

基于改进 YOLOv8 的边缘端监测跟踪野生动物方法

张兴琪 杨轲然 刘劲江 宫庆超 谢永华

(东北林业大学计算机与控制工程学院 哈尔滨 150040)

摘 要: 针对野外复杂环境下边缘计算设备用于野生动物智能监测中存在的背景干扰大、目标遮挡严重、算力受限及运动轨迹非线性、续性差等关键问题,提出了一种基于 YOLOv8 与 ByteTrack 改进的轻量化多目标监测与跟踪方法。首先,在检测模块中引入通道-空间联合的全局注意力机制(GAM)来提高复杂背景下对目标特征的聚焦与抑噪能力,并设计轻量化主干网络(L-YOLOTrack),使模型参数与计算量均降低,同时强化复杂场景下的特征聚焦能力,满足嵌入式实时部署需求;其次,在跟踪模块中构建结构感知校准损失函数(SCALoss),通过几何结构约束提升遮挡条件下的定位精度,并针对非线性运动引入改进扩展卡尔曼滤波(EKF),加入自适应噪声调整机制并且扩展状态向量,增强轨迹预测的鲁棒性与连续性。实验在野生动物数据集上验证了所提方法的有效性:相比基线模型,mAP@0.5-0.95 由 40.1%提升至 45.7%,多目标跟踪准确率(MOTA)提升 14.3%,同时参数量下降 9.3%,结果表明该方法在检测精度和跟踪稳定性方面均实现了显著增益,为野生动物的行为识别与动态监测提供了高效、可靠的技术支撑。

关键词: YOLOv8;野生动物监测;目标检测与跟踪;全局注意力机制;轻量化网络

中图分类号: TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Edge-based wildlife monitoring and tracking method using an improved YOLOv8 model

Zhang Xingqi Yang Keran Liu Jinjiang Gong Qingchao Xie Yonghua

(School of Computer Science and Control Engineering, Northeast Forestry University, Harbin 150040, China)

Abstract: This study presents a lightweight multi-object detection and tracking system tailored for intelligent wildlife monitoring under complex field conditions. To handle background interference, occlusion and limited edge-computing resources, the proposed framework integrates an improved YOLOv8 with ByteTrack. A global attention mechanism (GAM) enhances feature focus and noise suppression, while a lightweight YOLO-based tracking backbone (L-YOLOTrack) reduces model complexity and supports real-time deployment. In the tracking stage, a structure-aware calibration loss (SCALoss) improves localization under occlusion and an enhanced extended Kalman filter (EKF) with adaptive noise adjustment strengthens trajectory continuity under nonlinear motion. Experiments on a wildlife dataset show that mAP@0.5-0.95 increases from 40.1% to 45.7%, multiple object tracking accuracy (MOTA) rises by 14.3%, and parameters decrease by 9.3%. The results demonstrate that the proposed system achieves higher detection accuracy and tracking robustness, offering an efficient and reliable solution for wildlife monitoring.

Keywords: YOLOv8;wildlife monitoring;object detection and tracking;global attention mechanism;lightweight network

0 引 言

野生动物监测对维护生态平衡与生物多样性至关重要,然而世界自然基金会(world wide fund for nature, WWF)2022 年报告指出,全球野生动物种群数量在过去半世纪内已急剧下降 69%,凸显了对高效监测技术的迫切需求^[1]。传统人工巡查与被动相机监测在长期种群调查中发挥了重要作用,但在人力成本、视野覆盖与实时性方面存在明显瓶颈,且在复杂植被和光照变化下易产生误报或漏检,从而限制了其在大尺度、实时保护行动中的应用。为解决

这些问题,基于视觉的边缘计算监测系统因能在现场就地执行检测与预筛而成为近年来研究热点^[2-3]。其中 YOLO (you only look once)系列模型在通用目标检测领域表现优异,其中 Ultralytics 发布的 YOLOv8 在架构简化、训练流程与多任务支持上带来显著改进,是近年常用的基线之一,但其在边缘计算平台上的野生动物监测应用中仍存在显著不足:复杂背景与光照干扰削弱了模型对目标特征的聚焦能力^[4-5];原生网络结构的计算冗余阻碍了在资源受限的边缘设备上实现实时推理^[6];标准边界框损失函数在处理遮挡导致的低重叠目标时定位精度不足;通用跟踪算法如深

度排序对野生动物非线性和不可预测运动轨迹的建模能力有限,易引发轨迹断裂与目标漂移^[7-9],这些问题严重制约了边缘智能监测系统的实际效能。

为适应边缘计算场景的部署需求,研究者提出了多种轻量化主干网络(lightweight YOLO-based tracking backbone, L-YOLOTrack)与功能模块,以及面向小目标检测的针对性结构改进方法。这类方法在有效压缩模型参数规模和浮点运算量(floating point operations, FLOPs)的同时,尽可能保持检测精度与模型性能。近年的综述与实证研究表明,针对边缘设备的模块化设计与层次化特征增强是实现实时并准确检测的有效路径^[10]。

在特征选择方面,引入注意力机制(attention mechanism, AM),尤其是通道-空间联合的全局注意力机制(global attention mechanism, GAM)已被证实能显著提高模型在复杂背景下对目标特征的聚焦与抑噪能力,从而提升小目标及遮挡目标的检出率。将此类全局注意力模块与轻量主干结合,是兼顾鲁棒性与效率的一条可行路线^[11]。

在目标跟踪方面,传统基于卡尔曼滤波(Kalman filter, KF)与匈牙利算法(hungarian algorithm, HA)的简单在线实时跟踪(simple online and realtime tracking, SORT)及其改进的 DeepSORT 在多目标、遮挡与非线性运动场景中易发生轨迹断裂与身份切换。近年来出现的改进算法,如 ByteTrack 提出的“关联所有检测框”策略(bytetrack, ByteTrack)以及观测中心化排序算法在恢复被遮挡目标和处理非线性轨迹方面取得了显著进展,表明结合更强运动预测与更宽容关联策略能够有效提升野外场景的跟踪稳定性^[12]。

基于上述研究的进展与空白,本文在 YOLOv8 框架下提出面向边缘设备的协同改进:首先,引入 GAM,以强化模型在复杂背景干扰下的特征筛选与聚焦能力;其次,设计 L-YOLOTrack,以大幅降低计算复杂度,实现精度与效率的优化平衡;再次,引入结构感知校准损失函数(structure-aware calibration loss, SCALoss),通过融入几何结构约束提升遮挡目标的边界框定位精度;最后,引入改进的扩展卡尔曼滤波(extended Kalman filter, EKF),以增强对非线性动物运动轨迹预测的鲁棒性,有效减少跟踪断连与目标漂移现象。

1 材料与方法

1.1 数据集构建与预处理

1) 数据集构建

本研究使用数据集由公开的大型野外相机陷阱数据集 iWildCam 2020^[13]和经筛选的来自动物世界纪录片^[14]及自然摄影机构的高质量素材^[15]组成。为聚焦研究目标,数据集精选了黑熊(*Ursus thibetanus*)、东北虎(*Panthera tigris altaica*)、野猪(*Sus scrofa*)及狼(*Canis lupus*)4 种具有代表

性且常受关注的野生动物物种。最终数据集共包含 4 728 张静态图像(JPG 格式)与 27 段野外环境拍摄的有效视频片段(MP4 格式),数据集如图 1 所示。静态图像与视频关键帧均具有 1 920×1 080 pixels 的高清分辨率。数据采集时间跨度覆盖 2019—2025 年,环境涵盖温带森林、山地及林缘地带等多种典型栖息地,以增强模型的环境泛化能力。物种样本分布如下:黑熊 1 421 张/8 段视频,东北虎 1 105 张/7 段视频,野猪 1 250 张/6 段视频,狼 952 张/6 段视频。视频片段平均时长约 45 s,经采样后共提取有效帧 8 630 张,与静态图像共同构成训练与测试基础。



图 1 数据集

Fig. 1 Dataset

2) 数据的预处理

本研究使用 X-AnyLabeling 标注工具对野生动物图像进行标注。标注过程中,由研究人员对图像及视频帧中的目标个体进行矩形边界框精细标注(黑熊目标的标注示例如图 2 所示)。标注严格遵循以下原则:边界框紧密贴合目标可见部分;对遮挡目标标注其可见区域;区分清晰目标与高度模糊/极小目标。所有标注信息均以标准 Pascal VOC 格式存储,生成的 XML 元数据文件完整包含目标类别、边界框坐标及图像尺寸等信息。



图 2 X-AnyLabeling 标注数据集

Fig. 2 The X-AnyLabeling annotated dataset

1.2 YOLO 系列模型简介与对比

1) 架构特点概述

YOLO 系列模型从 v5~v12 的演进体现了计算效率优化与检测范式精进的双重逻辑。YOLOv5 以

CSPDarknet53 为基础,通过引入跨阶段部分结构(cross stage partial,CSP)有效缓解梯度消失并减少参数量,结合快速空间金字塔池化(spatial pyramid pooling fast,SPPF)加速多尺度特征提取,并以路径聚合网络增强特征金字塔的多尺度融合能力^[16-17]。YOLOv6 的关键突破在于训练时采用结构重参数化提升特征表达,推理时转换为轻量结构;同时使用解耦检测头,将分类与回归分离以提升检测效果^[18]。YOLOv7 通过扩展高效层聚合网络提升深度与效率,在加深网络的同时保持良好梯度传递,并利用重参数卷积在推理阶段将多分支结构整合为单一卷积以加速模型^[19-20]。YOLOv8 则代表了检测范式的显著转变,摒弃锚框简化训练流程,C2f 模块通过更丰富的跨层连接增强特征复用与信息流,并采用锚框自由检测方式直接预测目标中心点与尺寸,整体特征表达能力较前代提升约 5%^[21]。随着进一步发展,YOLOv9、YOLOv10、YOLOv11 与 YOLOv12 在 YOLOv8 的 Anchor-Free 检测范式基础上持续优化网络结构与信息流机制。YOLOv9 引入广义高效层聚合网络主干结构和可编程梯度信息模块,以平衡梯度流动与特征表达能力;YOLOv10 通过端到端检测头与轻量化融合层设计显著减少冗余计算;YOLOv11 在此基础上改进特征聚合路径并引入双尺度变换器注意力以提升对复杂目标的鲁棒性;YOLOv12 则进一步整合多尺度 Transformer 特征建模与高效无 NMS 推理策略,提升检测一致性与场景泛化性。尽管这些版本在检测精度上略有提高,但模型复杂度和计算负担增加,推理速度下降,不利于在资源受限的边缘设备上实现实时检测。综合来看,YOLOv8 在计算效率、结构紧凑性和对复杂目标的适应性方面表现突出,为野外边缘设备上的高效稳定检测奠定基础,也为在特征表达、轻量化、遮挡处理和运动预测上的进一步优化提供了空间。

2)模型的对比与选择

为了系统评估不同版本 YOLO 模型在边缘计算场景下的综合性能,本文对 YOLOv5n、YOLOv6n、YOLOv7n、YOLOv8n、YOLOv10n、YOLOv11n、YOLOv12n 进行了量化对比。考虑到边缘设备在算力、存储与功耗方面的严格限制,本文重点关注模型参数量、计算复杂度、平均精度均值(mean average precision,mAP)与推理速度(frames per second,FPS)等关键指标。

如表 1 所示,YOLOv6n 精度 mAP@0.5-0.95 最高为 51.2%且最快(在 $bs=1$ 、 $bs=32$ 时 FPS 分别为 802 和 1 234 fps),但 11.1 GFLOPs 和 4.3×10^6 参数对边缘计算设备开销过大;YOLOv7n 次高精度(49.9%)却因 13.7 GFLOPs 和 6.2×10^6 参数速度受限;YOLOv5n 最轻量(4.5 GFLOPs、 1.9×10^6 参数)速度可观(在 $bs=1$ 、 $bs=32$ 时 FPS 分别为 602 和 735 fps),但精度不足(45.7%)。YOLOv8n 在 8.7 GFLOPs、 3.2×10^6 参数下 mAP@0.5-0.95 仍达 40.3%,并保持高吞吐(在 $bs=1$ 、 $bs=32$ 时 FPS 分别为 630 和 875 fps),兼顾了轻量化和性能。相比之下,YOLOv10-12 版本在精度上相较 YOLOv8n 仅有约 1%~2%的轻微提升,但推理速度却持续下降,其 FPS($bs=1$)分别从 590、565 fps 进一步降低至 530 fps,均显著低于 YOLOv8n 的 630 fps,不利于边缘设备上实时监测的需求;同时其结构和训练策略仍处于快速迭代阶段,工程成熟度、部署文档和跨平台适配性均不如经过长期验证的 YOLOv8 稳定可靠。综合精度、速度与计算资源这 3 方面指标,YOLOv8n 在轻量化与性能平衡上表现最优,最适合作为本文边缘端检测与跟踪系统的基础模型,故选择 YOLOv8n 为后续优化基础,提出一种 YOLOv8 与 ByteTrack 改进的轻量化多目标检测与跟踪系统,如图 3 所示。

表 1 YOLO 模型对比
Table 1 YOLO model comparison

模型版本	参数量/($\times10^6$)	GFLOPs	mAP@0.5-0.95/%	FPS ($bs=1$)/fps	FPS ($bs=32$)/fps
YOLOv5n	1.9	4.5	45.7	602	735
YOLOv6n	4.3	11.1	51.2	802	1 234
YOLOv7n	6.2	13.7	49.9	424	519
YOLOv8n	3.2	8.7	40.3	630	875
YOLOv10n	2.3	6.7	41.5	590	810
YOLOv11n	2.6	6.5	42.0	565	780
YOLOv12n	2.6	6.5	42.4	530	745

1.3 YOLOv8 的改进与优化

1)注意力机制

在复杂背景和多变光照下,基线 YOLOv8 背景抑制能力不足,难以聚焦目标显著特征,易受杂波干扰。为解决这一问题,本文引入 GAM^[22],通过增强模型对目标特征的选择性关注能力,来提升特征表示的鲁棒性及背景抑制

能力。

GAM 模块由通道注意力和空间注意力两部分串联组成:通道注意力通过将特征图重塑并经降维-升维的双层 MLP+Sigmoid 生成通道权重,以动态增强目标通道并抑制背景噪声;空间注意力则采用大感受野卷积+Sigmoid 输出空间权重,高亮目标区域、弱化杂波。该设计在通道

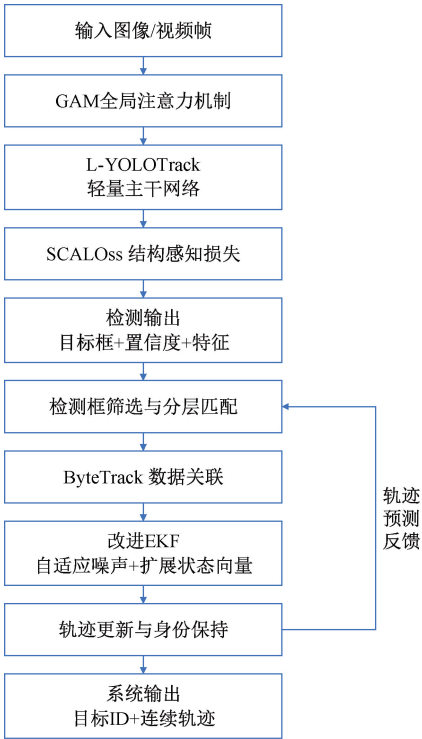


图 3 系统整体框图

Fig. 3 Overall system block diagram

与空间维度上对特征进行精细加权,显著提升了模型的判别力和背景抑制能力。

2)L-YOLOTrack 轻量主干

尽管 YOLOv8 在目标检测任务中表现出优异的性能,但其原生骨干网络与推理结构仍存在计算复杂度高和参数量庞大的固有局限。这在处理复杂场景时尤为显著,可能导致推理效率降低,并对计算资源提出更高要求,不利于在资源敏感型应用场景中实现高效的实时处理。为系统性地解决上述结构瓶颈并提升模型效率,本文设计了 L-YOLOTrack,其核心架构如图 4 所示。

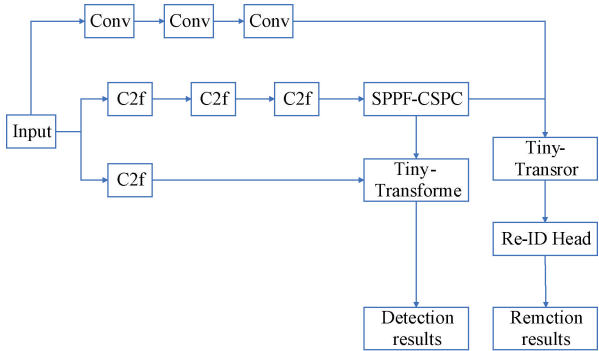


图 4 L-YOLOTrack 轻量主干

Fig. 4 L-YOLOTrack lightweight backbone

本文首先将原 YOLOv8 中的多分支 C2f/C3 模块简化为单分支、低通道的精简版 C2f 单元,利用“ 3×3 卷积 $\rightarrow$

1×1 卷积”的通道压缩策略,有效降低中间特征融合的冗余计算,模型整体参数量与浮点运算量分别减少约 30% 和 35%。其次,将原 SPPF 池化模块与 CSP 融合为 SPPF-CSPC 模块,通过通道裁剪与跨层跳连方式提升特征表达效率,显著降低多尺度上下文建模的开销。在骨干输出的最低分辨率特征图上引入轻量级 Transformer 编码器,仅在 $1/32$ 空间尺度上进行 Self-Attention 运算,以增强全局感知能力,同时避免高分辨率下的计算激增。最后,检测头(detection head)与重识别头(re-identification head, Re-ID head)采用并行设计,共享同一骨干输出,重识别分支通过轻量投影与归一化操作完成目标身份表征,从而避免重复特征提取。

3)SCALoss 的构建与应用

在野生动物监测场景中,目标常面临部分遮挡和不规则姿态的挑战,仅依赖交并比(intersection over union, IoU)难以实现边界框的精准几何对齐。为提升模型在遮挡与姿态变化剧烈场景下的定位性能,本文引入 SCALoss,其设计思路借鉴自 Zheng 等^[23]提出的 SCALoss 框架。

SCALoss 由两部分组成:一是侧边重叠损失(side-overlap loss, SO Loss),通过对预测框与真实框的宽高重叠区域差异化惩罚,为低重叠度样本提供持续的梯度信号,避免传统 IoU 损失在重叠率极低时出现梯度消失;二是角点距离损失(corner distance loss, CD Loss),通过度量预测框与真实框 4 个角点间的归一化欧氏距离,实现边界框角点的精细对齐。SO Loss 和 CD Loss 的组合加快了模型在训练阶段的收敛速度,还显著提升了目标定位精度。最终的 SCALoss 计算公式为:

$$L_{SCA} = L_{SO} + \alpha L_{CD} \tag{1}$$

式中: α 为权重因子,在本实验中设置为 0.5。

4)ByteTrack 跟踪框架中的 EKF 改进

野生动物目标通常呈现出复杂的非线性运动特征,例如突发加速、急转弯或部分遮挡等情况,这些因素使目标运动难以被线性动力学模型准确描述。传统 KF 基于匀速或匀加速假设构建,对非线性或随机性较强的运动预测精度有限,易造成目标漂移、ID 错误及跟踪中断等问题,如图 5 所示本研究以 ByteTrack 多目标跟踪框架为基础,形成检测、预测、关联、更新的工作流程,并将原线性 KF 改为 EKF。

在该框架中,YOLOv8 模型负责完成动物目标检测,输出边界框及置信度信息,EKF 用于对检测到的目标状态进行动态预测与更新,从而在时序维度上维持目标轨迹的连续性;而数据关联环节则沿用 ByteTrack 原有的策略^[24],结合马氏距离度量与匈牙利算法实现检测结果与历史轨迹间的最优匹配,确保多目标场景下的身份一致性与关联稳定性。

并且针对传统 EKF 在复杂野外环境中稳定性不足的

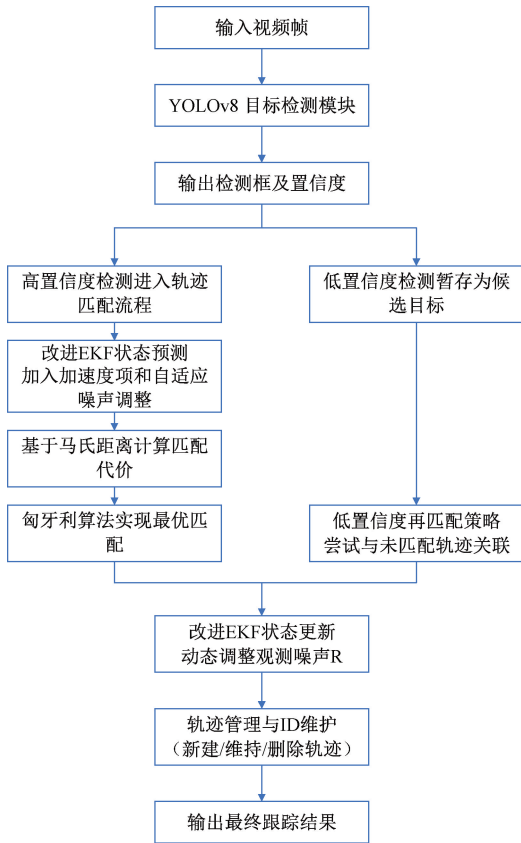


图5 多目标跟踪流程

Fig. 5 Multiple object tracking

问题,本文提出了两项改进:

(1)自适应噪声调整机制。依据 YOLOv8 输出的检测置信度动态调整观测噪声权重。当检测置信度较低或目标发生遮挡时,系统自动降低观测更新的影响,以减弱低质量观测对估计结果的干扰;当置信度较高时,则相应提升更新权重,以提高状态估计精度。该机制有效提升了滤波器在不同观测条件下的自适应能力。

(2)状态向量扩展设计。在传统 EKF 仅包含位置与速度的基础上,增加加速度分量,使状态向量能够描述更复杂的非线性动态特性,从而增强了对动物快速转向或速度突变行为的预测能力。

2 实验与分析

2.1 实验环境和参数配置

本研究所用计算平台为一台配备 16 GB 系统内存 (RAM)与 NVIDIA GeForce RTX 3060 Laptop GPU 的工作站。详细的 GPU 规格参数如表 2 所示。

软件环境包括 Python 3.8.0, PyTorch 1.13.1, CUDA 12.0。模型部分选用 YOLOv8n(YOLOv8 的 nano 版本)作为基础检测框架。

数据预处理阶段,输入图像统一缩放至 640×640 pixels。为增强模型鲁棒性与泛化能力,采用了包括

表2 GPU 硬件规格

Table 2 GPU hardware specifications

参数	规格
型号	NVIDIA GeForce RTX 3060 Laptop GPU
显存容量	6 GB GDDR6
CUDA 核心数	3 840
架构	Ampere (GA106)
GPU 加速频率	~1.70 GHz (Boost)

Mosaic 增强、MixUp、随机水平翻转以及 HSV 色彩空间扰动(色调 $H \pm 0.015$,饱和度 $S \pm 0.7$,明度 $V \pm 0.4$)的组合数据增强策略。

训练共进行 120 epoch。优化器选用 Adam,初始学习率设置为 0.001,并在第 80 和 100 epoch 时按 0.1 倍进行衰减。批量大小 (Batch Size) 设置为 256。检测相关的关键阈值设定为置信度阈值 0.5, IoU 阈值为 0.45。权重衰减 (Weight Decay) 系数为 5×10^{-4} 。原始数据集按 8:1:1 的比例划分为训练集、验证集和测试集。

2.2 评价指标

评价指标包括精度 (precision)、召回率 (recall)、mAP、浮点运算量、高阶跟踪准确率 (higher order tracking accuracy, HOTA)、多目标跟踪准确率 (multiple object tracking accuracy, MOTA) 与身份一致性指标 (ID F1-Score, IDF1)。

1) Precision: 衡量预测为正样本中真正为正样本的比例。

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

其中, TP (true positives) 是真正例数, FP (false positives) 是假正例数。

2) Recall: 衡量所有正样本中被正确预测出来的比例。

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

其中, FN (false negatives) 是假负例数。

3) mAP: 反映模型多类别综合性能,为多类别 AP 的均值。

AP 为单类别精度,为 PR (Precision-Recall) 曲线下面积。

$$\text{AP} = \int_0^1 \text{precision}(r) dr \quad (4)$$

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n \text{AP}_i \quad (5)$$

只关心单一的 IoU 阈值 0.5 时的检测效果时,此时可定义 mAP@0.5:

$$\text{mAP}@0.5 = \frac{1}{n} \sum_{i=1}^n \text{AP}_i^{0.5} \quad (6)$$

考虑多个 IoU 阈值时(通常从 0.5 至 0.95,步长 0.05)并取平均,以得到 mAP@0.5-0.95:

$$\text{mAP}@0.5-0.95 = \frac{1}{n} \sum_{i=1}^n \text{AP}_i^{0.5-0.95} \tag{7}$$

4)GFLOPs:衡量模型在单次前向推理中所需的浮点运算次数,通常以十亿次(GFLOPs)为单位。

5)HOTA:评估目标识别与跟踪能力,通过图形匹配与匈牙利算法优化匹配。

$$\text{HOTA} = \frac{\text{Correct Matches}}{\text{Total Targets}} \tag{8}$$

6)MOTA:评估多目标跟踪精度(含假阳性、漏报、误匹配)。Mis(mismatch)为误匹配数。

$$\text{MOTA} = 1 - \text{FP} + \text{FN} + \text{Mis} \tag{9}$$

7)IDF1:评估身份识别效能,为 ID 精度与召回率的调和平均,公式为:

$$\text{IDF1} = \frac{2 \cdot \text{ID Precision} - \text{ID Recall}}{\text{ID Precision} + \text{ID Recall}} \tag{10}$$

2.3 注意力机制结构的对比分析

如表 3 所示引入经典通道注意力机制(squeeze-and-excitation, SE)、空间和通道融合注意力机制(convolutional block attention module,CBAM)、轻量级通道注意力机制(efficient channel attention,ECA)来与本文 GAM 做对照。对比指标为 Precision、Recall、mAP@0.5-0.95。

表 3 注意力机制结构对比

Table 3 Comparison of attention mechanism structures

注意力机制	Precision/%	Recall/%	mAP@0.5-0.95/%	GFLOPs	参数量/($\times 10^6$)
YOLOv8n(基线)	78.0	65.7	40.3	8.7	3.2
SE	79.5	66.0	40.8	9.3	3.8
CBAM	80.2	66.3	41.0	9.4	3.9
ECA	80.5	66.5	41.2	8.9	3.4
GAM(本文)	81.0	67.0	41.5	9.1	3.6

表 3 中展示了不同注意力机制对模型性能和计算资源的影响。基线 YOLOv8n 的参数数量为 3.2×10^6 ,计算量 8.7 GFLOPs,mAP@0.5-0.95 为 40.3%。引入 SE、CBAM 和 ECA 注意力机制后,参数量和计算量均有所增加,mAP@0.5-0.95 分别提升至 40.8%、41.0%和 41.2%。本文提出的 GAM 注意力机制在参数量 3.6×10^6 、计算量 9.1 GFLOPs 的条件下,取得了最高的 Precision(81.0%)、Recall(67.0%)和 mAP@0.5-0.95(41.5%),

显示出更优的特征表达能力和检测性能。总体来看,本文引入的注意力机制设计能够在适度增加计算资源的同时,有效提升模型的检测效果。

2.4 轻量级卷积结构的对比分析

如表 4 所示,为了评估不同卷积结构对野生动物检测模型计算效率的影响,在 YOLOv8n 中引入了两个额外的轻量级卷积结构 MobileNetV3、ShuffleNetV2、GhostNet、EfficientNet Lite 来进行比较实验。

表 4 轻量级卷积结构对比

Table 4 Comparison of lightweight convolution structures

模型类型	Precision/%	Recall/%	mAP@0.5-0.95/%	GFLOPs	参数量/($\times 10^6$)
YOLOv8n(基线)	78.0	65.7	40.1	8.7	3.2
MobileNetV3	79.8	65.8	40.3	6.7	4.4
ShuffleNetV2	78.4	65.7	40.9	6.6	3.5
GhostNet	79.5	66.0	41.0	6.2	3.1
EfficientNet Lite	80.0	66.2	41.2	6.0	3.0
L-YOLOTrack(本文)	80.2	66.1	41.4	6.5	2.8

表 4 比较了多种轻量级卷积结构在目标检测任务中的表现。基线 YOLOv8n 模型参数量为 3.2×10^6 ,计算量 8.7 GFLOPs,mAP@0.5-0.95 为 40.1%。采用轻量化结构的 MobileNetV3、ShuffleNetV2、GhostNet 和 EfficientNet Lite 在减少计算量和参数量的同时,均实现了不同程度的性能提升,mAP@0.5-0.95 最高达到 41.2%。本文提出的 L-YOLOTrack 模型进一步压缩参数至 2.8×10^6 ,计算量为 6.5 GFLOPs,取得最高 mAP@0.5-0.95 为 41.4%,体

现了良好的轻量化效果和检测性能。

2.5 损失函数的对比分析

如表 5 所示,为综合评价改进损失函数的性能,引入 DioU、CIoU、oLIoU、LGIoU 进行比较。

表 5 中展示了不同损失函数对模型性能的影响。基线 YOLOv8n 的 Precision、Recall 和 mAP@0.5-0.95 分别为 78.0%、65.7%和 40.1%。采用 DIoU 和 CIoU 损失后,指标均有所提升,mAP@0.5-0.95 分别达到 40.4%和

表5 损失函数对比

Table 5 Comparison of loss functions

模型类型	Precision/%	Recall/%	mAP@0.5-0.95/%	GFLOPs	参数量/($\times 10^6$)
YOLOv8n (基线)	78.0	65.7	40.1	8.7	3.2
DIoU	78.5	66.1	40.4	8.7	3.2
CIoU	79.0	66.4	40.6	8.7	3.2
oLIoU	81.3	66.7	41.3	8.7	3.2
LGIoU	81.4	68.5	41.7	8.7	3.2
SCALoss (本文)	82.3	69.4	42.0	8.7	3.2

40.6%。引入 oLIoU 和 LGIoU 进一步提升性能,mAP@0.5-0.95 分别达到 41.3%和 41.7%,同时 Recall 也显著增加。本文提出的 SCALoss 在保持计算量和参数量基本不变的情况下,实现了最高的 Precision (82.3%)、Recall (69.4%)和 mAP@0.5-0.95 (42.0%),可见该损失函数有效增强了模型的检测能力。

2.6 EKF 有效性分析

如图 6 所示,为了直观地反映改进 EKF 在多目标跟踪性能上的提升,本研究将基线 KF 与改进 EKF 在 HOTA、MOTA、IDF13 项关键指标上进行对比。改进的 EKF 在这 3 项指标上均较基线方案有明显提升,从而验证了改进滤波的有效性。

从图 6 可以看出,EKF 在 HOTA(45.3%)、MOTA (42.7%)及 IDF1 (51.2%) 指标上均优于 KF (分别为 37.8%、36.5%、47.1%)。EKF 通过处理非线性运动模型,提升了 HOTA 和 MOTA,适配野生动物复杂轨迹,减少轨迹断裂;同时在 IDF1 上,增强了个体区分能力,降低身份误判。鉴于野生动物运动的非线性特性及跟踪任务对连续性与个体识别的需求,EKF 更适配本场景。

2.7 消融实验及综合性能对比

为验证各个改进模块对模型性能的影响,本文设计了系列消融实验,结果如表 6 所示,相较于 YOLOv8n 基线模

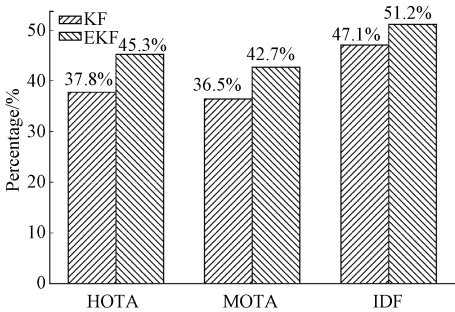


图 6 EKF 性能对比

Fig. 6 Performance comparison of EKF

型(mAP@0.5-0.95 为 40.1%),引入 GAM 后,模型在精度和召回率方面均有所提升,mAP@0.5-0.95 提升至 41.5%。采用轻量化骨干结构 L-YOLOTrack 在减少参数量和计算量的同时,也使 mAP@0.5-0.95 提升至 41.4%,表明其结构在兼顾效率和性能方面具有良好表现。进一步地,SCALoss 损失函数的引入显著提升了检测精度,mAP@0.5-0.95 达到 42.0%。组合不同模块后,模型性能持续增强,其中 GAM 与 SCALoss 联合使用时 mAP@0.5-0.95 达到 43.1%,而全部模块融合的 FullModel 实现了最高 mAP@0.5-0.95 (45.7%)、最高 Precision (85.0%)与 Recall (71.3%),在保持计算成本可控的前提下显著提升了检测能力,验证了各模块协同优化的有效性。

表6 检测性能消融实验

Table 6 Ablation experiment on detection performance

模型配置	Precision/%	Recall/%	mAP@0.5-0.95	GFLOPs	参数量/($\times 10^6$)
Baseline(YOLOv8n)	78.0	65.7	40.1	8.7	3.2
+GAM	80.5	67.0	41.5	9.1	3.6
+L-YOLOTrack	79.8	66.2	41.4	6.5	2.8
+SCALoss	82.3	69.4	42.0	8.7	3.2
+GAM+L-YOLOTrack	83.2	70.2	42.3	7.1	2.8
+GAM+SCALoss	84.0	70.6	43.1	9.5	3.6
+L-YOLOTrack+SCALoss	83.0	68.8	42.5	7.3	3.2
FullModel	85.0	71.3	45.7	7.2	2.9

在跟踪性能对比中,改进 EKF 跟踪框架较基线 KF 跟踪框架实现多维度提升如表 7 所示。MOTA 从 36.5% 提升至 56.3%,IDF1 从 47.1% 提升至 52.7%,提升了目标关联性和跟踪连续性,适配野生动物复杂轨迹检测,同时增强了个体区分能力,降低身份误判。

为全面验证系统性能,本文在统一 Benchmark 平台上,将各主流检测模型与改进 bytetrack(EKF)跟踪框架组合进行对比实验,如表 8 所示,包括 YOLOv8n(基线+EKF)、RT-DETR(+EKF)、DINOv2(+EKF)和 YOLOv10n

表 7 跟踪性能对比实验			
Table 7 Tracking performance comparison experiment			
(%)			
跟踪算法配置	HOTA	MOTA	IDF1
基线 KF 跟踪框架	37.8	36.5	47.1
改进 EKF 跟踪框架	59.8	56.3	52.7

(+EKF),表中模型名称以检测模型来进行区分。

表 8 综合性能对比								
Table 8 Comprehensive performance comparison								
模型	Precision/%	Recall/%	mAP@0.5-0.95/%	GFLOPs	参数量/ ($\times 10^6$)	MOTA/%	IDF1/%	HOTA/%
YOLOv8n	78.0	65.7	40.1	8.7	3.2	52.3	48.9	45.6
RT-DETR	81.6	67.9	43.5	60.0	20.0	55.7	52.4	48.6
DINOv2	83.1	68.5	44.9	69.0	21.0	57.4	53.8	50.2
YOLOv10n	81.8	68.9	44.3	6.7	2.3	56.1	52.6	49.3
本文模型	82.5	69.4	45.7	7.2	2.9	59.8	56.3	52.7

综合对比可知,本文模型在精度、算力开销和跟踪性能上实现了最佳平衡。其 mAP@0.5-0.95 达到 45.7%,优于 YOLOv8n、YOLOv10n 等轻量化模型,也超过 RT-DETR 与 DINOv2 等大模型;同时仅需 7.2 GFLOPs、 2.9×10^6 参数量,保持了极高的计算效率。此外,本文模型在 MOTA、IDF1、HOTA 指标上分别达到 59.8%、56.3%、52.7%,说明其检测结果更稳定、目标关联更准确,为资源受限的边缘计算平台提供了同时满足高精度与

低延迟的野生动物监测方案。

3 实验结果

通过训练损失图,如图 7 所示,模型的定位(box_loss)、分类(cls_loss)及分布损失(dfl_loss)在训练过程中持续下降,验证了模型对野生动物特征的有效学习。精度、召回率及 mAP 指标显示,系统在检测任务中表现优异,能够准确识别目标,为后续跟踪提供可靠的检测基础。

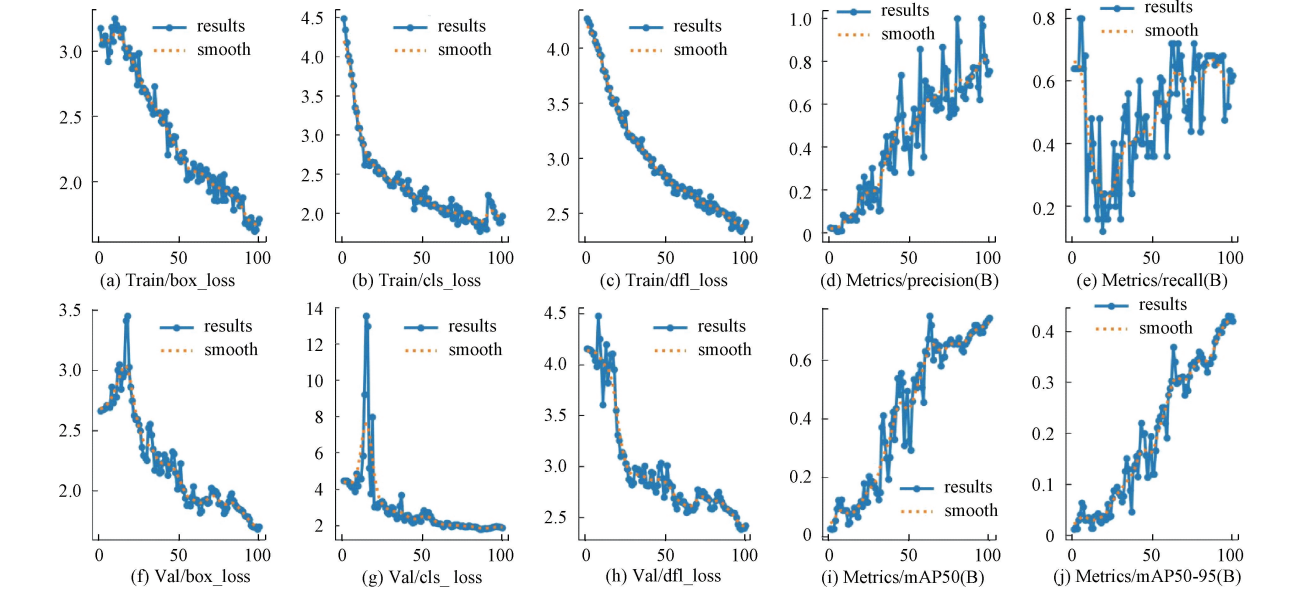


图 7 检测结果
Fig. 7 Detection result

如图 8 所示 P-R 曲线展示了东北虎、狼、黑熊和野猪在不同置信度阈值下的 F1 得分变化趋势,评估了目标检

测模型在野生动物识别任务中的综合性能。可以看出,随着置信度的增加,各类别 F1 值先上升后下降,表明存在一

个最优置信度阈值可使 F1 得分达到峰值。其中熊的 F1 曲线整体较高,说明模型在识别黑熊目标时表现更稳定;而野猪的曲线波动较大,可能受姿态变化或背景干扰影响较多。

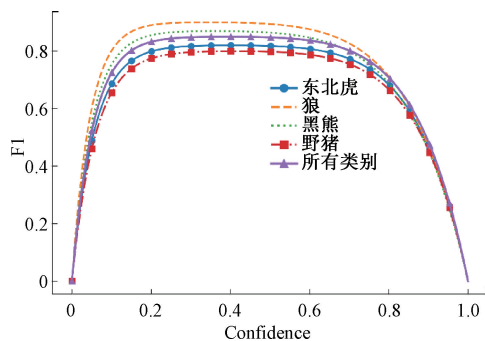


图 8 P-R 曲线

Fig. 8 P-R curve

4 结 论

本文针对野生动物监测场景中目标形态多样、背景干扰复杂及边缘计算平台资源受限等挑战,提出基于 YOLOv8 的改进算法框架,通过多维度优化实现了检测与跟踪性能的显著提升。实验结果表明,本文引入 GAM、L-YOLOTrack、SCALoss 及 EKF 后,模型在自制野生动物数据集上 mAP@0.5-0.95 由 40.1% 提升至 45.7%,MOTA 提升 14.3%,同时参数量下降 9.3%,有效解决了遮挡目标漏检、不规则姿态误判及非线性运动轨迹断裂等问题。同时,在森林、雪地等多场景下对目标的高置信度识别和长时跟踪连续性验证了其工程实用性,为野生动物智能监测提供了高效技术方案。

尽管基于 YOLOv8n 的边缘计算设备对野生动物检测与跟踪系统在精度、稳定性和实时性方面取得了显著成果,但仍面临若干亟待突破的瓶颈。当前模型主要依赖可见光数据,导致其在夜间红外或低光照场景下的目标检测性能显著下降。在目标遭遇长时间遮挡或快速机动时,尽管卡尔曼滤波能提供短期预测,但轨迹漂移与身份切换问题仍难以完全避免。训练数据中部分样本源自开源视频,其标签边界模糊与类别歧义问题,使得模型对标签误差较为敏感,进而影响系统性能的稳定性的。

未来研究可以从以下两方面推进:一是在感知能力上做深化。通过结合红外与可见光信息提升在不同光照条件下的识别稳定性,同时利用重识别技术改善遮挡情况下的身份保持,使长时间监测更可靠。二是在数据与模型适应性上做加强。借助弱监督或主动学习减少标注成本,并逐步扩充物种和场景类型,以得到更具代表性的数据基础,从而提升系统在复杂环境中的适用性。

参考文献

[1] World Wildlife Fund. Living Planet Report 2022:

Building a nature-positive society[R/OL]. 2022-10-14. <https://www.worldwildlife.org/publications/2022-living-planet-report/>.

[2] 王海群,王炳楠,葛超.重参数化 YOLOv8 路面病害检测算法[J].计算机工程与应用,2024,60(5):191-199.

Wang Haiqun, Wang Bingnan, Ge Chao. Reparameterized YOLOv8 pavement disease detection algorithm[J]. Computer Engineering and Applications, 2024, 60(5): 191-199.

[3] 蒙素素,康家荣,杨秀增.基于目标检测的动物识别研究[J].现代信息科技,2025,9(10):58-63.

Meng Susu, Kang Jiarong, Yang Xiuzeng. Research on animal recognition based on target detection[J]. Modern Information Technology, 2025, 9(10): 58-63.

[4] 程焕新,乔庆元,骆晓玲,等.基于改进 YOLOv8 的无人机航拍图像目标检测算法[J].无线电工程,2024,54(4):871-881.

Cheng Huanxin, Qiao Qingyuan, Luo Xiaoling, et al. Target detection algorithm for UAV aerial images based on improved YOLOv8[J]. Radio Engineering, 2024, 54(4): 871-881.

[5] Yuan C M, Liu J L, Wang H B, et al. Object detection in complex traffic scenes based on environmental perception attention and three-scale feature fusion[J]. Applied Sciences, 2025, 15(6): 3163.

[6] 王凌云,万旭东,刘兰兰,等.基于改进 YOLOX 的输电线路典型部件及缺陷轻量检测方法[J].电机与控制学报,2025(9):106-115.

Wang Lingyun, Wan Xudong, Liu Lanlan, et al. Light weight detection method for typical components and defects of transmission lines based on improved YOLOX[J]. Electric Machines and Control, 2025(9): 106-115.

[7] 冉宁,施高朗,张少康,等.基于 YOLOv8 的遥感小目标检测算法[J].电子测量与仪器学报,2025,39(5):197-207.

Ran Ning, Shi Gaolang, Zhang Shaokang, et al. A YOLOv8-based algorithm for remote sensing small object detection[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(5): 197-207.

[8] 周畅达,杨帆.基于元素相乘结构的实时多目标跟踪算法[J].电子测量技术,2025,48(3):43-51.

Zhou Changda, Yang Fan. A real-time multi-object tracking algorithm based on element-wise multiplication structure[J]. Electronic Measurement Technology, 2025, 48(3): 43-51.

[9] Subedi A. Improving generalization performance of YOLOv8 for camera trap object detection[PP/OL]. V1. (2024-12-18) [2025-09-10]. <https://doi.org/>

- 10.48550/arXiv.2412.14211.
- [10] Cao J S, Bao W S, Shang H X, et al. GCL-YOLO: A GhostConv-based lightweight YOLO network for UAV small object detection[J]. Remote Sensing, 2023, 15(20): 4932.
- [11] Chen W J, Liu J M, Liu T, et al. PCPE-YOLO with a lightweight dynamically reconfigurable backbone for small object detection[J]. Scientific Reports, 2025, 15: 29988.
- [12] 霍旭,盖绍彦,洪濡,等.基于多阶段关联的多目标跟踪算法[J].仪器仪表学报,2023,44(11):205-214.
Huo Xu, Gai Shaoyan, Hong Ru, et al. A multi-object tracking algorithm based on multi-stage association [J]. Chinese Journal of Scientific Instrument, 2023, 44(11): 205-214.
- [13] Beery S, Cole E, Gjoka A. The iWildCam 2020 competition dataset[PP/OL]. V1. (2020-04-21)[2026-01-05]. <https://arxiv.org/abs/2004.10340>.
- [14] 中央电视台. 动物世界[EB/OL]. [2026-01-05]. <https://tv.cctv.com/lm/dwsj/index.shtml>.
- [15] Minden Pictures. Wildlife and nature stock photography agency[EB/OL]. [2025-09-10]. <https://www.mindenpictures.com/>.
- [16] 李衍照,于镭,田金文.基于改进 YOLOv5 的金属焊缝缺陷检测[J].电子测量技术,2022,45(19):70-75.
Li Yanzhao, Yu Lei, Tian Jinwen. Metal weld defect detection based on improved YOLOv5[J]. Electronic Measurement Technology, 2022, 45(19): 70-75.
- [17] 吕禾丰,陆华才.基于 YOLOv5 算法的交通标志识别技术研究[J].电子测量与仪器学报,2021,35(10): 137-144.
Lyu Hefeng, Lu Huacai. Research on traffic sign recognition technology based on YOLOv5 algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(10): 137-144.
- [18] 赵振兵,冯烁,赵文清,等.融合知识迁移和改进 YOLOv6 的变电设备热像检测方法[J].智能系统学报,2023,18(6): 1213-1222.
Zhao Zhenbing, Feng Shuo, Zhao Wenqing, et al. Thermal image detection method for substation equipment by incorporating knowledge migration and improved YOLOv6 [J]. CAAI Transactions on Intelligent Systems, 2023, 18(6): 1213-1222.
- [19] 赵佰亭,张晨,贾晓芬. ECC-YOLO: 一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2024, 38(4):108-116.
Zhao Baiting, Zhang Chen, Jia Xiaofen. ECC-YOLO: An improved method for steel surface defect detection [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [20] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023:7464-7475.
- [21] 王迎龙,孙备,丁冰,等. BG-YOLO:复杂大视场下低慢小无人机目标检测方法[J]. 仪器仪表学报,2025, 46(2):255-266.
Wang Yinglong, Sun Bei, Ding Bing, et al. BG-YOLO: A detection method for low-slow-small UAV targets in complex wide-field scenarios [J]. Chinese Journal of Scientific Instrument, 2025, 46 (2): 255-266.
- [22] Liu Y C, Shao Z R, Hoffmann N. Global attention mechanism: Retain information to enhance channel-spatial interactions[PP/OL]. V1. (2021-12-10)[2025-09-10]. <https://doi.org/10.48550/arXiv.2112.05561>.
- [23] Zheng T, Zhao S, Liu Y, et al. SCALoss: Side and corner aligned loss for bounding box regression[C]// AAAI Conference on Artificial Intelligence, 2022, 36 (3): 3535-3543.
- [24] 陈美龙,赵新华,叶秀芬.一种基于 ByteTrack 的前视声呐多目标跟踪算法[J]. 仪器仪表学报,2025,46(7): 332-344.
Chen Meilong, Zhao Xinhua, Ye Xiufen. A multi-object tracking algorithm for forward-looking sonar based on ByteTrack[J]. Chinese Journal of Scientific Instrument, 2025, 46(7): 332-344.

作者简介

张兴琪,本科,主要研究方向为图像识别与目标跟踪、边缘端轻量模型优化。

E-mail:m1368456882@163.com

谢永华(通信作者),博士,教授,主要研究方向为无人机载智能视觉检测、机器学习与智能控制、机电一体化等。

E-mail:zdhxyh@163.com