



计算机科学

COMPUTER SCIENCE

融合ByteTrack的EAP-YOLOv8无人机Marker点检测与追踪

唐心亮, 潘晓润, 王建超, 苏鹤

引用本文

唐心亮, 潘晓润, 王建超, 苏鹤. 融合ByteTrack的EAP-YOLOv8无人机Marker点检测与追踪[J]. 计算机科学, 2026, 53(3): 266-276.

TANG Xinliang, PAN Xiaorun, WANG Jianchao, SU He. Integrate ByteTrack's EAP-YOLOv8 UAV Marker Point Detection and Tracking [J]. Computer Science, 2026, 53(3): 266-276.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于文本-图像多模态融合的变电所布局图纸图符检测方法](#)

Method for Symbol Detection in Substation Layout Diagrams Based on Text-Image Multimodal Fusion
计算机科学, 2026, 53(1): 206-215. <https://doi.org/10.11896/jsjcx.250200090>

[基于多目标追踪的视频无关人员自动识别](#)

Automatic Recognition of Irrelevant Individuals in Videos Based on Multi-object Tracking
计算机科学, 2025, 52(11A): 241100155-8. <https://doi.org/10.11896/jsjcx.241100155>

[基于特征增强与上下文融合的无人机小目标检测算法](#)

UAV Small Object Detection Algorithm Based on Feature Enhancement and Context Fusion
计算机科学, 2025, 52(11): 131-140. <https://doi.org/10.11896/jsjcx.241000017>

[基于小目标特征增强RT-DETR的SAR图像舰船目标检测方法](#)

Ship Detection Method for SAR Images Based on Small Target Feature Enhanced RT-DETR
计算机科学, 2025, 52(10): 151-158. <https://doi.org/10.11896/jsjcx.250100097>

[改进RT-DETR的遥感图像小目标检测算法](#)

Improved RT-DETR Algorithm for Small Object Detection in Remote Sensing Images
计算机科学, 2025, 52(8): 214-221. <https://doi.org/10.11896/jsjcx.241000019>

融合 ByteTrack 的 EAP-YOLOv8 无人机 Marker 点检测与追踪

唐心亮¹ 潘晓润¹ 王建超¹ 苏鹤²

¹ 河北科技大学信息科学与工程学院 石家庄 050018

² 河北工业大学电气工程 天津 300401

(375555@qq.com)

摘要 随着科技不断发展,无人机的应用越来越广泛,实现无人机的精准动作捕捉成为其核心技术。光学动作捕捉系统在对无人机进行检测与追踪时,由于受到复杂环境、飞行速度等多方面的干扰,会出现对无人机所粘贴的 Marker 点识别不准确的情况。为了解决这一问题,提出一种基于 YOLOv8 改进的目标检测算法 EAP-YOLOv8,以提高 Marker 点识别检测的准确率。首先,在骨干部分构建新型通道注意力机制 MAP-ECA,增强全局视角信息和不同尺度大小的特征,提升了小目标的检测能力;其次,在原有检测头的基础上利用多层次自适应特征融合形成新的检测头 D-SASFF,利用多尺度融合来强化小目标特征信息;最后,设计了损失函数 PLoUv3,通过改进加快了模型收敛速度,提高了小目标检测能力。为验证 EAP-YOLOv8 算法的有效性,在自制数据集上进行实验,结果表明,EAP-YOLOv8 算法在 mAP@0.5 和 mAP@0.5:0.95 上分别达到了 96.5% 和 50.2%,相较于其他算法有显著提升。在此基础上,通过结合多目标追踪算法 ByteTrack 显著提高了 Marker 点的追踪准确率。此外,在公开数据集 MOT16 上进行追踪实验,结果表明,所提模型在 HOTA, MOTA, MOTP 上追踪准确率分别达到了 37.60%, 25.64%, 80.76%, 相较于当前算法有显著提升,为后续实现无人机精准跟踪提供了有效途径。

关键词: EAP-YOLOv8; 无人机检测; Marker 点; 小目标检测; 多目标追踪; ByteTrack

中图分类号 TP391

Integrate ByteTrack's EAP-YOLOv8 UAV Marker Point Detection and Tracking

TANG Xinliang¹, PAN Xiaorun¹, WANG Jianchao¹ and SU He²

¹ School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China

² Electrical Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract With the development of science and technology, drones are more and more widely used, and the realization of accurate motion capture of drones has become its core technology. When the optical motion capture system detects and tracks the UAV, due to the interference of complex environment, flight speed and other aspects, the Marker point pasted by the UAV will be inaccurate. In order to solve this problem, an improved object detection algorithm EAP-YOLOv8 based on YOLOv8 is proposed to improve the accuracy of Marker point recognition detection. Firstly, a new channel attention mechanism MAP-ECA is constructed in the backbone part, which enhances the global perspective information and the characteristics of different scales, and improves the detection ability of small targets. Secondly, on the basis of the original detection head, the multi-level adaptive feature fusion is used to form a new detection head, D-SASFF, and the multi-scale fusion is used to strengthen the feature information of small targets. Finally, the loss function PLoUv3 is designed, which accelerates the convergence speed of the model and improves the detection ability of small targets. In order to verify the effectiveness of the EAP-YOLOv8 algorithm, experiments are carried out on the self-made dataset, and the results show that the EAP-YOLOv8 algorithm reaches 96.5% and 50.2% on mAP@0.5 and mAP@0.5:0.95, respectively, which is significantly improved compared with other algorithms. On this basis, the tracking accuracy of Marker points is significantly improved by combining the multi-target tracking algorithm ByteTrack, and the tracking experiments are carried out on the public dataset MOT16, and the results show that the new model reaches 37.60%, 25.64% and 80.76% on HOTA, MOTA and MOTP, respectively, which is significantly improved compared with the current algorithms, providing an effective way for the subsequent accurate tracking of UAVs.

Keywords EAP-YOLOv8, Drone detection, Marker point, Small target detection, Multi-target tracking, ByteTrack

到稿日期:2024-11-19 返修日期:2025-03-04

基金项目:河北省高等学校科学技术研究项目(QN2023185)

This work was supported by the Hebei Provincial Higher Education Science and Technology Research Project(QN2023185).

通信作者:王建超(15380349158@163.com)

近年来,随着科技的不断进步,无人机发展已日渐成熟,并广泛应用于航拍摄影、军事侦察、环境监控、灾害救援等领域。在复杂环境中,无人机控制技术的核心在于实现无人机的精准定位。光学运动捕捉系统在检测复杂结构的无人机定位方面有明显优势,通过将光学球形标识物(Marker点)标记在无人机的关键点处,得到无人机的位姿信息。由于光学动作捕捉系统受光照影响较大,在室外光照复杂、有遮挡等情况下,关键点目标会丢失,因此,对Marker点的识别至关重要。

目前最常见的Marker点检测方法有霍夫(Hough)变换检测方法^[1]和圆度识别检测方法^[2],但其计算量大、操作相对复杂。随着近几年卷积神经网络的快速发展,深度学习的目标检测算法在计算速度以及稳定性上都要优于传统目标检测算法,故逐渐取代了传统的目标检测算法。基于深度学习的目标检测方法主要分为区域提取的两阶段检测方法和单阶段检测方法。基于区域提取的方法,如R-CNN^[3],Fast R-CNN^[4],Mask R-CNN^[5]等,主要通过候选区域提取器生成目标候选区,并利用卷积神经网络(CNN)对每个区域进行特征提取和分类。单阶段检测方法,如YOLO^[6]和SSD^[7]等,则直接在图像中提取目标位置和类别信息,以实现高效检测。其中YOLOv8系列算法相较于YOLOv9和YOLOv10,在小目标的准确率和处理速度之间找到了平衡,成为很多目标检测任务的首选算法之一。Wang等^[8]针对无人机航拍场景,基于YOLOv8提出了一种新模型UAV-YOLOv8,改进后的模型由原来的3尺度检测改为5尺度检测,有效提升了对小物体的检测性能,但忽略了视角信息的丢失;Liu等^[9]从网络设计、模型压缩两个方面对YOLOv8进行轻量化改进,满足了军事飞机在目标检测方面的实时性要求,但在实际应用中,复杂的背景可能会导致检测精度下降;Yadav等^[10]利用YOLOv8高性能卷积神经网络和DeepSort跟踪算法提出了一个多目标检测与跟踪方法,实现了准确且实时的目标检测,然而该模型在检测小目标时会出现误检、漏检的情况,检测精度略有下降;Zhai等^[11]提出一种优化YOLOv8的微型无人机检测方法,通过切断大目标探测头和冗余网络层,提高了模型在无人机检测方面的整体性能,但随着模型的改进,参数量和计算资源也大大增加;Hu等^[12]提出的无人机目标检测算法PC-YOLOv8-n通过均衡网络检测精度与计算量,增强了网络的类别检测能力,并对长尾分布数据有一定泛化能力,但

是模型的检测精度在无人机快速运动、快速检测的环境中还有待增强;Liu等^[13]在ByteTrack和SORT追踪算法上引入功率自适应卡尔曼滤波PAKF,提高了在线多目标跟踪的精度和鲁棒性;You等^[14]提出了CDS-YOLOv8和改进ByteTrack的车辆检测跟踪算法,提高了复杂场景下的检测精度和跟踪性能,但是模型在通道间信息交互方面有所限制,可能会出现小目标识别精度不高的问题;Yasir等^[15]针对船舶跟踪救援场景,基于YOLOv8和ByteTrack提出了一种鲁棒MOT算法,提升了检测准确率和追踪实时性,但模型在计算量和参数量方面远远高于原模型。上述方法在不同层面上提高了小目标检测追踪的准确性和有效性,但也存在一定问题:首先,通过尺度变化等方法可以更容易捕捉小目标,但是在追求尺度的过程中,图像裁剪、缩放等处理形式容易导致视角信息丢失,从而导致检测精度下降;其次,优化网络层、模型轻量化可以提升检测精度,但是模型的参数量和计算量也有所增加;最后,在多目标跟踪研究中,结合DeepSort等算法可以实施精准监测,满足实时性、准确性,但在光照弱、有遮挡等环境下适应性较弱,容易出现误检、漏检的情况,反映出算法在复杂目标的特征提取与关联上存在不足。

针对上述问题,本文提出一种改进算法EAP-YOLOv8,具体改进如下:

- 1)构建一个新型跨通道注意力机制MAP-ECA,通过重塑通道和批次维度实现通道间的信息交互,增强全局视角信息。
- 2)设计一种新型检测头D-SASFF,用于增强每个层级特征的语义和细节信息,减少模型检测时的误检和漏检。
- 3)构建一种新型损失函数PIoUv3,该损失函数更加关注中高质量锚框,从而提高了小目标的检测精度。
- 4)融合EAP-YOLOv8和ByteTrack算法在公共数据集上进行实验,验证了所提算法的有效性。

1 相关工作

1.1 YOLOV8

YOLOv8的核心优势在于在不进行复杂处理的前提下快速识别定位图像中的目标。YOLOv8在可扩展性基础上通过增加新的骨干网络、新的无锚点检测头和新的损失函数,成为一种适用于广泛目标检测的选择,并以高效的检测能力在计算机视觉领域占据重要地位,在需要实时处理的场景中具有广泛应用能力。每个模块的结构如图1所示。

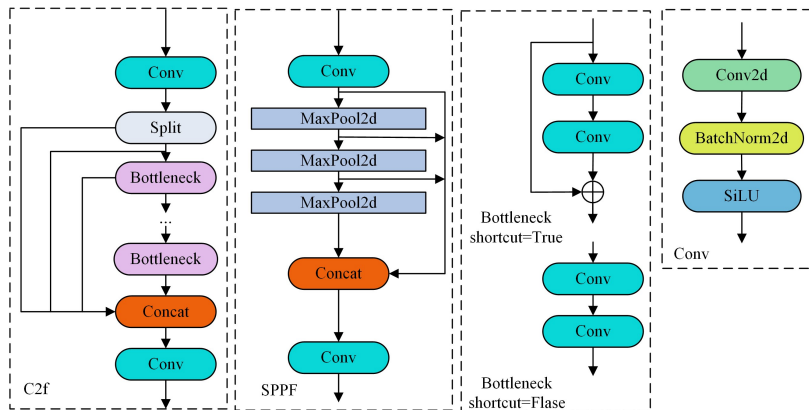


图1 YOLOv8 每个模块结构

Fig. 1 Structure of each module of YOLOv8

1.2 多目标追踪

多目标跟踪任务可以分为目标检测、目标描述、多目标匹配 3 个部分。多目标匹配是多目标跟踪的核心,需要构建相似度矩阵并优化算法来关联目标,其中匈牙利算法^[16]是最常用的数据关联算法之一,例如 SORT,DeepSort,ByteTrack 等

多目标追踪算法就是利用匈牙利算法来提升匹配精度。相较于其他跟踪算法,ByteTrack 算法的优势在于它通过关联每一个检测框来解决遮挡问题,减少了对象丢失和轨迹碎片化。

ByteTrack 的基本框架如图 2 所示。

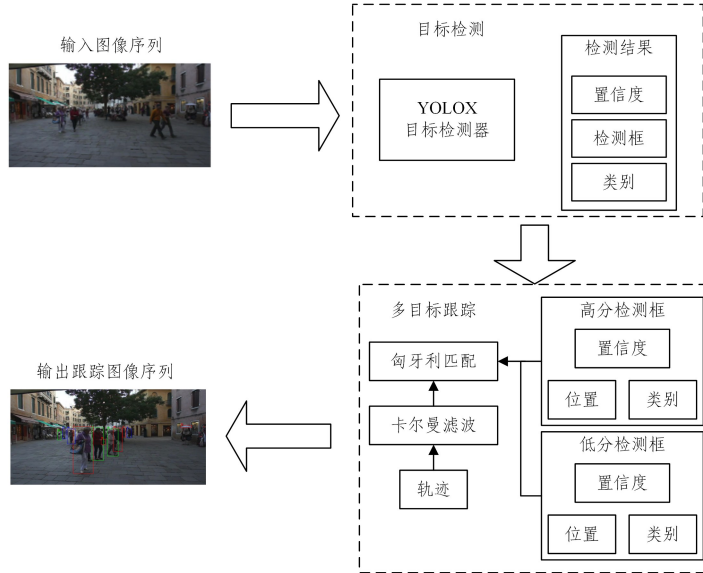


图 2 ByteTrack 基本框架

Fig. 2 Basic framework of ByteTrack

2 EAP-YOLOv8 的结构

2.1 EAP-YOLOv8 算法

针对 YOLOv8 算法在检测小目标时效果不佳的问题,本

文提出一种改进算法 EAP-YOLOv8,通过构建新型注意力机制 MAP-ECA、设计新的检测头 D-SASFF,以及构建新的损失函数 PIoUv3 来提升小目标检测的准确性。EAP-YOLOv8 的结构图如图 3 所示。

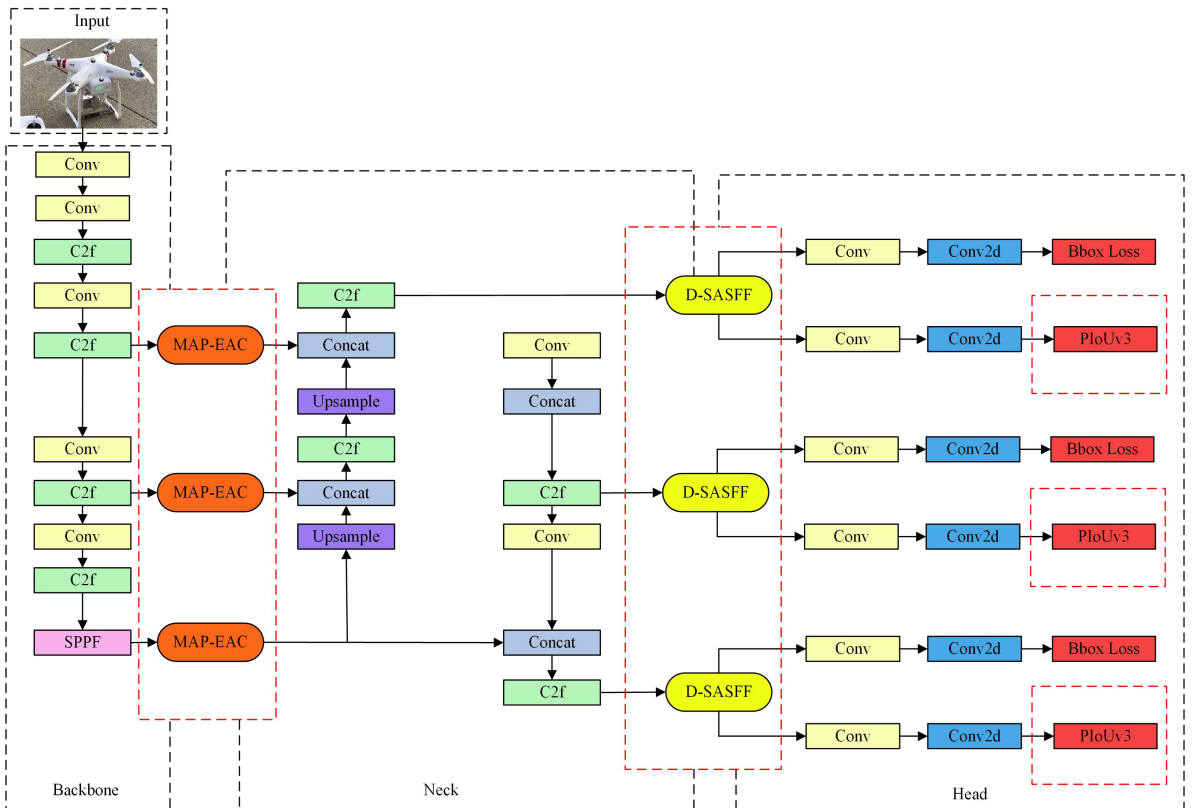


图 3 EAP-YOLOv8 的结构

Fig. 3 Structure of EAP-YOLOv8

首先,在骨干网络 C2f 和 SPPF 模块的输出端加入改进的 MAP-ECA 模块,当多次卷积之后的特征图从输出端进入 MAP-ECA 后,模块会通过自适应池化操作对重塑后的特征进行处理,强化相关性强的通道特征,弱化冗余的通道特征,再将特征图重塑回原始尺寸,恢复空间结构。经过 MAP-ECA 处理的特征图能够针对不同尺度下聚集的特征进行自适应优化,增强全局视角信息,提升小目标的检测精度。

其次,在颈部网络部分结合特征融合 ASFF 与多层次融合 SDI 模块形成新的检测头 D-SASFF。检测头对不同层级的特征图进行并行采样,通过不同尺度的卷积核来提取不同感受野下的特征,再用卷积神经网络动态地计算权重系数,将加权处理后的不同层级特征进行多层次融合,增强语义细节信息,减少误检漏检。

最后,在头部网络部分设计新型损失函数 PIoUv3 替换原有的损失函数 CIoU,通过引入高阶函数与自适应缩放因子来平衡高宽比约束在损失函数中的权重。模型在每次迭代调整预测框时会优先调整高宽比,从而强化锚框的精准性,进一步提升小目标的检测能力。

2.2 MAP-ECA 模块

在 YOLOv8 的 Backbone 模块中,C2f 模块采用特征融合的方式,将来自不同层级的特征图按照通道的维度进行拼接操作,保留空间语义信息来提高目标检测能力;SPPF 模块用于卷积层和 Bottleneck 层之后,对特征图进行不同尺度的池化操作来捕捉上下文的信息,提高检测精度。一方面,经过 C2f 中两个卷积层的处理操作后,增加了模型复杂度;另一方面,经过不同尺度的池化操作后,容易丢失局部特征信息。为了解决这一问题,考虑引入通道注意力机制^[17],通道注意力机制可以增强全局视角信息和不同尺度大小的特征,利用一维卷积提升计算效率。为解决在对无人机关键点检测时出现检测精度低的问题,进一步增强全局视角信息,降低模型复杂度,本文在 C2f 模块和 SPPF 模块后引入一个新的注意力机制 MAP-ECA,其结构如图 4 所示。

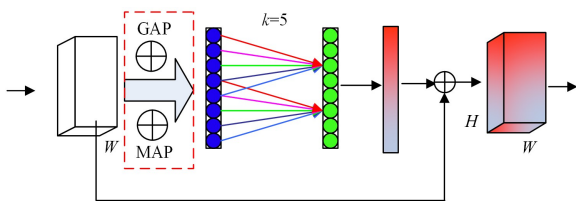


图 4 MAP-ECA 网络结构

Fig. 4 Structure of MAP-ECA network

虽然 ECA 模块通过高效的 channel attention 机制和自动调节的 kernel size 机制有效地捕捉了长距离的依赖关系,可以对不同的通道进行特征加权,从而提升识别精度,但其在处理全局上下文依赖性和通道空间方面存在一定的限制,容易造成计算结果的不稳定性。改进后的 MAP-ECA 是将高效多尺度注意力 EMA 的全局池化操作 MAP 加入到高效通道注意力模块 ECA^[18]的输入层上形成的新型注意力机制。通

过对输入层进行自适应平均池化(AdaptiveAvgPool2d)和自适应最大池化(AdaptiveMaxPool2d)的相加操作,局部连接和一维卷积产生通道注意,完成通道权重计算。该模块增强了全局视角信息,利用一维卷积提升了计算效率,保证了计算的稳定性;在保持高效轻量级的同时,提升了多尺度提取能力。MAP-ECA 首先在输入端进行自适应最大以及平均池化操作,一个分支进行一维全局池化,另一个通过 3×3 的卷积进行特征提取,之后通过 sigmoid 函数进行归一化调制,实现跨通道信息的交互合并;然后通过一维卷积快速完成通道权重的计算,保留了原有的通道一对一权重更新。卷积核 k 代表了在一个通道权重的计算过程中参与的近邻数目,表达式为:

$$k = \varphi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{\alpha}{\gamma} \right\rfloor_{\text{odd}} \quad (1)$$

当 $\gamma=2, b=1$ 时,通过映射 φ ,高维通道具有更长的相互作用,低维通道具有更短的相互作用。得到卷积核 k 后,MAP-ECA 将一维卷积应用于输入特征上,表示为:

$$[out = Conv1D_k(in)] \quad (2)$$

其中, out 表示输入特征 in 经过一维卷积操作后得到的输出特征, $Conv1D_k$ 表示核大小为 k 的卷积操作。本文在输入端加入的自适应平均池化运算与平均池化运算类似,区别在于自适应平均池化运算中任何输入大小、输出大小均为指定的 $H \times W$ 大小,但是输入和输出特征的数目不会变化,可以适应不同的输入尺寸,增强了全局视角信息和不同尺度大小的特征,提高了模型的性能。为了证明通道注意力机制 MAP-ECA 在算法中的优越性,通过在 YOLOv8 模型中加入不同注意力机制 MSFE 和 MSAM 进行对比实验,具体结果如表 1 所列。

表 1 不同注意力机制的对比实验

Table 1 Comparative experiments of different attention mechanisms

Model	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	Para	GFLOPs
YOLOv8	92.6	47.4	3.0×10^6	8.1
YOLOv8+MSFE	94.9	49.2	3.4×10^6	8.4
YOLOv8+MSAM	95.7	49.1	3.0×10^6	8.2
YOLOv8+MAP-ECA	95.8	49.4	3.0×10^6	8.1

表 1 中的结果表明:本文添加的 MAP-ECA 通道注意力机制在检测精度 $mAP@0.5$ 和 $mAP@0.5:0.95$ 方面均高于其他注意力机制,分别达到 95.8% 和 49.4%;与 MSFE 注意力机制相比,所提方法在检测精度上分别提升了 0.9 个百分点和 0.2 个百分点,在参数量和计算复杂度上分别减少了 0.4×10^6 与 0.3 GFLOPs,表明 MAP-ECA 在精度和资源消耗方面均优于 MSFE;所提方法与 MSAM 在参数量相同的情况下,检测精度和计算复杂度均更优。由此可知,在面对通道数较多的抽象信息时,通道注意力机制 MAP-ECA 的加入提升了 YOLOv8 对无人机关键点检测的准确性。

2.3 D-SASFF 检测头

YOLOv8 在特征融合部分通过在 FPN 后面增加一个自

底向上的金字塔,将浅层和深层特征图进行聚合,进一步增强了多尺度特征的表达能力。但是在进行特征融合的过程中,不同尺度的特征图在进行上下采样的过程中会导致部分关键信息的丢失,影响最终的检测精度。特别是在进行小目标检测时,小目标特征信息较弱,会因大目标或背景的遮挡而检测

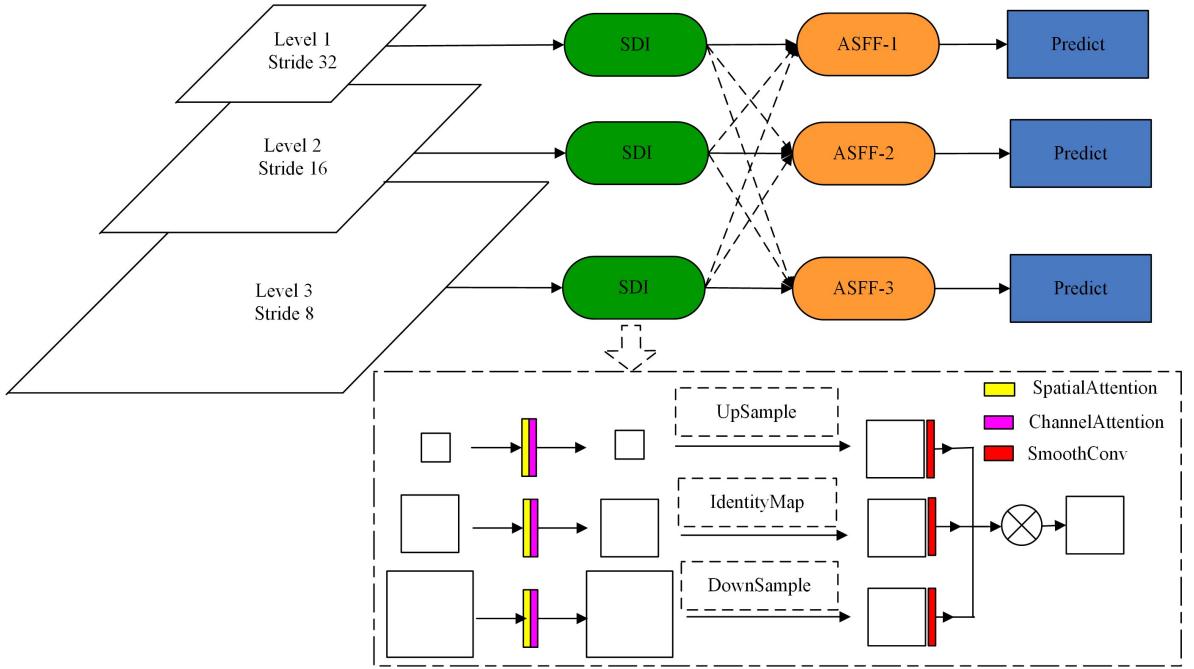


图5 SASFF网络结构

Fig. 5 Structure of SASFF network

ASFF的核心思想是引入一个自适应权重网络动态调整不同尺度特征的融合^[20],允许模型根据每个特定位置和尺度的上下文来灵活地选择最有用的特征,以提高检测精度。但是在复杂场景中,小目标的特征信息在经过多层卷积后会变得模糊不清,ASFF模块难以充分捕捉,故在此基础上加入多层次融合模块SDI,通过高层特征和低级特征的融合进行多层次检测,提升小目标特征信息的捕捉能力。D-SASFF模块共分为两步:第一步整合编码器生成的层级特征图,来增强图像中的语义信息和细节信息;第二步通过自适应空间特征融合对每一张图片进行卷积、池化等处理,从而进行加权融合。

在D-SASFF模块的第一步中,首先进行上采样操作,通过增加低层级特征图的尺寸与当前层级的尺寸进行匹配,将更细节的信息带入当前层级;然后通过身份映射到下采样,通过减小高层级特征图的尺寸与当前层级的尺寸匹配,将高层级的语义信息带入当前层级;最后通过哈达玛积(Hadamard Product)将高层级语义信息的特征和低层级细节信息的特征进行融合,增强了每个层级特征的语义和细节。

在D-SASFF模块的第二步中,先对第一步中整合过后的特征图进行恒等缩放,采用 1×1 卷积层对每一层通道进行降维压缩,然后使用最近邻插值来提高分辨率。将分辨率比为 $1:2$ 的特征图层使用卷积核为 3×3 、步长为 2 的卷积进行降采样;将分辨率比为 $1:4$ 的特征图层先进行最大池化(Max-

pooling)操作,再通过 3×3 、步长为 2 的卷积进行降采样。在自适应融合部分先对前三个特征图处的向量进行加权融合,得到融合后空间位置 (i, j) ,系数是由网络自适应学习到的,在所有通道间是共享的。以ASFF-1为例,融合后的ASFF-1的输入来自于权重 α, β, γ 与level1, level2和level3特征相乘后再相加的结果,计算式为:

SASFF结构如图5所示。

参数 α, β, γ 经过拼接之后通过softmax函数进行处理,确保它们的输出值在 $[0, 1]$ 内,且输出值总和为 1 。

$$y_{ij}^l = \alpha_{ij}^l \cdot \chi_{ij}^{1-l} + \beta_{ij}^l \cdot \chi_{ij}^{2-l} + \gamma_{ij}^l \cdot \chi_{ij}^{3-l} \quad (3)$$

参数 α, β, γ 经过拼接之后通过softmax函数进行处理,确保它们的输出值在 $[0, 1]$ 内,且输出值总和为 1 。

$$\alpha_{ij}^l = \frac{e^{\lambda_{\alpha_{ij}}^l}}{e^{\lambda_{\alpha_{ij}}^l} + e^{\lambda_{\beta_{ij}}^l} + e^{\lambda_{\gamma_{ij}}^l}} \quad (4)$$

YOLOv8引入SASFF后进一步融合Marker点的特征,可以更加关注无人机关键点的特征信息,提高检测能力。为了进一步验证新检测头D-SASFF在算法中的有效性,在模型中加入检测头Dyhead和RepHead进行对比实验,具体结果如表2所列。

表2 不同检测头的对比实验

Table 2 Comparative tests of different detection heads

Model	mAP@		Para	GFLOPs
	0.5/%	0.5; 0.95/%		
YOLOv8	92.6	47.4	3.0×10^6	8.1
YOLOv8+Dyhead	95.7	47.9	5.1×10^6	15
YOLOv8+RepHead	95.2	48.7	4.0×10^6	8.4
YOLOv8+SASFF	95.4	49.1	4.4×10^6	10.3

由表2可知,加入不同检测头对YOLOv8模型的性能产

生了显著影响。在 $mAP@0.5$ 方面,SASFF 与 Dyhead 相差无几,但在 $mAP@0.5;0.95$ 上 SASFF 较 Dyhead 提升了 1.2 个百分点,SASFF 与 RepHead 相比, $mAP@0.5$ 与 $mAP@0.5;0.95$ 都分别提升了 0.2 个百分点和 0.4 个百分点;在参数量和计算复杂度方面,检测头 DyHead 拥有最大的参数量和最高的计算复杂度,SASFF 居中,RepHead 优势最大。综合来看,SASFF 在保持较高精度的同时实现了较低的计算负载,更适合用于无人机关键点的检测。

2.4 PIoUv3 损失函数

YOLOv8 在目标检测模块采用交叉熵损失和 IOU 损失结合的设计,受到现有损失函数的影响,锚框在回归过程中膨胀导致收敛速度减小^[21]。为此,基于损失函数 PIoU 设计了一种强有力的损失函数 PIoUv3 来代替传统的 CIoU 损失,加快收敛速度。

PIoU 损失函数通过结合目标尺寸自适应惩罚因子和基于锚框质量的梯度调节函数来引导锚框沿着有效路径回归。通过惩罚因子 P 来解决锚框的膨胀问题,计算式如下:

$$P = \left(\frac{dw_1}{w_{gt}} + \frac{dw_2}{w_{gt}} + \frac{dh_1}{h_{gt}} + \frac{dh_2}{h_{gt}} \right) / 4 \quad (5)$$

其中, dw_1, dw_2, dh_1, dh_2 是预测框和目标框的对应边缘之间距离的绝对值, w_{gt} 是目标框的宽度, h_{gt} 是目标框的高度。

但是在摄像机捕捉无人机关键点时,由于无人机处在高速运动的状态下,飞行背景复杂多样且机身的关键点易被遮挡,原有的 PIoU 损失函数难以将模型优化到理想状态。而在 WIoU 损失函数中,通过构造一个非单调聚焦系数 β 就可以解决这一问题,所以本文基于 WIoU 中的动态非单调聚焦机制设计了一个超参数控制的非单调注意力函数 v3,通过与 PIoU 相结合得到一个新的损失函数 PIoUv3。PIoUv3 是在 PIoU 上增加注意力层得到的,它增强了对高质量锚框的聚焦能力,提高了检测能力。注意函数和 PIoUv3 的计算式为:

$$q = e^{-P}, q \in (0, 1] \quad (6)$$

$$u(x) = 3x \cdot e^{-x^2} \quad (7)$$

$$L_{PIoUv3} = u(\theta q) \cdot L_{PIoU} = 3 \cdot (\theta q) \cdot e^{-(\theta q)^2} \cdot L_{PIoU} \quad (8)$$

其中, $u(\theta q)$ 表示注意函数, θ 是用来控制注意力函数的超参数。用 q 替换惩罚因素 P ,当 $q=1$ 时, $P=0$,表示锚框和目标框之间完全对齐。通过关注每个锚框的回归过程中的中等质量阶段,PIoUv3 更加关注中等质量的锚框,简化了调整过程。为了表明损失函数 PIoUv3 在本文算法中的优越性,加入 3 种不同的损失函数 DIoU, WIoU, PIoUv3 来替换 YOLOv8 模型中的损失函数进行对比。对比结果如表 3 所列。

表 3 不同损失函数的对比实验结果

Table 3 Comparative experiment results of different loss functions

Model	$mAP@0.5/\%$	$mAP@0.5;0.95/\%$	Para	GFLOPs
YOLOv8	92.6	47.4	3.0×10^6	8.1
YOLOv8+DIoU	94.6	48.4	3.0×10^6	8.1
YOLOv8+WIoU	94.6	48.5	3.0×10^6	8.1
YOLOv8+PIoUv3	95.6	49.1	3.0×10^6	8.1

在加入损失函数 PIoUv3 之后,在参数量和计算复杂度

不变的情况下,PIoUv3 较其他两个损失函数 DIoU^[22] 和 WIoU,模型的 $mAP@0.5$ 和 $mAP@0.5;0.95$ 均为最优,检测精度较原有的模型分别增加了 3 个百分点和 1.7 个百分点。由此可见,损失函数 PIoUv3 可以在提升模型精度的基础上加快收敛速度,增强检测能力。

2.5 检测与追踪

本文采用 ByteTrack 追踪算法,并与改进的 EAP-YOLOv8 算法相结合,形成一种新的追踪方法。ByteTrack 是一个高效的追踪算法,与其他非 ReID 的算法一样,通过使用目标追踪获得的 bbox 进行追踪,利用卡尔曼滤波预测边界框^[23]与匈牙利算法结合进行目标和轨迹的匹配。在对低分框的使用上,ByteTrack 利用高低分框之间的区别和匹配有效地优化了追踪过程中的遮挡问题。EAP-YOLOv8 与 ByteTrack 追踪算法融合流程如图 6 所示。

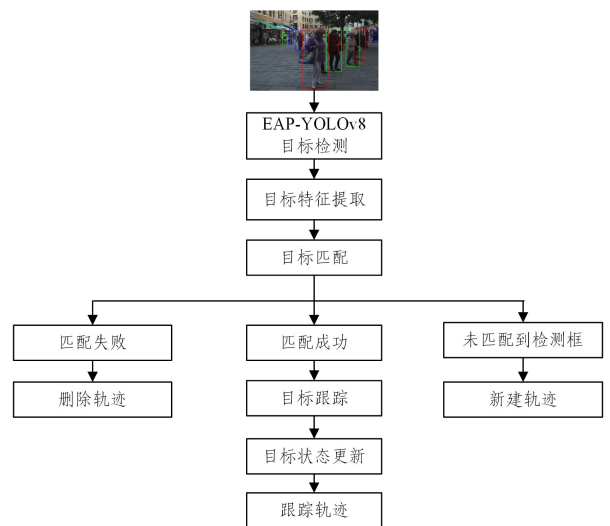


图 6 EAP-YOLOv8 与 ByteTrack 追踪算法融合流程

Fig. 6 Fusion process of EAP-YOLOv8 and ByteTrack tracking algorithm

3 实验结果与分析

本文以小型旋翼无人机为研究对象,通过对无人机机身各部分标记关键点来实现目标检测,将改进的算法应用于检测中,并与原始算法进行比较。同时,将改进的算法结合 Byte-Track 跟踪算法来实现对目标物体的检测与追踪。

3.1 数据集

目前对无人机动作捕捉的研究较少,无人机关键点的数据集也较少。因此选取 Kaggle 的公开数据集 Drone-Dataset 中的 500 张无人机照片,通过标定无人机关键点来完成自制数据集的构建。考虑到数据样本背景单一化,采用模糊、亮度、裁剪、旋转、平移、镜像等数据增强的方式将数据集扩充至 4384 张,其中训练集 3332 张,验证集 1052 张。利用 Labelimg 软件对无人机关键点进行标注,再转换成适合于 YOLOv8 的 txt 形式。在追踪数据集方面,由于无人机关键点视频数据集较少,因此采用公共数据集 MOT16^[24] 进行追踪实验, MOT16 数据集训练集和测试集均包含 7 个视频序列。

3.2 实验环境及参数设置

本实验的硬件环境为 GPU RTX 3090(24GB) * 1, CPU 14 v CPU Intel^(R) Xeon^(R) Gold 6330 CPU @ 2.00 GHz, 内存 80GB。软件环境为 PyTorch 1.10.0, Python 3.8 (ubuntu 20.04), Cuda 11.3。

本实验的参数设置如下: 图像输入尺寸 $imgsz$ 为 640×640 , 训练轮数 Epoch 设置为 100 轮, 批次大小 batch 设置为 16。

3.3 模型评价指标

为了测试本文提出的 EAP-YOLOv8 改进模型的目标检测性能, 采用基于深度学习的验证指标精确率(P)、召回率(R)、平均精度(mAP)、参数量(Parameters)以及每秒浮点运算次数(FLOPS)来评估模型。计算式分别如下所示:

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$AP = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_i \quad (12)$$

其中, P 是评估模型预测正确的正样本的比例, R 是评估模型能够找出所有真实正样本的比例, TP 是预测为正例且真实值为正例的个数, FP 是预测为正例但真实值为负例的个数, FN 是预测为负例但真实值为正例的个数; AP 表示每个类别 PR 曲线与坐标轴围成的面积, $P(R)$ 表示召回率 R 处的准确率; mAP 是指在所有类别上计算 AP 的平均值, 分为 mAP_{50} 和 mAP_{50-90} , mAP_{50} 表示在 50% 的 IoU 阈值下的 mAP 值, mAP_{50-90} 表示在 50%~95% 的 IoU 阈值范围内的 mAP 平均值。

MOTA, MOTP, HOTA 分别是多目标跟踪精度和多目标跟踪准确度以及更高维度跟踪精度, 这 3 个指标都表示多目标跟踪器的性能, HOTA^[25] 引入了更高维跟踪精度指标, 能够更加全面、均衡地衡量多目标跟踪器的性能。IDF1 是指正确的目标检测数与真实数和计算检测数的平均数的比率。计算式如下:

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \in (-\infty, 1] \quad (13)$$

其中, m_t, fp_t, mme_t 分别代表 FN, FP 以及跟踪过程中身份交换的数量 $IDSW$, g_t 表示视频序列中真实目标的数量。

$$MOTP = \frac{\sum_t d_{t,i}}{\sum_t c_t} \quad (14)$$

其中, c_t 是在第 t 帧时匹配的总数量, $d_{t,i}$ 表示假边界框与真边界框之间的距离值。

$$HOTA = \int_0^1 HOTA_\alpha \approx \frac{1}{19} \sum_{\alpha \in \{0.05, \dots, 0.95\}} HOTA_\alpha \quad (15)$$

$$IDF1 = \frac{IDTP}{IDTP + 0.5IDFP + 0.5IDFN} \quad (16)$$

其中, α 可以是 IoU 也可以是距离; $IDTP$ 表示视频中检测目标被正确分配的数量, $IDFN$ 表示视频中漏检分配的数量, $IDFP$ 表示视频中检测目标被错误分配的数量。

3.4 消融实验

为了验证本文提出的各种改进方法对 YOLOv8 性能的影响, 在自制的数据集上进行了消融实验。提出的 3 种改进策略包括在主干网络中加入跨通道交互的注意力机制 MAP-ECA、在原有检测头上加入多层次自适应空间特征融合的检测头 D-SASFF, 以及加入优化损失函数 PIoUv3。在 YOLOv8n 模型的基础上进行消融实验的结果如表 4 所列。

表 4 改进前后 YOLOv8 算法消融实验结果

Table 4 Ablation experiment results of YOLOv8 algorithm before and after improvement

算法	MAP-ECA	SASFF	PIoUv3	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	Para	GFLOPs	FPS
YOLOv8				96.3	86.7	92.6	47.4	3.0×10^6	8.1	588
YOLOv8_A	✓			97.7	92.0	95.8	49.4	3.0×10^6	8.1	667
YOLOv8_B		✓		97.4	92.0	95.4	49.1	4.4×10^6	10.3	526
YOLOv8_C			✓	97.2	91.1	95.6	49.1	3.0×10^6	8.1	588
YOLOv8_D	✓	✓	✓	98.2	94.0	96.5	50.2	4.4×10^6	10.3	526
YOLOv8_E	✓	✓		97.1	92.5	96.1	50.1	4.4×10^6	10.3	556
YOLOv8_F	✓		✓	97.5	92.3	94.7	49.4	4.4×10^6	10.3	530
YOLOv8_G		✓	✓	96.7	92.0	95.3	49.7	4.4×10^6	10.3	530

通过实验结果可以看出, 在引入 MAP-ECA, SASFF 和 PIoUv3 之后, 模型的多个指标都有所上升。相较于基础模型, 模型 A 通过加入 MAP-ECA 模块, 突出了低层细节的特征权重, 保证了 Marker 点特征的完整传递, 使得精度提升了 1.4 个百分点, 召回率提升了 5.3 个百分点, 平均精度 $mAP@0.5$ 和 $mAP_{0.5:0.95}$ 分别提升了 3.2 个百分点和 2 个百分点; 模型 B 通过采用新型检测头 D-SASFF, 实时优化锚框位置、大小和形状, 降低了漏检率和误检率, 精确率、召回率和平均精度分别提升了 1.1 个百分点、5.3 个百分点、2.8 个百分点、1.7 个百分点; 模型 C 通过采用损失函数 PIoUv3 替换原

有的损失函数 CIoU, 加快了模型的收敛速度, 平均精度 $mAP@0.5$ 提升了 3 个百分点; 模型 E, F, G 采用单一变量原则进行消融, 检测精度均低于模型 D, 模型 D 同时结合了 MAP-ECA, SASFF, PIoUv3 这 3 个模块, 模型性能到达最佳, 精确率、召回率、 $mAP@0.5$ 和 $mAP_{0.5:0.95}$ 分别达到 98.2%, 94%, 96.5% 和 50.2%, 说明了 3 个模块之间的协同作用能够在模型处理 Marker 点的识别检测任务时精准地完成目标定位, 精确地提取目标特征, 提高了检测的精确性和稳定性。本实验证明了该算法可以适用于无人机 Marker 点的识别检测任务。为了更直观地展示 EAP-YOLOv8 算法的有效性, 对训练结果进行

了指标曲线对比,以证明每个模块的必要性,结果如图 7 所示,

其中(a)和(b)为边界框损失,(c)和(d)为平均精确率。

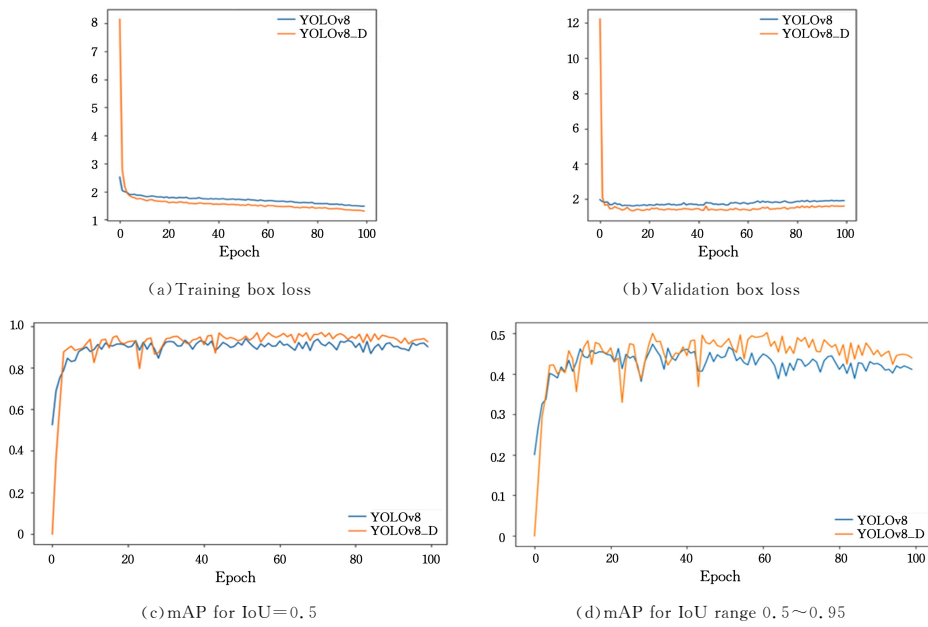


图 7 训练结果指标曲线对比

Fig. 7 Comparison of training result metric curves

3.5 综合对比实验

3.5.1 目标检测算法对比实验

为了客观地验证 EAP-YOLOv8 算法在无人机关键点检测^[26]上的有效性,将其与当前新型算法 PE-YOLO^[27], GOLO-YOLO^[28]等进行对比。PE-YOLO 算法中的 PENet 通过拉普拉斯金字塔将图像分解成多个分辨率的组件,增强了图像细节和低频信息,能够在复杂背景下进行准确识别;GOLO-YOLO 通过一种创新的聚合-分发 (Gather-and-Distribute, GD) 机制来提高信息融合效率,可以更有效地融合多尺度特征,实现低延迟和高准确性之间的理想平衡。所有实验都在相同数据集、相同环境下进行,具体结果如表 5 所列。

表 5 不同模型的对比实验结果

Table 5 Comparative experiment results of different models

Model	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	Para	GFLOPs
Yolov5	87.6	40.8	1.8×10^6	4.1
Yolov7	91.3	44	3.65×10^7	103.2
Yolov8	92.6	47.4	3.0×10^6	8.1
Yolov9	95.4	48.5	60.4×10^6	263.9
PE-YOLO	90.8	46.6	3.1×10^6	30.5
GOLD-YOLO	93.6	47.3	8.1×10^6	17.6
EAP-YOLOv8	96.5	50.2	4.38×10^7	10.3

由表 5 可知,在平均精确度 $mAP@0.5$ 和 $mAP@0.5:0.95$ 方面,本文模型 EAP-YOLOv8 较其他几个模型都有所提升,分别达到 96.5% 和 50.2%,在所有模型中是最高的。与 YOLO 系列基本模型相比,本文算法增强了全局视角信息,融合了多层次的目标特征,强化了锚框的精准性,在参数量和计算量合理控制的前提下,算法平均精度高于其他系列模型,综合性能优势明显;与新型算法相比,在相同数据集下,

EAP-YOLOv8 通过弱化无用信息,融合特征图的细节信息,减少了误检和漏检的情况,在平均精度和计算量方面表现出色。综上,EAP-YOLOv8 相较于其他的模型具有更高的检测精度和计算效率。

3.5.2 多目标跟踪算法对比实验

为验证 EAP-YOLOv8 在融合 ByteTrack 跟踪模型后对算法性能的影响,在数据集 MOT16 上进行了对比实验。通过在 EAP-YOLOv8 的输出端加入两个不同的跟踪算法模型 ByteTrack 和 Deepsort^[29]来进行实验,具体结果如表 6 所列。

表 6 多目标跟踪消融实验结果

Table 6 Ablation experiment results of multi-target tracking

Model	HOTA	MOTA	MOTP	IDF1
YOLOv8+Deepsort	32.16	25.92	79.86	35.97
EAP-YOLOv8+Deepsort	32.68	28.80	80.06	36.24
YOLOv8+ByteTrack	34.82	26.23	80.56	40.08
EAP-YOLOv8+ByteTrack	37.60	25.64	80.76	43.68

从表 6 中可以看出,采用改进算法 EAP-YOLOv8 和 ByteTrack 结合的模型在高纬度跟踪精度 HOTA 和多目标跟踪准确度 MOTP 上均有所提升。与原算法结合 ByteTrack 相比,HOTA 从 34.82% 提升到 37.60%,MOTP 从 80.56% 提升到 80.76%,同样 IDF1% 从 40.08% 提升到 43.68%。通过对目标识别检测算法的改进,Marker 点的跟踪轨迹不易丢失或出现偏差,目标特征也得到了加强。同时与 Deepsort 跟踪算法相比,ByteTrack 能在跨帧匹配目标时快速锁定特征,确保了跟踪的连贯性与准确性。在实验数据方面,ByteTrack 算法在跟踪性能上都优于 Deepsort。为了更直观地显示本文跟踪算法的优越性,对对比结果进行统计分析,训练结果如图 8 所示。

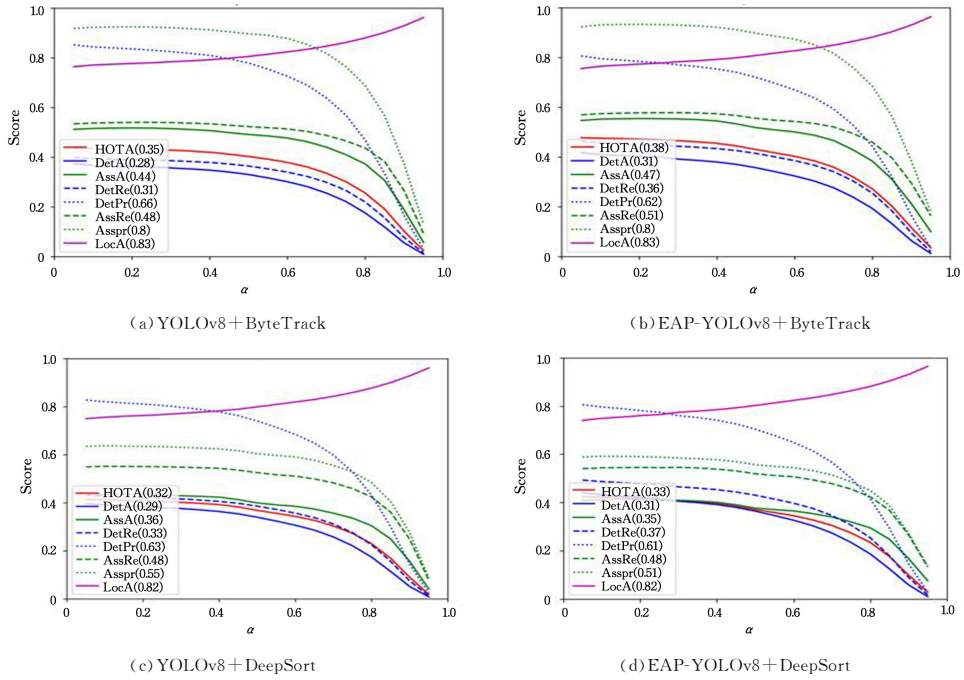


图 8 多目标跟踪算法消融实验结果

Fig. 8 Ablation experiments results of multi-target tracking algorithms

3.5.3 不同检测算法多目标跟踪算法对比实验

关键点检测方法的性能在一定程度上会影响跟踪检测的效果。如表 7 所列,4 种不同检测算法在追踪性能上均不同。改进后的 EAP-YOLOv8 检测算法较其他检测方法在 HOTA 和 IDF1 方面的表现较为明显,分别达到了 37.60% 和 43.68%,意味着该算法在检测和追踪的准确性方面表现较为出色。

表 7 不同检测算法的多目标跟踪对比结果

Table 7 Comparison of results of multi-target tracking with different detection algorithms

	(%)			
Model	HOTA	MOTA	MOTP	IDF1
YOLOv5+ ByteTrack	35.81	23.32	80.86	40.99
YOLOv8+ ByteTrack	34.82	26.23	80.56	40.09
BiP2+ ByteTrack ^[30]	—	24.39	75.96	—
CEM ^[31]	—	25.20	75.80	—
EAP-YOLOv8+ByteTrack	37.60	25.64	80.76	43.68

3.6 实验结果可视化

为了更清晰地展示检测效果,在目标检测结果的基础上进行可视化,多方位展示所提算法的有效性。由图 9 可知,与原模型相比,EAP-YOLOv8 在精度和准确性上都有所提升,有效减少了漏检与误检。

在跟踪算法方面,将 EAP-YOLOv8 与 ByteTrack 融合来检测算法性能。图 10 给出了 MOT16 的可视化结果,MOT16 共有 7 个序列,与原模型相比,新算法在视频中的每一帧都有所提升,选取其中 4 帧图片进行分析。由图 10 可知,EAP-YOLOv8 可以检测出更多小目标,大大减少了漏检的情况,提高了追踪效果。从检测和追踪可视化效果来看,EAP-YOLOv8 的性能优于现有算法。

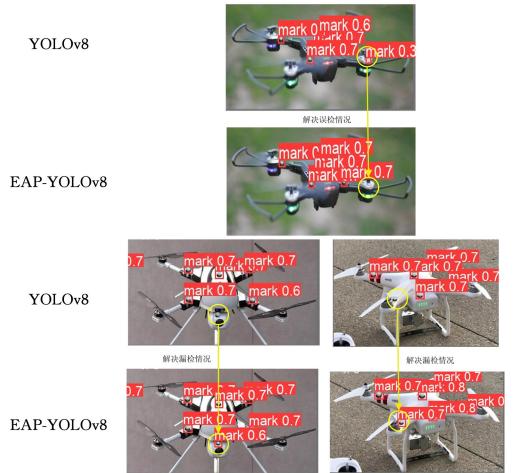


图 9 目标检测结果可视化

Fig. 9 Visualization of object detection results



图 10 目标跟踪结果可视化

Fig. 10 Visualization of target tracking results

结束语 本文提出了一种改进的 YOLOv8 无人机关键点检测算法,旨在解决无人机在空中时关键点漏检以及精度不高的问题。本文采用改进的注意力机制 MAP-ECA 和引入多层次自适应空间特征融合形成新的检测头 D-SASFF,提高了模型对小目标的检测能力和准确性;通过设计 PIoUv3 损失函数加快了收敛速度。实验结果表明,相较于其他模型,EAP-YOLOv8 在参数量和计算量提升较小的基础上,mAP@0.5 和 mAP@0.5:0.95 指标分别提高了 3.9 个百分点和 2.8 个百分点。同时在数据集 MOT16 上验证了 EAP-YOLOv8 与 ByteTrack 追踪算法结合后的有效性,通过与其他算法模型的对比,验证了 EAP-YOLOv8 算法具有很好的检测追踪效果。

本文模型能够满足光学动捕系统对无人机的实时检测任务,但是所提算法还存在两点局限:一是模型在无人机快速移动时的检测精度不高;二是模型在处理移除重叠的检测框方面存在不足。在后续的研究中,将继续优化算法模型,解决检测快速移动物体精度不高、重叠框的问题,并且尝试在无人机的视频序列数据集上进行追踪实验。

参 考 文 献

- [1] BALLARD D H. Generalizing the Hough transform to detect arbitrary shapes[J]. *Pattern Recognition*,1981,13(2):111-122.
- [2] LYU X,LIU F,REN P,et al. An image processing approach to measuring the sphericity and roundness of fracturing proppants [J]. *IEEE Access*,2019,7:16078-16087.
- [3] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014.
- [4] REN S 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*,2017,39(6):1137-1149.
- [5] LI W K,ZHANG S Q. Mask feature fusion:a new paradigm of instance segmentation[J]. *Computer Engineering*,2025(2):126-138.
- [6] XU Y W,LI J,DONG Y F,et al. YOLO series object detection algorithm review[J]. *Computer Science and Exploration*,2024,18(9):2221-2238.
- [7] LIU W,ANGUELOV D,ERHAN D,et al. SSD:Single Shot Multibox Detector[C]//European Conference on Computer Vision. Berlin:Springer,2016:21-37.
- [8] WANG G,CHEN Y,AN P,et al. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios[J]. *Sensors*,2023,23(16):7190.
- [9] LIU L,ZHANG S,BAI Y A,et al. Improving YOLOv8's Lightweight Military Aircraft Detection Algorithm[J]. *Computer Engineering and Application*,2024,60(18):114-125.
- [10] YADAV A,CHATURVEDI P K,RANI S. Object Detection and Tracking using YOLOv8 and DeepSORT [EB/OL]. https://www.publications.scrs.in/uploads/final_menuscript/6200cb6958c1909085022aae542b8792.pdf.
- [11] ZHAI X,HUANG Z,LI T,et al. YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection[J]. *Electronics*,2023,12(17):3664.
- [12] HU J F,LI B C,ZHU H,et al. Lightweight UAV target detection algorithm for improved YOLOv8[J]. *Computer Engineering and Applications*,2024,60(8):182-191.
- [13] LIU Y,LI Y,XU D,et al. Adaptive Kalman Filter with power transformation for online multi-object tracking[J]. *Multimedia Systems*,2023,29:1231-1244.
- [14] YOU L,CHEN Y,XIAO C,et al. Multi-Object Vehicle Detection and Tracking Algorithm Based on Improved YOLOv8 and ByteTrack[J]. *Electronics*,2024,13(15):3033.
- [15] YASIR M,LIU S W,XU M M,et al. YOLOv8-BYTE:Ship tracking algorithm using short-time sequence SAR images for disaster response leveraging GeoAI[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2024, 128: 103771.
- [16] ARIOKA K,SAWADA Y. Improved Kalman Filter and Matching Strategy for Multi-Object Tracking System[C]//2023 62nd Annual Conference of the Society of Instrument and Control Engineers(SICE). IEEE,2023:772-777.
- [17] XU C J,WANG X F,YANG Y D. Attention-YOLO:YOLO detection algorithm with attention mechanism[J]. *Computer Engineering and Applications*,2019,55(6):13-23,125.
- [18] WANG Q,WU B,ZHU P,et al. ECA-Net:Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11534-11542.
- [19] LIU S,HUANG D,WANG Y. Learning spatial fusion for single-shot object detection[J]. *arXiv*:1911.09516,2019.
- [20] LIU Y,CHEN J,LU P,et al. MFID-Net:multi-scaled feature-fused image dehazing via dynamic weights[J]. *Displays*,2023,78:102416.
- [21] SEKHARAMANTRY P K,MELGANI F,MALACARNE J. Deep learning-based apple detection with attention module and improved loss function in YOLO[J]. *Remote Sensing*,2023,15(6):1516.
- [22] ZHENG Z,WANG P,LIU W,et al. Distance-IoU loss : Faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020:12993-13000.
- [23] LIU Y,LI Y,XU D,et al. Adaptive Kalman Filter with power transformation for online multi-object tracking[J]. *Multimedia Systems*,2023,29(3):1231-1244.
- [24] LIANG X G,LI H,CHENG Y Z,et al. Multi-target tracking based on spatio-temporal embedding perception and multi-task collaborative optimization[J]. *Computer Engineering and Appli-*

cations, 2024, 60(6):282-292.

- [25] LUITEN J, OSEP A, DENDORFER P, et al. Hota: A higher order metric for evaluating multi-object tracking[J]. International Journal of Computer Vision, 2021, 129:548-578.
- [26] GE Q B, LI K, ZHANG X G. Relative Pose Estimation Algorithm for UAV Based on Multi-Key Point Detection Weighted Fusion[J]. Acta Automatica Sinica, 2024, 50(7):1-15.
- [27] YIN X, YU Z, FEI Z, et al. Pe-yolo: Pyramid enhancement network for dark object detection[C]// International Conference on Artificial Neural Networks. Cham: Springer, 2023:163-174.
- [28] WANG C C, HE W, NIE Y, et al. Gold-YOLO: Efficient object detector via gather-and-distribute mechanism[J]. arXiv: 2309.11331, 2023.
- [29] SHENG W, SHEN J, HUANG Q, et al. Multi-objective pedestrian tracking method based on YOLOv8 and improved DeepSORT. [J]. Mathematical Biosciences and Engineering, 2024, 21(2):1791-1805.
- [30] LIN M Y. Research on vehicle pedestrian detection and tracking algorithm based on improved YOLOv5[D]. Liuzhou: Guangxi

University of Science and Technology, 2023.

- [31] MILAN A, LEAL-TAIXÉ L, REID I, et al. MOT16: A benchmark for multi-object tracking[J]. arXiv:1603.00831, 2016.



TANG Xinliang, born in 1977, Ph.D, researcher. His main research interests include intelligent manufacturing and image processing.



WANG Jianchao, born in 1991, Ph.D, lecturer. His main research interests include intelligent information processing and machine vision.

(责任编辑:何杨)