

一种面向嵌入式边缘计算设备的反无人机目标检测方法

廖嘉仪

(上海大学, 上海 200444)

摘要: 随着无人机技术广泛应用, 其滥用带来的安全风险日益显著。而传统检测和预警技术成本高、易受干扰。深度学习计算机视觉技术为无人机检测提供了高效、精准、低成本优势的新方案。然而, 边缘设备上的无人机检测面临复杂背景、目标小且快、运动模糊及精度与速度平衡等挑战, 本文基于 YOLOv8 模型, 针对边缘设备限制, 引入轻量级动态上采样器 DySample 加强特征融合, 并应用三维注意力机制 SimAM 自适应调节特征权重, 聚焦于目标区域, 旨在提升无人机检测的准确性和鲁棒性, 有效应对边缘计算环境下的检测难题。在 Det-Fly 与 TIBNet 数据集实验数据表明, 本文提出的方法分别提升了 4.5%、3.17%, 且在边缘计算板上效果提升 4.04%、4.34%, 显著提升检测效率, 满足实时反无人机检测的实际需要。

关键词: 反无人机; 嵌入式; 边缘设备; 目标检测; 特征融合

中图分类号: TP391.41 **文献标志码:** A

1 引言

随着工业技术的成熟, 无人机逐渐成为主流, 它们在各大领域得到了广泛应用^[1]。在民用领域, 可以监控人流量, 提高公共交通安全; 在灾害监控领域, 可以监控灾害地点并引导救援人员开展救援活动; 在军事领域可以对敌方无人机进行有效军事侦察与打击。然而, 无人机技术的滥用给关键基础设施和飞行安全带来了严重地安全风险, 尤其在军事重地, 非法或无意入侵无人机严重地影响了国家信息安全^[2], 故反无人机的检测和跟踪至关重要。然而大多数现有的检测和预警技术都基于雷达、射频 (RF) 和声学传感器, 这些技术往往存在成本高、

易受噪声干扰等局限性, 结果可靠性弱。

近年来深度学习发展迅速, 基于计算机视觉的无人机检测和监控已被证明是一种可行的解决方案, 其具有高效的检测速度和精准的检测精度, 并且相对于传统检测方式能够大大降低检测的成本。在边缘计算设备中应用反无人机检测至关重要, 它能够实时、高效地识别并应对潜在的无人机威胁, 减少数据传输延迟, 保护关键区域免受非法侵入或监视, 确保安全与隐私。在领域中诸多学者在经典基础检测框架 RCNN^[3]、YOLO^[4] 上进行改进, 并取得了优秀的效果。褚昭晨^[5] 等基于分块增强特征提取, 将无人机目标各部分特征分块处理, 提取目标更细粒度的特征。于吉锴^[6] 等通过引入差分卷积来减

少模型尺寸,将深浅层特征进行融合,提升小目标检测精度。Zhou^[7]等为对抗复杂的环境合并 SPP 模块到初始检测头以提升准确性并压缩网络。

然而在诸多研究下,基于边缘设备的无人机检测仍存在一些难点,边缘设备视角下的无人机如图 1 所示,①复杂的背景与无人机融为一体,导致摄像头无法捕捉物体;②无人机体积小,飞行速度快,导致摄像头无法准确定位;③算法通常需要部署至边缘计算设备上处理帧,需要平衡模型检测精度与速度。因此需要设计新的检测算法以解决上述问题。



图 1 边缘设备视角下的无人机

本文针对上述问题,以 YOLOv8 模型作为基准研究模型,针对边缘计算设备算力不高、存储空间有限的约束,首先采用轻量级动态上采样器 DySample,充分促进通道间的信息交互,加强全局上下文的特征重组于融合;然后采用三维注意力机制 SimAM,自适应调节每个空间位置的特征权重,强调目标区域,从而帮助模型更好地聚焦于感兴趣的目标区域,提高目标检测的准确性和鲁棒性。

2 基于 DS-YOLOv8 的反无人机目标检测方法

2.1 整体框架

本文提出方法的整体框架如图 2 所示。在 YOLOv8 的基础上,主要修改了 YOLOv8-Neck 部分,将原先的线性上采样 Upsample 层替换为轻量级动态上采样器 DySample,促进模型对上下层特征融合并有利于模型特征的表达。下采样过程卷积特征提取时添加 SimAM 注意力模块,加强模型对语义信息的提取。

2.2 基于 DySample 的上采样模块

DySample^[8]是一种超轻量级且高效的动态上采样器,相比于其他动态上采样器,DySample 不需要引入耗时的动态卷积与用于生成动态内核的额外子网络,而采用点采样方式对数据进行上采样,其过程如图 3 所示。

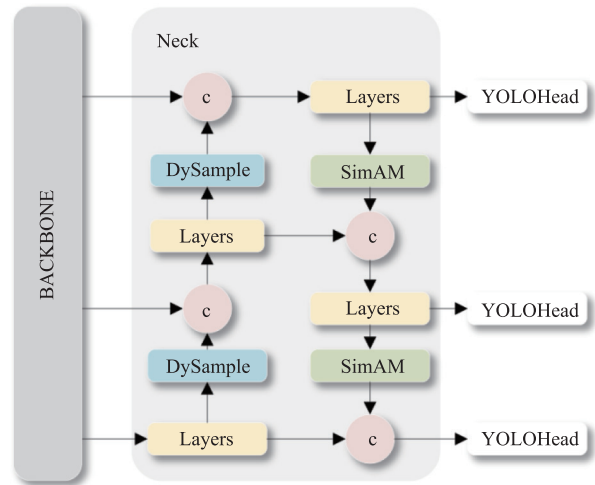


图 2 方法整体框架

给定一个维度为 $C \times H1 \times W1$ 的特征图 X ,缩放因子 s ,和一个输出通道为 $2s2$ 的线性层 L ,该线性层用于生成大小为 $2s2 \times H \times W$ 的偏移量 O ,然后通过像素重排去除网格效应,生成大小为 $2 \times sH \times sW$ 的中间层,并通过与初始特征图经过像素重排后的网络 G 进行像素叠加,获取采样集合 S ,其公式如下:

$$O = \text{Linear}(x) \quad (1)$$

$$S = G + O \quad (2)$$

最后初始的特征图与采样集 S 使用双线性插值,根据 S 的坐标对 X 重新采样成大小为 $C \times sH \times sW$ 的特征图 X_1 ,其过程 grid_sample 公式如下:

$$X_1 = \text{grid_sample}(X, S) \quad (3)$$

DySample 的每个输出像素的位置都是由其对应输入区域内的内容直接影响的,能够精细控制网络提取特征的细节,从而保持局部特征连续性。其中内部采用的空间重排操作能够促进通道间的信息交互,更有利于全局上下文的融合和特征图的表征能力,有效提升网络提取特征能力。

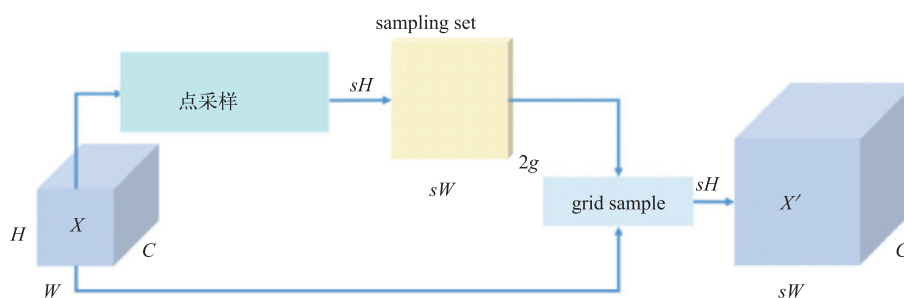
2.3 基于无参数化构建的三维注意力机制 SimAM

由于无人机图像在特定环境的特殊性,单一的通道注意力机制(图 4a)与空间注意力机制(图 4b)无法满足无人机检测精度,通常采用空间-通道混合的三维注意力机制来获取无人机在图像的相对位置权重。但叠加多种复杂注意力机制会导致模型网络结构冗余,最终降低模型检测效率,而 SimAM^[9](图 4c)是一种无参的三维注意力方法,它直接估计三维权重获取每个神经元的重要性来获取网络感兴趣的部分。

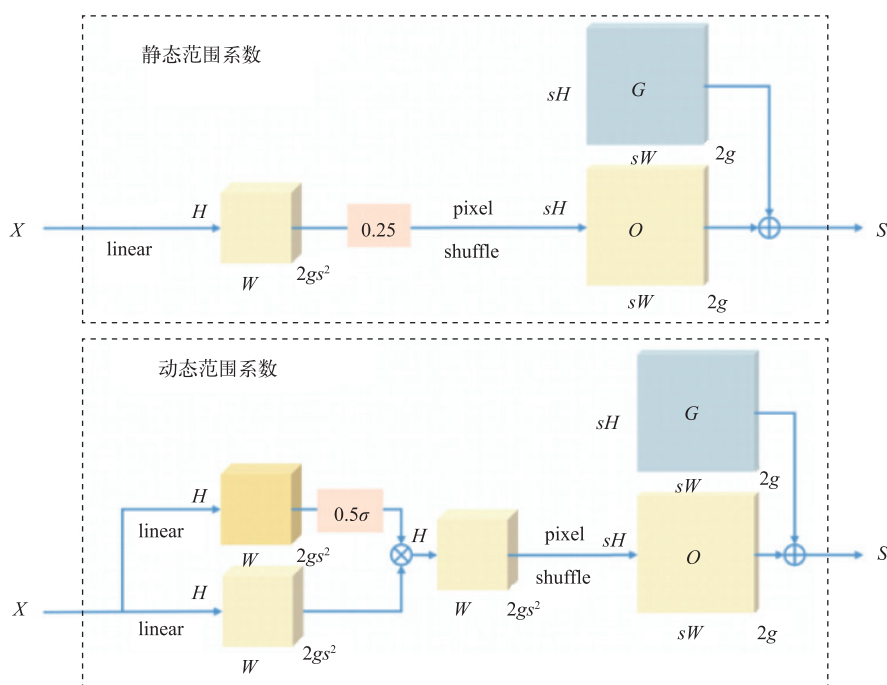
SimAM 首先通过全局平均池化获取每个通道特征，然后通过一个 1D 卷积执行通道之间的交互，将空间与通道的信息相结合，生成三维的注意力权重，最终采用 Sigmoid 激活函数用于生成通道权重，并将其与原始特征

相乘，重新加权特征，使模型能够自适应调整特征，提升网络的特征表达能力。其公式如下：

$$\bar{X} = \text{sigmoid}\left(\frac{1}{E}\right) \otimes X \quad (4)$$



a) 基于DySample的上采样



b) DySample采样点生成器

图3 DySample 上采样步骤

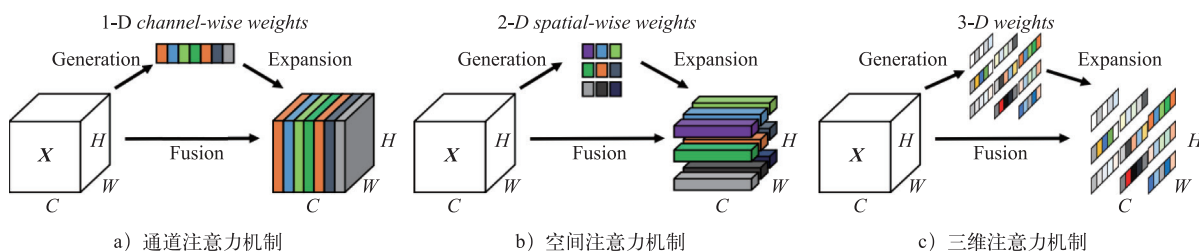


图4 通道、空间、三维注意力机制

3 实验结果与分析

3.1 数据集

本文采用 Det-Fly^[10] 与 TIBNet^[11] 数据集进行训练、验证、测试。Det-Fly 由机载摄像头收集用于空对空无人机探测，其中无人机占图像大小 5% 以下的占整体数据集的 79.86%，TIBNet 由一个地面摄像头收集，用于地对空无人机探测，其中所有无人机目标大小均在图像大小 5% 以下。

3.2 评价指标

本文主要使用目标检测中常用的指标平均精度均值（mean Average Precision, mAP）来评估检测器的总体性能。

为了测试模型在边缘计算设备上的运行速度，在工业中常采用 FPS(Frames Per Second, 每秒帧数) 来衡量模型的推理速度，并使用 Parameter 来表示模型参数大小，GFLOPs 表示每秒执行浮点数运算次数，反映模型算法计算性能。

3.3 实验设备

本文所采用的实验平台详细信息见表 1、表 2。总共训练 100 轮，batchsize 设为 16，每 20 轮保存一次实验结果，并实时保留最优模型。

表 1 训练使用系统信息

系统	GPU	CPU
Ubuntu18.04	RTX 3060 × 2	i7-11700k

表 2 边缘计算设备信息

系统	GPU	CPU
Ubuntu18.04	NVIDIA Maxwell GPU	64-bit Quad-core ARM A57

3.4 注意力机制选择实验

通过探索不同的注意力机制在 YOLOv8 中的性能变化，本文主要选取当前主流的注意力机制网络 CBAM、SENet、Non-Local 进行对比实验，其实验结果见表 3。

表 3 不同注意力机制在不同数据集的实验结果

Model	mAP (%)	Det-fly		mAP (%)	TIBNet	
		Params	FLOPs		Params	GFLOPs
YOLOv8n	91.11	3.01M	8.1G	89.72	3.01M	8.1G
+SENet ^[12]	91.76	3.27M	8.3G	89.93	3.27M	8.3G
+CBAM ^[13]	91.45	3.34M	8.4G	89.97	3.34M	8.4G
+Non-Local ^[14]	92.14	3.46M	8.6G	90.15	3.46M	8.6G
+SimAM	93.43	3.01M	8.2G	92.17	3.01M	8.2G

由实验结果可知，仅添加 SENet 与 CBAM 注意力模块能够微量地提升模型检测效果，但计算量与模型参数提升，而 Non-Local 是一种空间、通道注意力融合模块，虽然使模型在两数据集上的效果分别提升 1.03%、0.43%，但计算量与模型参数大幅度提升，模型检测效率降低。SimAM 是一种无参的注意力模块，模型参数大小不变，计算量仅上升 0.1G 的前提下模型在两数据集的精度分别提升 2.32%、2.45%。这种直接获取神经元权重的三维网

络能够使模型“无痛涨点”，更加关注目标在图像的位置、空间信息，并提取特征与特征之间的相对位置权重，达到提升模型检测效率的目的。

3.5 主流检测网络性能对比实验

为验证改进后的算法在反无人机检测中是否存在优势，将本文的模型 DS-YOLOv8 与目前该领域的优秀检测算法在两数据集上进行比较，实验结果见表 4。

表 4 不同模型在不同数据集的检测结果

Model	mAP (%)	Det-fly		mAP (%)	TIBNet	
		Params	FLOPs		Params	GFLOPs
YOLOv8n	91.11	3.01M	8.1G	89.72	3.01M	8.1G

(续)

		Det-fly		TIBNet		
EdgeYOLO-Tiny ^[15]	88.84	5.82M	27.2G	87.61	5.82M	27.2G
YOLOv9-Tiny ^[16]	94.02	2.01M	7.7G	91.43	2.01M	7.7G
EDGS-YOLOv8 ^[17]	93.44	4.31M	9.9G	90.85	4.31M	9.9G
ADMNet ^[18]	94.82	3.90M	9.1G	89.72	3.90M	9.1G
DS-YOLOv8	95.61	3.04M	8.4G	92.89	3.04M	8.4G

由上述实验可知,在参数与计算量在一个量级的网络中,本文的模型在两数据集中表现最好,相比于基模型分别提升了4.5%、3.17%,且计算量与模型参数没有显著变化。模型效果在两数据集对比于次优网络分别提升0.79%、1.46%,且参数与计算量比ADMNet低,模型检测效率提升显著。

3.6 上采样模块性能对比实验

通过探究不同的上采样模块在YOLOv8中的性能变化,本节主要选取当前优秀的上采样机制,CARAFE、A2U、SAPA-B和FADE进行对比实验,其实验结果见表5。

由实验结果可知,FADE在两数据集上mAP分别提

升较多,但在FLOPs和Params上增加最多,使用较多的计算资源来换取检测精度,CARAFE与SAPA-B方法在提升少量计算量的情况下,提升少量的检测精度。而A2U方法提升了计算量反而降低了检测的精度。本文采用的DySample方法在所有结果中表现最好,对比于基线模型,在两数据集上mAP分别提升2.34%、1.32%,且在FLOPs与Params两参数上增加最少。综上所述,针对少量像素的小目标,DySample采用的点采样方式上采样能够更细粒度地还原小目标的特征,进而有利于模型全局上下文融合和特征图的表征能力,有效提升网络效果。

表5 不同上采样模块在不同数据集的实验结果

		Det-fly		TIBNet		
Model	mAP (%)	Params	FLOPs	mAP (%)	Params	GFLOPs
YOLOv8n	91.11	3.01M	8.1G	89.72	3.01M	8.1G
+CARAFE ^[19]	91.89	3.42M	9.8G	90.07	3.42M	9.8G
+A2U ^[20]	90.46	3.11M	8.5G	89.45	3.11M	8.5G
+SAPA-B ^[21]	92.70	3.14M	9.1G	90.17	3.14M	9.1G
+FADE ^[22]	93.31	3.38M	10.8G	90.84	3.38M	10.8G
+DySample	93.45	3.02M	8.3G	91.04	3.02M	8.3G

3.7 边缘设备测试

为进一步验证本文模型的使用价值,本文将模型部署至边缘计算板,通过摄像头获取物理世界图像,以实现实时检测效果。本文先使用YOLOv8n模型作为初始部

署模型,对两个数据集进行测试,将YOLOv8n.pt文件转换为onnx文件,再通过TensorRT转化为支持边缘计算板推理的引擎YOLOv8n.trt。效果见表6。

表6 模型在边缘计算设备实验效果

Model	数据集	mAP (%)	FPS	Params	GFLOPs
(a) YOLOv8n	Det-fly	91.11	8	3.01M	8.1G
(b) YOLOv8n-1/4		82.69	25	0.36M	1.4G
(c) DS-YOLOv8-1/4		86.73	25	0.37M	1.5G

(续)

Model	数据集	mAP (%)	FPS	Params	GFLOPs
(a) YOLOv8n	TIBNet	89.72	9	3.01M	8.1G
(b) YOLOv8n-1/4		78.57	26	0.36M	1.4G
(c) DS-YOLOv8-1/4		83.91	26	0.37M	1.5G

YOLOv8n 在边缘计算板上精度与在服务器上的结果相差不大,但是在边缘计算板的运行 FPS 仅有 8 ~ 9 帧,尽管模型参数量已经足够小,但仍然无法满足实时检测效果。为了进一步探究小模型在边缘计算板的效果,本文将 YOLOv8n 中所有卷积通道数降低为原来的 1/4 进行训练,结果如表 6 中 (b) 所示,模型参数与计算量大幅

度降低,在 FPS 值上提升至 25 帧,满足模型实时性检测需求,但在 mAP 指标上降低较多。通过添加本文优化的方法如表 5 (c),模型精度在两数据集上分别提升 4.04%、4.34%,且 FPS 几乎没有发生变化,仍保持实时检测所满足的帧率,检测精度也能完美符合实际需要。其检测结果如图 5 所示。



图 5 反无人机目标检测视觉效果

4 总结

本文提出一种面向边缘计算设备的反无人机侦察的目标检测网络 DS-YOLOv8,为了当前检测无人机存在的实际问题,引入了 DySample 来控制网络提取特征的细节,促进通道间的信息交互,更有利于全局上下文的融合和特征图的表征能力。此外本文还引入一种无参数的三维注意力机制网络 SimAM,使网络更加注重于特征与特征之间的相对位置关系,提升模型检测效率。实验结果表明,本算法对比于其他优秀算法具有显著优势,能够很好地运行且精度与速度满足显示实际需要,能够很好应用于现实反无人机的移动实时检测领域。IM

参考文献

- [1] 刘新锋,陈梦雅,李成龙,等.基于计算机视觉技术的无人机检测跟踪方法[J].济南大学学报(自然科学版),2024,38(4):445-455.
- [2] 薛珊,王亚博,吕琼莹,等.基于 YOLOX-drone 的反无人机系统抗遮挡目标检测算法[J].工程科学学报,2023,45(9):1539-1549.
- [3] HMIDANI O, ALAOUI E M I A comprehensive survey of the R-CNN family for object detection[C]//2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet). IEEE,

- 2022; 1-6.
- [4] JIANG P, ERGU D, LIU F, et al. A review of Yolo algorithm developments[J]. *Procedia Computer Science*, 2022, 199: 1066-1073.
- [5] 褚昭晨, 宋韬, 金忍, 等. 基于视觉图像的空对空多无人机目标跟踪[J]. *航空学报*, 2024, 45 (14): 20-35.
- [6] 于吉锴, 黄真亮, 江乐旗, 等. 基于差分卷积和深层次特征融合的表面缺陷分割[J]. *失效分析与预防*, 2024, 19 (3): 149-157.
- [7] ZHOU X, YANG G, CHEN Y, et al. ADMNet: Anti-drone real-time detection and monitoring[C]//2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2023: 3009-3016.
- [8] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample[C]//IEEE/CVF International Conference on Computer Vision. IEEE, 2023: 6027-6037.
- [9] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks[C]//International conference on machine learning. PMLR, 2021: 11863-11874.
- [10] ZHENG Y, CHEN Z, LV D, et al. Air-to-air visual detection of micro-uavs: An experimental evaluation of deep learning[J]. *IEEE Robotics and Automation Letters*, 2021, 6 (2): 1020-1027.
- [11] SUN H, YANG J, SHEN J, et al. TIB-Net: Drone detection network with tiny iterative backbone[J]. *IEEE Access*, 2020 (8): 130697-130707.
- [12] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2018: 7132-7141.
- [13] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//European Conference on Computer Vision (ECCV), 2018: 3-19.
- [14] XIANG S, LIANG Q. Remote sensing image compression with long-range convolution and improved non-local attention model[J]. *Signal Processing*, 2023 (209): 109005.
- [15] LIANG S, WU H, ZHEN L, et al. Edge YOLO: Real-time intelligent object detection system based on edge-cloud cooperation in autonomous vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23 (12): 25345-25360.
- [16] WANG C Y, YE H I, LIAO H Y M. Yolov9: Learning what you want to learn using programmable gradient information[J/OL]. (2024-02-29) [2024-04-11]. <https://arxiv.org/abs/2402.13616?context=cs.CV>.
- [17] HUANG M, MI W, WANG Y. EDGS-YOLOv8: An improved YOLOv8 lightweight UAV detection model[J]. *Drones*, 2024, 8 (7): 337.
- [18] ZHOU X, YANG G, CHEN Y, et al. ADMNet: Anti-drone real-time detection and monitoring[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2023: 3009-3016.
- [19] WANG J Q, CHEN K, XU R, et al. CARAFE: Context-aware reassembly of features[C]//2019 IEEE International Conference on Computer Vision (ICCV), IEEE, 2019: 3007-3016.
- [20] DAI Y T, LU H, and SHEN C H. Learning affinity-aware upsampling for deep image matting[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2021: 6841-6850.
- [21] LU H, LIU W Z, YE Z X, et al. SAPA: Similarity-aware point affiliation for feature upsampling[J/OL]. (2022-12-27) [2024-04-11]. <https://arxiv.org/abs/2209.12866v1>.
- [22] LU H, LIU W Z, FU H T, et al. FADE: Fusing the assets of decoder and encoder for task-agnostic upsampling[C]//European Conference on Computer Vision (ECCV), Springer, 2022: 231-247.

收稿日期: 2024-10-11