

多尺度特征融合的无人机红外图像目标检测方法

常聚忠, 戴迪, 粟梦涵, 刘虎, 吴连顺
国网湖北省电力有限公司直流公司, 湖北 宜昌 443001
DOI: 10.61369/TACS.2025140006

摘 要 : 针对无人机红外图像中目标较小、背景杂波较强及硬件算力限制等问题, 研究基于多尺度特征融合方法的无人机红外图像目标检测算法, 通过融合不同尺度下的特征图谱增强小目标检测效果。关键技术包括改进的特征金字塔结构、引入注意力模块以及改进的特征融合策略, 在识别精度和鲁棒性方面均有所提高; 实验结果表明: 在浅层微粒度信息的基础上融入深层语义信息可显著降低漏检率, 并能适应复杂多变的场景变化情况。本工作为无人机的在线红外检测提供了一种快速有效的解决方案, 并推动了遥感应用于安防和救援等领域。

关 键 词 : 多尺度特征融合; 无人机红外检测; 目标检测; 特征金字塔; 注意力机制

Multi-Scale Feature Fusion-Based Target Detection Method for UAV Infrared Images

Chang Juzhong, Dai Di, Su Menghan, Liu Hu, Wu Lianshun
DC Company, State Grid Hubei Electric Power Co., Ltd., Yichang, Hubei 443001

Abstract : Addressing the challenges of small target size, strong background clutter, and limited hardware computational power in UAV infrared images, this study investigates a target detection algorithm based on multi-scale feature fusion. By integrating feature maps at different scales, the algorithm enhances the detection performance of small targets. Key technologies include an improved feature pyramid structure, the introduction of an attention module, and an enhanced feature fusion strategy, which collectively improve recognition accuracy and robustness. Experimental results demonstrate that incorporating deep semantic information into shallow fine-grained information significantly reduces the missed detection rate and enables adaptation to complex and variable scene changes. This work provides a fast and effective solution for online infrared detection by UAVs and promotes the application of remote sensing in fields such as security and rescue.

Keywords : multi-scale feature fusion; UAV infrared detection; target detection; feature pyramid; attention mechanism

引言

虽然无人机红外成像具有全天时的优点, 在低空遥感领域得到了广泛应用, 但对于目标探测而言仍存在显著困难: 红外图像分辨率低、目标间对比度小导致难以提取特征信息; 因热辐射噪声和其他杂乱场景造成的环境干扰易产生误判。最后, 考虑到无人机有限的计算资源, 需要找到一个既能够得到小目标有效检测又不牺牲实时性能的方法。单次尺度检测器难以达到这种目的。因此, 利用多尺度信息进行特征级融合是可行的选择。

一、多尺度特征融合的无人机红外图像目标检测技术

针对军事侦察、灾害救援以及环境监测等领域, 无人机红外图

像目标检测技术具有重要意义, 在夜视和遮挡条件下可以利用目标物体热辐射穿透烟云阴影实现对目标的探测。然而传统方法面临着巨大的挑战: 目标尺度跨度大, 目标为长距小尺寸物体, 易受背景

项目信息:

国网湖北省电力有限公司科技资助项目: 反无人机全域被动智能防御技术研究及设备研制 (521521240005)。

作者简介:

常聚忠 (1983.04—), 男, 汉族, 河南省滑县人, 本科学历, 三峡大学, 国网湖北省电力公司直流公司就职, 高级工程师, 主要从事特高压换流站运维管理工作。

戴迪 (1983.4—) 男, 汉族, 湖北省武汉市人, 硕士研究生, 三峡大学, 国网湖北省电力公司直流公司就职, 主要从事直流检修工作。课题: 参与国网项目课题结题2项、省公司科技项目课题结题5项。

粟梦涵 (1995.12—) 女, 土家族, 湖北省建始县人, 硕士研究生, 中国科学院大学, 国网湖北省电力公司直流公司就职, 主要从事直流检修工作。课题: 参与国网项目课题结题2项、省公司科技项目课题结题1项。

刘虎 (1983.10—), 男, 汉族, 湖北省宜昌市人, 本科, 武汉大学, 国网湖北省电力公司直流公司就职, 高级工程师, 主要从事直流输电研究工作。课题: 主持省级课题结题2项, 市级课题结题5项。

吴连顺 (1989.12—), 男, 汉族, 河南省商丘市人, 研究生, 武汉理工大学, 国网湖北省电力有限公司直流公司就职, 主要从事特高压换流站运维工作。作为主要参加人完成《可移动升降式降温风机》QC课题, 并获得湖北省电力有限公司三等奖。

噪声影响导致无法辨识；红外图像分辨率低，缺少特征信息；受复杂背景影响误检率高。多尺度特征融合的方法可以有效提升目标检测的精度以及鲁棒性，已成为解决以上问题的重要手段；其核心思想是模拟人眼逐级观测的过程，从粗到精、逐步累积出丰富且互补的表征信息以实现复杂环境下的目标检测的有效性。

FPN 特征金字塔网络作为多级特征融合的前提，可自动提取不同尺寸目标的不同特征信息。例如小尺寸物体（如 3×3 像素的飞行器）可能在高层特征图上模糊，而底层特征图则可以保留边缘角点的细节；对于大型物体（例如汽车），高阶语义信息被用来解释全局结构；多模态信息融合过程通常包括三个层次：底层特征主要检测局部纹理信息；中层特征处理物体的部分；高层特征解释整个场景。该级联模块能有效增强模型对于不同尺度特征的捕捉能力，同时消除了感受野固定的缺点。其中包含着三个方面技术优势：(1) 特征冗余性可减少漏检概率，例如对于城市废墟场景的小目标，当其发生遮挡后可通过不同尺度特征进行互补检测；(2) 背景抑制性提升，即通过增强对比度模块实现对目标与热噪声干扰的有效区分；(3) 在提升计算性能的基础上进行网络压缩以适应无人机的实时性要求，并实现推理时间在毫秒级，相比于单尺度方案可以有效提升小目标的检测精度并减少漏检率，可用于移动场景中的无人机巡检作业中。

（一）特征金字塔网络的改进

针对无人机图像尺寸变化大的问题，BIMAFPN 引入了跨层双向链接进行特征融合，在融合过程中不仅能够实现不同层级的特征图同时处理，而且还可以用可训练的权重自适应融合低层的纹理信息和高层的语义信息。将该方法应用到无人机在森林火灾预警中的实际场景中。BIMAFPN 能够有效地捕获由于树叶抖动引起的动态噪声，并提高火源检测效果；其优势在于能够在计算开销以及效率上进行权衡，使用状态空间模式来捕获长期相关性并减少特征丢失的风险。一个可行的方法就是使用特征金字塔网络中上级和下级之间的融合来弥补单一尺度功能上的不足。为提高无人机红外探测性能，改进了 FPN 结构，有利于提升小目标表征能力。将 PAFPN 的三阶立方体扩展并与深层清晰特征以及高层模糊特征相融合，能有效提高目标定位精度。

（二）双向多分支辅助特征金字塔

针对无人机图像尺寸变化大的问题，BIMAFPN 引入了跨层双向链接进行特征融合，在融合过程中不仅能够实现不同层级的特征图同时处理，而且还可以用可训练的权重自适应融合低层的纹理信息和高层的语义信息。将该方法应用到无人机在森林火灾预警中的实际场景中。BIMAFPN 能够有效地捕获由于树叶抖动引起的动态噪声，并提高火源检测效果；其优势在于能够在计算开销以及效率上进行权衡，使用状态空间模式来捕获长期相关性并减少特征丢失的风险。

（三）超分辨率预处理与多尺度结合

由于红外图像的分辨率较低导致目标不清晰，超分辨率重建能够有效放大小目标特征。采用 EDSR 模型进行原图超分辨率重建，并使用 Feature Pyramid Network 对其进行改进。例如在边境巡逻夜间摄像头中，该方法可以有效地优化两个像素级人体特征并扩大其检测规模，相当于在目标对象本身就具有模糊性的前提下为其戴上一副“眼镜”，从而使得后续的多尺度融合工作更加精准有效。

（四）通道压缩与轻量化设计

针对无人机的嵌入式系统特点，采用通道压缩降低特征图通道数减少计算复杂度，并引入多尺度融合的思想，选用轻量级主干网提取基础特征信息后，结合部分感知空间注意力 (PSA) 增强关键位置的特征表达能力。在本文所使用的交通流计数场景下，该方法具有较高的精度的同时计算速度也显著提升，在无人机巡检这种对实时性有较高要求的应用中具有一定优势。

（五）动态注意力检测头

通过动态注意力机制我们可以自适应地增强特征来聚焦于我们的目标，而 DyHead 在头中引入了空-信道 joint-aware 的方式来增加对小对象的关注，例如在救援场景下，这有助于缓解建筑物的遮挡效应，并且增加了存活者热量辐射特性识别率。核心是它能够从学习到的特征图像关系中生成权重值，让其模型更为稳固。

（六）多模态数据融合

单个红外模式会因为外界因素而不能很好地识别目标物，但如果配合可见光 (RGB) 模式则能够弥补这个缺点。在多模式信息融合下提高对于复杂背景下物体的识别效果，在融合了植物表面纹理特征的同时也融合了温度特征。也可以减少误判的情况发生。该方法的应用过程可以分成两步进行，即特征层的集成或者是判决层的集成，在确保能够达到较高准确率的检测结果的前提下对计算机系统资源进行有效的配置。

二、多尺度特征融合的无人机红外图像目标检测方法应用

为了更好地解决红外图像中的无人机目标特征缺失等问题，采用选择性坐标注意力机制以不对称卷积核多方向获取不同尺度和形态的目标特征，充分表征无人机复杂的外形，并将无人机特征位置信息进行编码，让网络更好地利用这些信息来有意识地对待每个位置上的特征并可以基于它的位置信息自适应地调整特征的重要程度，放大有用部分的特征表示的同时抑制掉无效部分的贡献。

（一）基于改进 YOLOv8 的多尺度融合

虽然 YOLOv8 引入网格化预测提升模型检测速度，但对红外小目标探测仍存在不足之处。为此，采用在 neck 增加 3 个层次 PAFPN 来传递一级特征；采用 sandwich fusion 模块实现低级别空间信息的有效融合与共享。实验表明该改进方法可以显著提升 VIsDrone 数据库中的召回率，并且适用于小轿车类无人机图像的检测。

（二）SR-DETR 的超分辨率与特征增强

SR-DETR 的创新性在于其采用了超分辨率重建技术和多层级特征增强技术对红外图像低分辨进行处理，并采用 EDSR 技术对图像进行超分辨率重建，再通过 EMA-VSS 模块提取全局上下文信息特征，将融合后的多层级特征送入 Transformer 进行目标检测头。经测试该流程对一帧图像的处理时间小于 20ms，能够实现无人机上的实时目标检测的要求。

（三）PSI-YOLO 的轻量化与通道压缩

PSI-YOLO 是为少样本场景专门设计的一种网络，融合多尺度聚合和通道压缩操作，在骨干中使用轻量化结构提取浅层信息，并在 neck 中使用 SLIM-Neck 结构 (GSConv + GSBottleneck) 实现跨尺度连接；最后，检测头部应用内置损失函数以改进边界框的位置精确度。在对森林火灾监控的过程中，这

种策略既能保证轻量级的同时也能提高小型物体的识别能力。

(四) IDOU-YOLO 的双向特征金字塔

IDOU-YOLO 是在 YOLOv7 基础上改进提升而来的,在主干网络的基础上构建两个方向上的多尺度子网络辅助主干网络,实现浅层特征和深层特征自适应融合的功能;该方法的优势有:可以通过不同层级间的连接增加特征复用性,并减少运算量;此外在 detector 的末端引入了 dynamic attention 模块来增强对感兴趣区域的权重分配效果,并且实验表明该方法能够有效提升 HIT-UAV 数据集上的 mAP 值。

(五) MASF-YOLO 的多尺度特征聚合

MASF-YOLO 使用多尺度平行卷积以及特征融合的方式提升小目标检测的有效性。其中,本文创新点为提出一种多尺度特征聚合块(MFAM),利用权重聚合方式,显著提升了特征利用率;针对城市场景监控场景中高频回波干扰的问题。从而降低误报率。

(六) 基于 Transformer 的多尺度注意力机制

虽然 Transformer 结构采用自注意机制得到全局的特征关系,但其需兼容不同尺度的红外信息。一方面可将多尺度的注意力嵌入骨干网络中并动态融合特征权重;另一方面可以结合路径聚集网络(PAN)来增强特征聚合能力。这种策略对于复杂场景下增强检测框的鲁棒性具有较好的效果。

虽然显著图融合方法能检测出大多数无人机,但由于存在具有类似特性的其他物体存在于图像中,所以该方法会出现误检现象;当无人机受树木或其他物体阻挡时,这种方法同样无法检测到。相比显著图融合方法,非监督领域的算法减少了误检情况的发生,没有出现漏检的情况,但其依旧无法实现对于被树影遮挡的无人机进行检测。该方法没有出现错检情况,且能够准确检测出树下隐藏的无人机,说明我们提出的方法能够快速准确地定位所有的无人机目标在红外图像上的位置,并避免漏检和错检的情况,为实际应用提供了更加可靠高效的解决方案。

三、应用场景与实战效能验证

对于夜间安保巡逻、灾害救援以及高速公路监控等多个复杂的应用场景而言,这种多尺度特征融合方法都具有极大的优势。比如,从无人机上的红外热图像中精准识别出可疑人员或者车辆,并能借助于该网络强大的多尺度属性将可疑人员或车辆同周围干扰性的热点区域区分开来,因此该方法几乎不会出现误判现象。对于复杂的红外图像来说,在存在大目标的情况下,由于大目标在红外图像中的目标明显且占据大量像素点,所以很容易获得其主要信息而精准定位;但是对于一些小目标而言,虽然他们占据的像素数量很小,并且特征可能并不清晰,但本文方法在采用优化后算法及模型架构仍能准确识别并对其进行定位,因此可在无人机红外图像中对目标进行有效检测与追踪。

例如,针对城市监视应用,本系统能够同时检测出高楼建筑楼顶无人机及街区内行走的人群与流动人口,适应目标尺寸的极端变化;针对救灾应用,如震后搜救中,在厚烟层中使用红外图像搜索幸存者热源特征,通过多尺度融合增强小目标的检测性,从而提升了救灾效果。在交通管理环境下,无人机摄影摄像测量中的车流量分析可以实现精准的数据信息,其中的技术优势在于对堵塞道路

中小型汽车目标的有效性,并且能够在阴雨天或者雾气环境中稳定工作。结论表明该方法可有效提升无人机平台的自主性,在社会治安防控等方面具有广阔的应用前景;随着后续算法研究深入,其在农作物病虫害防治等领域的智能应用也将成为可能。

为实现高效多尺度融合,我们尝试对模块设计进行创新研究,目前主要集中在特征筛选及轻量化结构方面。基于特征筛选可自适应性地筛选出重要特征,在火灾识别中将更多关注于热点区域,而不关注无效背景。另一种做法是采用轻量化架构,即利用深度可分离卷积或通道剪枝在降低运算量的同时不牺牲精度。

例如,结合了局部感受野空间注意力机制的组卷积可以在减少参数的同时不影响其性能,在资源受限的无人机平台上具有实际意义;另外,改进损失函数也是一个关键点,采用非单增关注方法如 Powerful-IOU 替代 IoU 可以更好地收敛 BBoxes 的位置。特别是在处理小型物体和高密度环境时表现尤为出色。这一改进不仅提升了模型的收敛速度,并增强了对恶劣环境的鲁棒性,在道路上路口的视频中可以较好地检测不同尺度目标,避免受光照变化的影响导致误检等问题。未来的趋势为融合多模态数据(如可见光和红外图像),以加深对图像内涵的理解程度。

四、结束语

基于多尺度特征融合的目标检测方法适用于无人机红外图像,能够结合不同尺度下的特征信息实现小目标识别和鲁棒性提升。后续研究可进一步探讨多模式融合及边界的处理方法的应用,推动无人机红外探测在更多领域的应用,如环境监测以及智能交通等领域,为低空遥感发展奠定基础。

参考文献

- [1] 崔令飞,郭永红,修全发,等.基于国产嵌入式智能计算平台的无人机检测方法[J].兵工学报,2022,43(21):146-154.
- [2] 张学思,张婷,刘兆英,等.基于轻量级卷积神经网络的海面红外显著性目标检测方法[J].山东大学学报(工学版),2022,52(2):41-49.
- [3] 张翔培,贺占庄,马钟,等.一种自标签特征点异源图像目标检测算法[J].北京理工大学学报,2021,41(5):558-568.
- [4] 赵兴科,李明磊,张弓,等.基于显著图融合的无人载热红外图像目标检测方法[J].自动化学报,2021,47(9):2120-2131.
- [5] 宋子壮,杨嘉伟,张东方,等.基于无监督域适应的低空海面红外目标检测[J].光学学报,2022,42(4):127-134.
- [6] WANG Q, ZHOU L, XU C, et al. Progress and perspectives on UAV visual object tracking[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 20214-20239.
- [7] 叶恒斌,朱允攀,周永康,等.轻量级目标检测算法综述[J].红外技术,2025,47(3):289-298.
- [8] lightweight object detection algorithms[J]. Infrared Technology, 2025, 47(3): 289-298.
- [9] WU X, LI W, HONG D, et al. Deep learning for unmanned aerial vehiclebased object detection and tracking: a survey[J]. IEEE Geoscience and Remote Sensing Magazine, 2021, 10(1): 91-124.
- [10] REN S, HE K, 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, 2025, 39(6): 1137-1149.