° ,表示为 \widehat{A}_2 ,形状为 $H \times C \times W$ 的张量。将张量通过 Z-Pool 层后,得到了 \widehat{A}_2^* 。再通过普通卷积和批量归一化层,得到 $1 \times C \times W$ 的中间张量。中间张量通过 Sigmoid 函数以获得注意力权重。这些权重按元素乘以 \widehat{A}_2 ,然后沿 W 轴顺时针旋转 90° ,以获得中间分支的输出。公式如下:

$$A_{middle} = \widehat{A}_2 \sigma(\phi_2(\widehat{A}_2^*)) \quad (3)$$

第 3 个分支用于通道维度 C 和高度维度 H 之间的交互。输入张量沿着 H 轴逆时针旋转 90° ,表示为 \widehat{A}_3 。Z-Pool 采用 \widehat{A}_3 作为输入,得到 $2 \times H \times C$ 的形状,表示为 \widehat{A}_3^* 。通过卷积层和批量归一化,再有 Sigmoid 激活层将中间张量作为输入并生成注意力权重,然后这些权重按元素乘以 \widehat{A}_3 ,接下来,注意力机制将生成的张量沿 H 轴顺时针旋转 90° ,以获得最后分支的输出。

$$A_{bottom} = \widehat{A}_3 \sigma(\phi_3(\widehat{A}_3^*)) \quad (4)$$

式中: σ 表示 Sigmoid 函数, ϕ_1 、 ϕ_2 和 ϕ_3 表示三元组注意力的 3 个分支中核大小为 k 定义的标准二维卷积层。

最后三支的细化张量通过平均运算聚合得到一个特征图。公式如下:

$$y = \frac{1}{3} (\widehat{A}_1 \omega_1 + \widehat{A}_2 \omega_2 + \widehat{A}_3 \omega_3) \quad (5)$$

式中: ω_1 、 ω_2 和 ω_3 是 3 个交叉维度的注意力权重。

Bottleneck_TripletAt 模块是以 Triple Attention 为基础,对传统 Bottleneck 模块进行重构。该模块将两个普通卷积和 Triple Attention 注意力顺序连接,最后通过残差连接机制进行特征融合。C2f_TripletAt 模块将原网络 C2f 模块中的 Bottleneck 替换为改进的 Bottleneck_TripletAt 模块。该模块帮助模型更有效地融合不同层次的特征。Bottleneck_TripletAt 结构如图 7 所示,C2f 模块结构如图 8(a)所示,改进后的模块如图 8(b)所示。

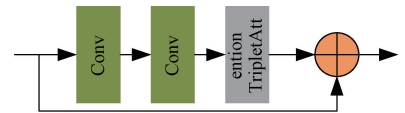


图 7 Bottleneck_TripletAt 结构

Fig. 7 Structure of Bottleneck_TripletAt

2.4 Efficient_Head 模块

YOLOv8 的解耦头是网络不可或缺的结构之一。目前,许多研究者在小目标检测方面增加一个不同规模的额

外检测头。这可能会增加计算资源的大量消耗,不利于部署在无人机平台上^[15]。YOLOv8 采用解耦双分支结构,在提高了特征提取的准确性的同时,也提高了 YOLOv8 在各种对象检测挑战中的实用性。在保持原始解耦头精度的基础上,为了减少参数数量和计算复杂性,本文设计了一种新颖的解耦头,如图 9 所示。本文提出了一种共享参数的方法,该方法将两个分支中的四个普通卷积简化成了单分支的两个卷积,但是预测和分类仍然是解耦的状态。这种方法不仅降低了网络结构的冗余,而且还能提高算法的精度。

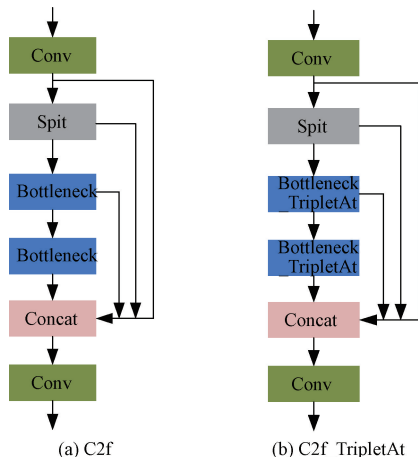


图 8 C2f 与 C2f_TripletAt 结构

Fig. 8 Structure of C2f and C2f_TripletAt

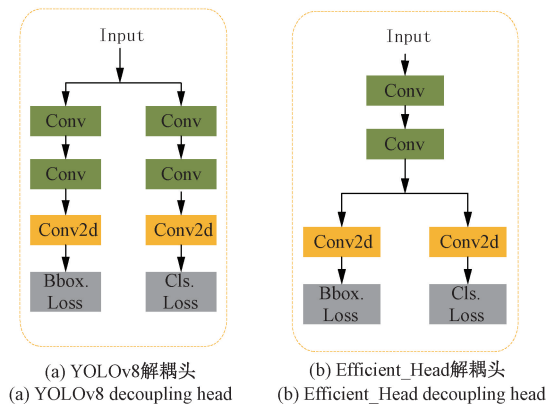


图 9 YOLOv8 与 Efficient_Head 解耦头结构

Fig. 9 Structure of the decoupling head between YOLOv8 and Efficient_Head

3 实验结果及对比分析

3.1 数据集介绍

为了评估 SDE-YOLOv8 检测算法在无人机红外航拍图像目标检测领域的有效性,选用 HIT-UAV 数据集进行验证。HIT-UAV 是一个专为无人机红外目标检测设计的数据集。该数据集包含了各种真实的场景和天气条件,为了能够更好的适应红外目标检测的任务^[16]。这些图像的分辨率为 640×512 ,由五种不同的类别组成:人、汽车、自

行车、其他车辆和“DontCare”,是从 60~130 m 的高空,30 至 90 度的角度捕获的。这些图像在不同的高度下具有不同的大小和形状。这使模型更好地适应不同尺度的输入,以提高模型的泛化能力和鲁棒性^[17]。

本文将 HIT-UAV 数据集分为了 3 部分。数据集包含 2 898 张图像,按 7:2:1 的比例分为训练集、验证集、测试集。本文随机选取 HIT-UAV 数据集中的 2 029 张图片为训练集,579 张图片为验证集,290 张图片为测试集。无人机红外数据集部分图像如图 10 所示。(注:其中第 1 行依次为:道路上的汽车;十字路口的汽车;道路上的其他车辆,第 2 行依次为:汽车和道路上的人;赛道上的人;游乐场里的人,第 3 行依次为:停车场内的汽车;社区内的汽车;道路上的汽车和自行车)



图 10 HIT-UAV 数据集的部分图像

Fig. 10 Part of the image of the HIT-UAV dataset

3.2 实验环境

本文实验使用 Pytorch 框架,采用 GPU 进行实验。实验所处软硬件环境为 Windows11 操作系统、Pytorch—2.0.1 深度学习框架、NVIDIA GeForce RTX 4060 Laptop GPU、Python—3.9.13 和 CUDA11.8。训练基本参数设置如下:总轮数为 300、批量大小为 16、线程为 4、图片尺寸为 640、学习率为 0.01、参数为 0.01、动量因子为 0.937。

3.3 评价指标

本文通过 6 个指标来评估 SDE-YOLOv8 的性能,即模型的准确率(P)、召回率(R)、平均精度(mAP)、参数(M)、每秒帧数(FPS)和浮点运算($GFLOPs$)。其中, mAP 是评估模型优劣的关键指标,参数和 $FLOPs$ 分别表示模型的大小和计算的复杂程度, FPS 表示模型检测的实时性的主要指标。(注:文中 mAP 均为 mAP_{50})以下是评估这些参数的公式:

$$P = \frac{TP}{TP + FP} \quad (6)$$

式中: P 表示正确检测的正样本占有所有预测为正样本的比例。 TP 表示正确检测到的正样本的数量, FP 表示错误把负样本标记为正样本的数量。

$$R = \frac{TP}{TP + FN}$$

(7)

式中: R 表示正确检测的正样本占有所有真正为正样本的比例。 FN 表示错误把正样本标记为负样本的数量。

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i$$

(8)

式中: mAP 是指在所有类别上计算的 AP 的平均值。该指标越高,表明模型检测的效果越好。

$$FPS = \frac{N}{T}$$

(9)

式中: FPS 表示目标检测的速度。其中, N 表示已处理图像的数量, T 表示总共处理需要的时间。

参数越小,该网络的所需的硬件设备配置就越低,更容易在低端设备上部署和移植。GFLOPs 数值越小,所需推理和检测的计算资源就越小。

4 实验过程

4.1 消融实验

为了验证本文提出模块的有效性,在 HIT-UAV 数据集上进行了消融实验。基于 YOLOv8n 原算法进行逐项消融,实验结果如表 1 所示。v8n 表示原始算法 YOLOv8n。

表 1 消融实验
Table 1 Ablation experiments

网络模型	$P/\%$	$R/\%$	Parameters/M	GFLOPs/G	$mAP50/\%$	$FPS/bit=16$
v8n	89.2	76.4	3.0	8.1	79.5	336
v8n-1	85.5	78.3	2.57	7.9	81.9	303
v8n-2	81.1	73.4	3.0	8.5	77.9	297
v8n-3	86.4	76.1	3.0	8.5	80.5	360
v8n-4	89.5	73.8	2.6	8.5	79.3	208
v8n-5	84.6	77.7	2.5	7.9	81.4	232
v8n-6	88.1	75.9	2.5	7.9	81.0	313
v8n-7	90.4	80.2	3.3	7.9	83.7	323

4.2 对比试验

为了证明本文提出改进算法 SDE-YOLOv8 的优越性,在 HIT-UAV 数据集上与 Faster R-CNN^[18]、RT-DETR^[19]、SSD^[20] 和 YOLO 系列算法,YOLOv3-tiny、YOLOv5n、YOLOv6、YOLOv7-tiny、YOLOv8n、YOLOv9-t、YOLOv10 n 进行了对比,对比实验结果如表 2 所示(粗体字表示最优结果)。在对比实验中,本文保持了一致的训练设置,并且所有的模型在训练过程中均未使用预训练权重^[21]。通过表 2 结果可以看出,本文的算法在无人机红外数据集上的精度达到了 83.7%。相较于 Faster R-

由表 1 的实验结果可知:原始算法的 $mAP50$ 为 79.5%。v8n-1 表示在 v8n 的基础上使用了 SCDown 模块,相比于 v8n 参数降低了 14.3%,FLOPs 减少了 2.5%, $mAP50$ 增加了 2.4%。实验结果证明了该模块更好地保留了图像信息的同时,也提高了算法的精度。v8n-2 表示在 v8n 的基础上使用了 DySample 模块。v8n-3 表示在 v8n 的基础上使用了 C2f_TripletAt 模块,相比于 v8n 在参数量保持不变的情况下, mAP 增加了 1%,FPS 达到了 360,较原始模型增加了 24,充分的证明了该模块通过空间和通道维度的交互可以有效的融合多尺度的语言信息。v8n-4 表示在 v8n 的基础上使用了 Efficient_Head 模块,相比于 v8n,参数降低了 13.3%,证明了创新的检测头可以优化结构,减少参数量。v8n-5 在 v8n-1 的基础上增加了 DySample 模块,相比于 v8n, $mAP50$ 增加了 1.9%。虽然 DySample 模块在单独使用时效果不佳,但是与 SCDown 模块结合之后,不仅能突出目标的细节特征,而且也能提高目标检测精度。v8n-6 在 v8n-5 的基础上添加了 C2f_TripletAt 模块,相比于 v8n, $mAP0$ 增加了 1.5%。v8n-7 在 v8n-6 的基础上增加了 Efficient_Head,相比于 v8n,P 增加了 1.2%,R 增加了 3.8%,FLOPs 降低了 2.5%, $mAP50$ 达到了 83.7%,FPS 达到了 323。实验结果验证了改进算法的有效性,不仅提高了目标检测精度,而且也满足实时性的要求。(注:粗体字表示最优结果)

CNN、RT-DETR 和 SSD 算法, mAP 值分别提升了 13.5%、3.6% 和 11.6%。与轻量级 YOLO 系列算法对比,SDE-YOLOv8 在参数和 FLOPs 方面有着显著的减少,使得算法在无人机红外目标检测的任务中具有出色的检测效率和响应速度^[22]。尽管 YOLOv5n、YOLOv8n 和 YOLOv10 n 的参数和浮点运算具有较少的数值,但是它们对红外图像下的小目标检测效果不佳。以上实验结果表明,本文提出的算法在保持出色的检测效果的同时,也最大限度的降低了模型的参数和复杂度,提高模型的检测效率,实现了更好的检测性能。

表 2 对比实验

Table 2 Comparative experiments

网络模型	$P/\%$	$R/\%$	Parameters/M	GFLOPs/G	$mAP50/\%$	FPS/bit=16
Faster R-CNN	73.7	67.5	41.2	156.3	70.2	86
RT-DETR-18	81.2	78.5	20	58.3	80.1	120
SSD	75.1	67.8	24.5	87.9	72.1	90
YOLOv3-tiny	83.9	68	12.1	19	74.5	171
YOLOv5n	75.3	82.3	2.6	7.5	81.5	294
YOLOv6n	79.2	69.6	4.2	11.8	72.3	384
YOLOv7-tiny	84.7	70.1	6	13.2	76.1	163
YOLOv8n	89.2	76.4	3	8.1	79.5	336
YOLOv9-t	81.7	70.2	2.25	9.3	77.7	239
YOLOv10n	92.3	71.3	2.7	8.2	80.3	272
SDE-YOLOv8	90.4	80.2	3.38	7.9	83.7	323

4.3 检测结果分析

为了更直观的突出改进算法的优势,从测试集中选取 HIT-UAV 数据集在噪声干扰、光照波动和复杂背景 3 种不同场景的图像以验证算法的性能。选取了 YOLOv8n 算法与 SDE-YOLOv8n 算法进行对比。检测结果如图 11 所示。

YOLOv8n 检测的结果。通过对比第 1 行的检测结果可以看出,SDE-YOLOv8n 在噪声干扰的环境下检测“Persons”和“Bicycle”的目标时,准确率比原算法有了明显的提升。第 2 行的两张照片展示了在光照变化的情况下,SDE-YOLOv8n 依然能够正确的检测出目标。第 3 行展示了 SDE-YOLOv8n 在目标密集的情况下的检测结果,尽管此类场景挑战性较高,但是 SDE-YOLOv8n 仍然能有效地检测出大多数目标的位置和类别。虽然 SDE-YOLOv8n 优于原算法的精度,但是也出现了漏检和误检的情况,但是总体的检测效果比原算法有明显的改善。

为了进一步的说明改进算法的有效性,将改进前后算法的 mAP 以及各种目标的 mAP 进行了对比,对比结果如图 12、13 所示。结果显示,改进后的算法在检测各种目标时的平均精度比原算法有所提高,再次验证了改进算法的有效性。

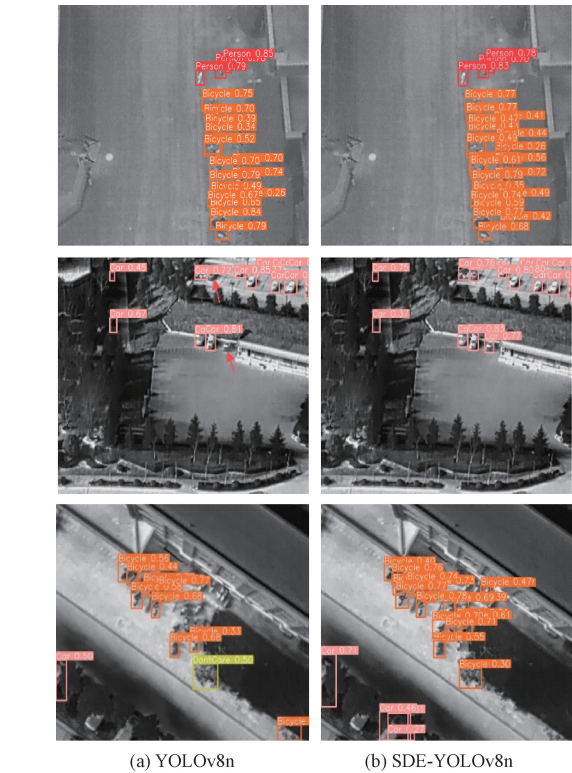


图 11 两种算法在 HIT-UAV 数据集上的对比图

Fig. 11 Comparison of the two algorithms on the HIT-UAV dataset

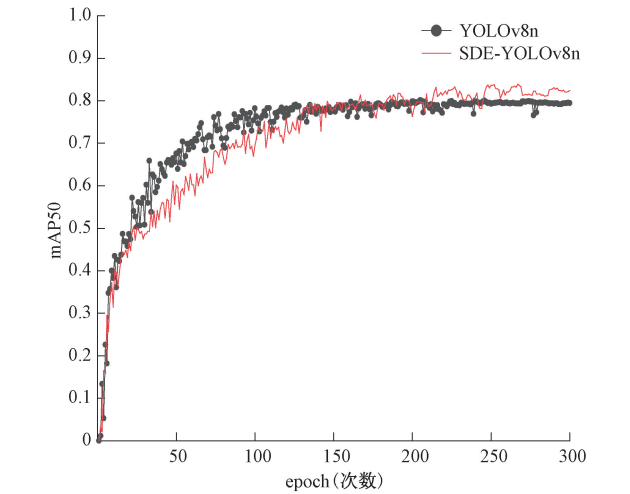


图 12 两种算法在 HIT-UAV 数据集上的 mAP 曲线图

Fig. 12 mAP curves of the two algorithms on the HIT-UAV dataset

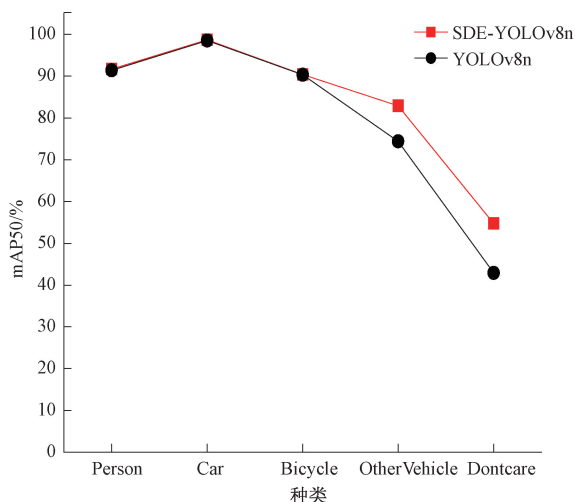


图 13 两种算法在 HIT-UAV 数据集上的 mAP 折线图

Fig. 13 mAP line chart of two algorithms on the HIT-UAV dataset

5 结 论

本研究在 YOLOv8n 的基础上,提出了一种专为红外小目标在复杂背景下检测而设计的新型无人机目标检测算法 SDE-YOLOv8n。首先,SCDown 模块实现了高效的下采样操作,既缩减了模型规模,又保证了检测精度。接着,DySample 模块的独特之处在于其高效地降低了计算负担。在 C2f 模块中,融入 Triplet Attention 机制,有助于融合多尺度信息,有效抑制背景干扰,提升红外目标检测的精确度。最后,本研究改良了检测头,通过简化结构和减少参数,实现了更高的检测速度。实验结果显示,相较于 YOLOv8n,SDE-YOLOv8n 在 mAP50 上有显著提升至 83.7%,同时浮点运算次数下降 2.5%、FPS 提升到 323。尽管如此,SDE-YOLOv8n 在多角度检测方面仍有待拓展。未来工作计划将针对不同红外图像数据集进行模型训练,旨在增强模型的鲁棒性,突破现有限制,持续优化检测。

参考文献

[1] 李冰,杜喜英,王玉莹,等. 基于改进 YOLOv8n 的变电设备红外图像实例分割算法[J]. 电子测量技术,2024, 47(10):151-159.
LI B, DU X Y, WANG Y Y, et al. An instance segmentation algorithm for power equipment infrared images based on improved YOLOv8n[J]. Electronic Measurement Technology, 2024, 47(10): 151-159.

[2] 沈凌云,郎百和,宋正勋,等. 基于 DCS-YOLOv8 模型的红外图像目标检测方法[J]. 红外技术,2024,46(5): 565-575.
SHEN L Y, LANG B H, SONG ZH X, et al. Infrared image target detection method based on DCS-YOLOv8 model[J]. Infrared Technology, 2024, 46 (5): 565-575.

[3] PAN L, LIU T, CHENG J H, et al. AIMED-Net: An enhancing infrared small target detection net in UAVs with multi-Layer feature enhancement for edge computing[J]. Remote Sensing, 2024, 16(10): 1776.

[4] 党宏社,高宛蓉,孟饶辰,等. 融合多尺度特征的轻量级航拍目标检测算法[J]. 电子测量技术,2024,47(16): 92-99.
DANG H SH, GAO W R, MENG R CH, et al. Lightweight aerial target detection algorithm fused with multi-scale features[J]. Electronic Measurement Technology,2024,47(16):92-99.

[5] CAO L, WANG Q, LUO Y H, et al. YOLO-TSL: A lightweight target detection algorithm for UAV infrared images based on triplet attention and slim-neck[J]. Infrared Physics & Technology, 2024, 141: 105487.

[6] WANG S, JIANG H P, LI ZH J, et al. PHSI-RTDETR: A lightweight infrared small target detection algorithm based on UAV aerial photography[J]. Drones, 2024, 8(6): 240.

[7] AIBIBU T Y, JIN H L, ZENG Y L, et al. An efficient rep-style gaussian-wasserstein network: improved UAV infrared small object detection for urban road surveillance and safety [J]. Remote Sensing, 2023, 16(1): 25.

[8] ZHANG Q Q, LI ZH, AN J SH. Real-time recognition algorithm of small target for UAV infrared detection[J]. Sensors, 2024, 24(10): 3075.

[9] 何自芬,陈光晨,陈俊松,等. 多尺度特征融合轻量化夜间红外行人实时检测[J]. 中国激光,2022, 49 (17): 130-139.
HE Z F, CHEN G CH, CHEN J S, et al. Multi-scale feature fusion lightweight nighttime infrared pedestrian real-time detection [J]. Chinese Laser, 2022, 49(17): 130-139.

[10] 樊强强,史再峰,孔凡宁,等. 面向航摄图像目标检测的轻量级特征融合网络[J]. 激光与光电子学进展, 2023, 60(10): 275-284.
FAN Q Q, SHI Z F, KONG F N, et al. Lightweight feature fusion network for aerial image target detection[J]. Advances in Laser and Optoelectronics, 2023, 60(10): 275-284.

[11] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458,2024.

[12] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C]. IEEE/CVF International Conference on Computer Vision, 2023:

- 6027-6037.
- [13] YASIR M, LIU SH W, PIRASTEH S, et al. YOLO ship tracker: Tracking ships in SAR images using lightweight YOLOv8 [J]. International Journal of Applied Earth Observation and Geoinformation, 2024, 134: 104137.
- [14] 王友伟, 郭颖, 邵香迎. 基于改进级联算法的遥感图像目标检测[J]. 光学学报, 2022, 42(24): 203-211.
WANG Y W, GUO Y, SHAO X Y. Object detection in remote sensing images based on improved cascade algorithm[J]. Acta Optica Sinica, 2022, 42(24): 203-211.
- [15] 张寅, 朱桂熠, 施天俊, 等. 基于特征融合与注意力的遥感图像小目标检测[J]. 光学学报, 2022, 42(24): 140-150.
ZHANG Y, ZHU G Y, SHI T J, et al. Small target detection in remote sensing images based on feature fusion and attention[J]. Acta Optica Sinica, 2022, 42(24): 140-150.
- [16] 文思予, 张上, 张朝阳, 等. 基于改进 YOLOv8s 的多尺度检测算法 [J]. 现代电子技术, 2024, 47(15): 133-138.
WEN S Y, ZHANG SH, ZHANG ZH Y, et al. Multi-scale detection algorithm based on improved YOLOv8s[J]. Modern Electronics Technique, 2024, 47(15): 133-138.
- [17] 付军军, 陈林, 童华敏, 等. 基于改进 EfficientNet 的海上风机叶片早期缺陷检测及分类[J]. 电子测量技术, 2024, 47(16): 110-119.
FU J J, CHEN L, TONG H M, et al. Early defect detection and classification of offshore wind turbine blades based on improved EfficientNet[J]. Electronic Measurement Technology, 2024, 47(16): 110-119.
- [18] REN SH Q, HE K M, 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, 2016, 39(6): 1137-1149.
- [19] ZHAO Y A, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [20] LIU W, ANGUELOV D G, ERHAN D M, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, 2016: 21-37.
- [21] DING B L, ZHANG Y H, MA SH. A lightweight real-time infrared object detection model based on YOLOv8 for unmanned aerial vehicles[J]. Drones, 2024, 8(9): 479.
- [22] LI L X, LIU X Y, CHEN X, et al. SDMSEAF-YOLOv8: A framework to significantly improve the detection performance of unmanned aerial vehicle images [J]. Geocarto International, 2024, 39(1): 2339294.

作者简介

张瑞芳, 硕士, 副教授, 主要研究方向为图像处理、电路系统。

E-mail: 2002069@glut.edu.cn

刘占占(通信作者), 硕士研究生, 主要研究方向为图像处理、深度学习。

E-mail: 2120231350@glut.edu.cn

程小辉, 博士, 教授, 主要研究方向为嵌入式系统与物联网。

E-mail: cxiaohui@glut.edu.cn

赵虹, 博士, 副教授, 主要研究方向为工业过程控制、网络控制、智能控制。

E-mail: zhaohong@glut.edu.cn