

基于红外偏振成像的复杂干扰环境下 无人机目标检测方法

乔新博¹, 郭 阳¹, 赵永强^{1*}, 刘 亮²

(1. 西北工业大学自动化学院, 陕西西安 710072; 2. 中国人民解放军海军航空大学, 山东烟台 264001)

摘 要: 现有雷达、射频、光学等探测手段在复杂干扰环境中检测无人机时存在误警率高、检测精度低等问题。针对这些问题, 本文利用无人机与背景红外辐射偏振信息差异较大的特点, 提出基于分焦平面红外偏振成像的复杂干扰环境无人机目标检测方法。该方法构建了一个基于注意力机制与去马赛克蒸馏的目标检测网络(Attention Network with Distillation Demosaic, ANDD), 可直接以分焦平面偏振相机获取的偏振马赛克图像为输入, 实现对无人机目标的高精度实时检测。首先设计预训练偏振解码蒸馏网络(Polarization DeMosaic Distillation Network, PDMDN)从偏振马赛克图像中快速获取噪声较少的多维偏振信息; 其次, 使用骨干网络从多维偏振信息中提取目标的纹理、形状等初步特征; 最后, 建立偏振注意力特征(Polarization Attention, PAT), 利用多维偏振信息实现复杂干扰环境无人机目标检测。为评估算法性能, 基于分焦平面红外偏振成像系统建立了复杂干扰环境下无人机目标检测数据集, 并在该数据集上与近年最先进的 6 种目标检测方法进行了比较, 验证了本文算法的有效性。

关键词: 红外偏振; 分焦平面; 深度学习; 去马赛克; 注意力网络; 目标检测

基金项目: 国家自然科学基金(No.61771391); 深圳市科技重大专项(No.KJZD20230923114159039)

中图分类号: TN219 **文献标识码:** A **文章编号:** 0372-2112(2025)11-3920-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250496

An Infrared Polarization Based UAV Detection Method for Complex Environment

QIAO Xin-bo¹, GUO Yang¹, ZHAO Yong-qiang^{1*}, LIU Liang²

(1. School of Automation, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China;

2. Naval Aeronautical University, Yantai, Shandong 264001, China)

Abstract: While detecting UAV under complex interferences, the existing methods based on radar, radio frequency and vision system obtains a high false alarm rate and low accuracy. Aimed at these problems, an UAV detection method for complex scenes is constructed based on the infrared division of focal plane (DoFP) polarization imager. By utilizing the polarization difference between UAV and background, the constructed attention network with distillation demosaic (ANDD) directly obtains the detection result from the mosaic image obtained from the DoFP imager with high accuracy in real time. Firstly, the method constructs the multi-polarized information with low noise using the pretrained polarization demosaic distillation network (PDMDN). Then, the primary features are extracted from the backbone network. Finally, the polarization attention (PAT) network is designed to obtain the UAV detection result by further applying the polarized feature. To verify the efficiency of our proposed ANDD network, an infrared polarization UAV detection dataset with complex interferences is constructed using an infrared DoFP imager. Experiences with the state-of-the-art methods demonstrate that the proposed method obtains a considerable result, demonstrate the effectiveness of the ANDD method.

Key words: infrared polarization; division of focal plane; deep learning; demosaic; attention network; object detection

Foundation Item(s): National Natural Science Foundation of China (No.61771391); Major Science and Technology of Shenzhen (No.KJZD20230923114159039)

1 引言

随着无人机技术的发展与低空空域的逐步开放,无人机逐渐在智能物流、安全监控等领域得到广泛应用^[1-3],但无人机操作简便、成本低廉的特性也使其对公共安全与社会秩序带来了一定的安全隐患^[4,5].为解决该问题,近年来人们在反无人机领域开展了大量研究.然而现有方法在检测处于复杂干扰环境中的无人机时仍存在许多问题:基于雷达的方法在复杂地形与城市环境中易受反射干扰^[6];射频探测与声学探测分别会受电磁干扰与噪声干扰^[7,8];现有基于光学成像系统的方法如可见光成像与红外成像则难以在复杂场景下检测尺寸变化剧烈、缺乏外观信息或尺度较小、分辨率较低的无人机目标^[9,10].因此,如何在复杂干扰环境下检测无人机目标仍是一个亟待解决的难题.

随着偏振成像技术的发展,红外偏振成像技术为复杂干扰环境下的无人机目标检测问题提供了一种可行的解决方案.与传统红外成像技术相比,红外偏振成像技术在获取目标红外辐射信息的同时还能够获取包含目标表面纹理、材质、形状等理化特性的偏振信息^[11-15].红外偏振成像能够有效抵抗烟雾、扬尘、耀光与复杂背景等干扰对成像的影响^[16,17],并利用目标与背景间的偏振信息差异凸显目标^[18,19].由于无人机表面大量使用塑料与金属等偏振信息较强的材料且表面光滑度较高,使无人机具有较强的偏振信号^[20,21],因此可以采用红外偏振成像技术检测无人机目标以解决复杂干扰环境下的无人机目标检测问题.

经过增强后的传统红外成像与红外偏振成像拍摄的无人机图像对比如图 1 所示,可以看到无论以天空为背景还是以复杂楼房为背景,红外偏振成像都能够利用偏振信息大幅提升无人机目标与背景间的对比度,降低无人机目标的检测识别难度.

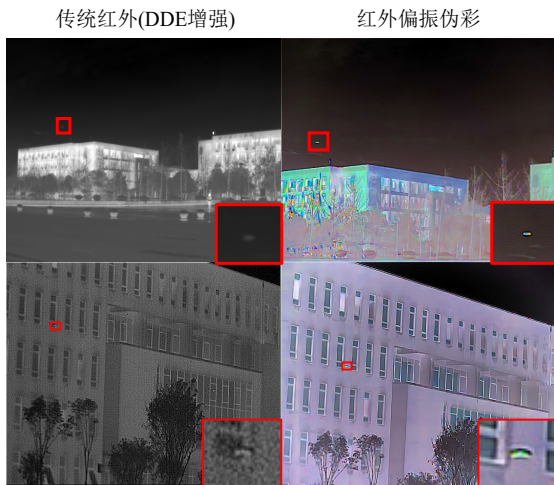


图 1 DDE 增强传统红外图像与红外偏振伪彩图像对比

传统红外偏振成像方法如分时成像方法^[17]、分振幅成像方法^[22,23]与分孔径成像方法^[24-26]分别存在不能实时成像、体积过大、通道间互相干涉等缺陷,难以有效应用于无人机目标检测任务.近年来随着微纳加工技术的发展,具有体积小、质量低、实时性强等优势的分焦平面红外偏振成像方法^[27,28]逐渐在目标检测等领域得到广泛关注.

分焦平面红外偏振成像方法获取的偏振马赛克图像需要经过去马赛克处理,恢复出与原始图像大小相同的不同偏振方向图像以计算图像偏振信息.现有目标检测算法并未考虑偏振马赛克图像中包含的偏振信息,若将红外偏振马赛克图像直接应用于无人机目标检测过程,则难以利用图像中的偏振信息提升目标检测性能;而若对图像进行去马赛克与偏振信息解算后再进行检测,会使解算后得到的偏振信息中包含的大量噪声影响目标检测过程,降低算法检测性能.因此,如何有效利用分焦平面红外偏振成像方法获取的偏振马赛克图像提升无人机目标的检测性能,仍是一个亟待解决的难题.

针对以上问题,本文设计了一种基于注意力机制与去马赛克蒸馏网络的目标检测网络(Attention Network with Distillation Demosaic, ANDD).为解决目标检测算法难以有效利用马赛克图像中的偏振信息检测目标的问题,本文在 ANDD 网络内构建了偏振解码蒸馏网络(Polarization DeMosaic Distillation Network, PDMDN)与偏振注意力特征网络(Polarization Attention, PAT),其中 PDMDN 网络在预训练时能够从教师网络中学习如何获取噪声更低、目标显著性更强的偏振信息,使网络能够充分利用马赛克图像中的偏振信息;PAT 网络则整合了全局、局部空间注意并构造了偏振注意力网络,使网络能够利用不同层级偏振特征的注意力信息快速定位目标位置,减弱偏振信息中的噪声对检测性能的影响,提升目标检测性能.为验证本文算法的准确性与实时性,本文利用自研的分焦平面红外偏振成像设备^[29]采集、标注了大量不同干扰场景下的不同种类无人机数据,构建了红外偏振马赛克无人机检测数据集,并在该数据集上与近年最先进的目标检测算法进行了比较,验证了算法的有效性.

2 偏振成像基础

2.1 偏振参数计算

偏振成像技术通过获取光在不同偏振方向的强度信息并对其解算以获取场景偏振信息.通常,偏振成像设备会获取光在 0° 、 45° 、 90° 、 135° 四个不同方向的强度信息,并利用光在不同方向上的强度信息计算 Stokes 矢量以量化表述偏振信息^[26]:

$$\mathbf{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(I_0 + I_{45} + I_{90} + I_{135}) \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ I_R - I_L \end{bmatrix} \quad (1)$$

其中, $I_0, I_{45}, I_{90}, I_{135}$ 分别表示光在 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 偏振方向上的强度信息; S_0 表示入射光的强度信息; S_1 表示光的水平、垂直分量的强度差; S_2 表示光的 $+45^\circ$ 与 -45° 分量的强度差; S_3 表示圆偏振右旋光与左旋光的强度差, 由于该值较小, 因此通常假定为 0.

$$\text{DoLP} = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \quad (2)$$

$$\text{AoP} = \frac{1}{2} \arctan\left(\frac{S_2}{S_1}\right) \quad (3)$$

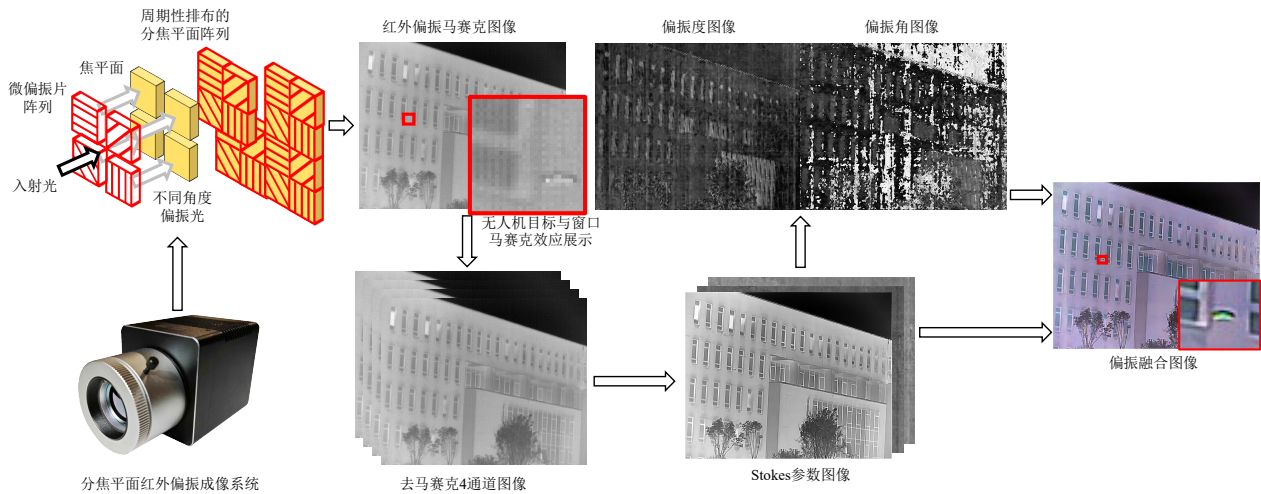


图2 分焦平面红外偏振成像系统工作原理与偏振马赛克图像处理流程

由于分焦平面红外偏振成像系统的每个像元只能获取一个偏振方向的信息, 致使获取的各通道图像的长、宽均为马赛克图像的 $1/2$, 因此需要对图像进行去马赛克处理, 以复原出与原始图像分辨率相同的偏振图像, 消除该成像方式引入的瞬时视场误差^[30,31]. 这之后, 通过计算 Stokes 参数, 解算图像偏振度、偏振角, 即可得到图像偏振信息. 然而从图 2 可以看到, 尽管马赛克图像与去马赛克图像中没有显著的噪声, 但偏振度、偏振角的计算过程对噪声的放大使偏振度、偏振角图像中出现了大量噪声. 与此同时, 现有去马赛克算法难以在获取准确偏振信息的同时保障算法实时性^[32], 致使难以在目标检测等领域得到有效应用.

3 本文方法

本文提出的基于注意力机制与去马赛克蒸馏的目标检测网络整体运行流程如图 3 所示. 网络共分为

其中, 线偏振度 (Degree of Linear Polarization, DoLP) 表示光的偏振强度, 偏振角 (Angle of Polarization, AoP) 表示光的偏振方向. 利用偏振度、偏振角即可区分具有不同形状、材质、表面粗糙度的无人机目标与背景. 由于线偏振度、偏振角分别存在平方开方与反三角函数计算, 因此线偏振度与偏振角的计算会大幅增强图像噪声, 降低偏振信息的有效性.

2.2 分焦平面红外偏振成像方法

图 2 展示了分焦平面红外偏振成像系统获取场景偏振信息的原理. 为获取光在不同偏振方向上的强度信息, 系统在焦平面上以每 2×2 的像素为一组集成了具有不同偏振方向的微偏振片阵列并将其周期性排布, 构建了分焦平面阵列. 由此, 分焦平面红外偏振成像系统便可以马赛克图像的形式实时获取光在不同偏振方向的强度信息.

三个部分: 偏振解码蒸馏网络、检测骨干网络与偏振注意力特征网络. 解码网络将从输入的 14 位红外偏振马赛克图像 I_M 中获取去马赛克后的多维偏振信息图像 I_D , 以提升小型无人机目标的显著性, 并将增强后的多维信息图像送入检测骨干网络获取三层初步特征 F , 之后, 网络将采用偏振注意力模块利用注意力机制构建注意特征图 F_A , 并通过检测头获取检测结果 I_R .

3.1 偏振解码蒸馏网络 PDMDN

为直接从马赛克图像中获取偏振信息, 提升目标与背景对比度从而提升检测算法性能, 本文构建了偏振解码蒸馏网络. 该网络不与检测算法进行同步训练, 而是使用独立的数据与教师网络通过知识蒸馏方法单独训练, 以保障算法在获取准确偏振信息的同时具有较高的计算速度.

为抑制解算过程中产生的噪声, 提升偏振信息有效性, 文献[32]系统地分析了分焦平面红外偏振成

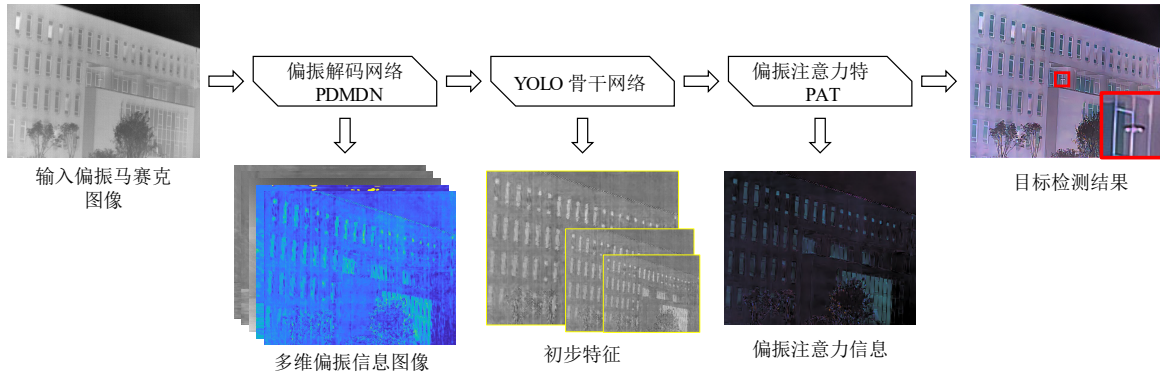


图3 ANDD网络整体流程图

像的加性噪声与泊松噪声,构建了噪声水平估计算法并将该算法应用于提出的PJNDMnet,使网络能有效抑制偏振信息的解算对噪声的放大.然而该算法的实时性较差,无法支持实时检测、识别等任务,因此本文采用文献[32]的PJNDMnet作为PDMDN的教师网络,并使用该文献构建的分焦平面去马赛克数据集进行训练,以提升PDMDN网络的去噪去马赛克性能.

PDMDN网络的结构如图4所示,该子网络共分为三层.输入的偏振马赛克图像 I_M 经过初步处理后将首先经过重构模块拆分为4张不同偏振通道的图像 $I_{Channel}$ 送入网络的第二层:

$$I_{Channel} = \text{Shuffle}(I_M) \quad (4)$$

其中, $I_{Channel} = [I_0, I_{45}, I_{90}, I_{135}]$, 分别为 0° 、 45° 、 90° 、 135°

通道的图像,长宽均为输入偏振马赛克图像的1/2;然后,这4张不同偏振通道图像将进一步拆分,送入网络第三层构建每个通道的去噪增强信息 F_E ,并与不同通道图像一同送入网络第二层构建4通道特征信息 $F_{Channel}$:

$$F_{Channel} = \text{SecondLayer}([F_E, I_{Channel}]) \quad (5)$$

其中, $F_{Channel} = [F_0, F_{45}, F_{90}, F_{135}]$ 表示第二层网络从4通道图像中获取的4通道特征信息,该特征信息将与输入的偏振马赛克图像一同送入第一层的主干网络,并作为辅助信息重构出去马赛克多维偏振信息图 I_D :

$$I_D = \text{FirstLayer}([F_{Channel}, I_M]) \quad (6)$$

其中, $I_D = [I_{S_0}, I_{DoLP}, I_{AoP}]$, 分别为去马赛克后的 S_0 、DoLP、AoP图像,以不同通道的形式保存到 I_D 中.

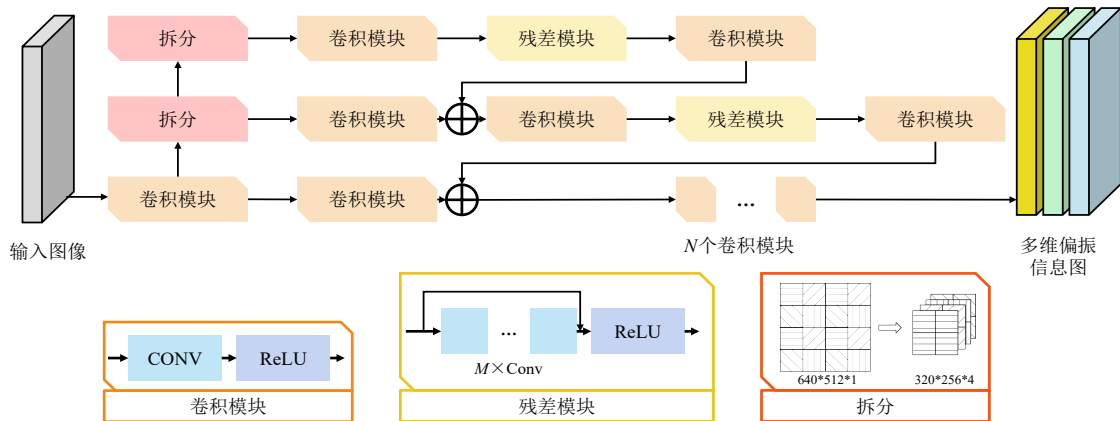


图4 PDMDN网络结构示意图

由于PDMDN网络不与检测算法共同训练,因此需要独立的损失函数训练以计算其损失值 L_{PDMSN} :

$$L_{GT} = \lambda_1 \|I_{S_0}^{GT} - I_{S_0}\|_1 + \lambda_2 \|I_{DoLP}^{GT} - I_{DoLP}\|_1 + \lambda_3 \|I_{AoP}^{GT} - I_{AoP}\|_1 \quad (7)$$

$$L_{Teacher} = \lambda_4 \|I_{S_0}^T - I_{S_0}\|_1 + \lambda_5 \|I_{DoLP}^T - I_{DoLP}\|_1 + \lambda_6 \|I_{AoP}^T - I_{AoP}\|_1 \quad (8)$$

$$L_{PDMSGN} = L_{GT} + L_{Teacher} \quad (9)$$

其中, $I_{S_0}, I_{DoLP}, I_{AoP}$ 为PDMDN输出的多维信息图; $I_{S_0}^{GT}, I_{DoLP}^{GT}, I_{AoP}^{GT}$ 为真值图像的多维信息图; $I_{S_0}^T, I_{DoLP}^T, I_{AoP}^T$ 为教师网络输出的多维信息图; $\lambda_1, \lambda_2, \dots, \lambda_6$ 为权值,总和为1; $\|\cdot\|_1$ 表示L1损失; L_{GT} 与 $L_{Teacher}$ 分别为多维信息图与真值图像、教师网络输出结果之间的损失值.

由此,PDMDN网络便可从偏振马赛克图像中解算出噪声更少的去马赛克偏振信息,提升偏振信息的有效性.

3.2 偏振注意力特征网络PAT

PDMDN网络输出的多维偏振信息图将通过基于Yolo v8的骨干网络进行初步特征提取,以获取包含语义信息、纹理信息等的不同层级初步特征 $F=[F_1, F_2, F_3]$ 之后,初步特征 F 将输入偏振注意力特征网络利用图像的注意力特征快速确定目标的准确位置,减弱偏振信息中的噪声对检测结果的影响.

PAT网络共分为3个子网络,以分别获取全局空间注意信息、多维偏振注意信息与局部空间注意信息.输入PAT网络的初步特征将首先通过全局空间注意网络获取图像的全局注意信息 A_G 以确定目标的大致位置,之后全局注意特征将分别送入多维偏振注意网络与局部空间注意网络,以获取多维偏振注意 A_p 与局部空间注意 A_L ,利用偏振信息提升检测准确性.

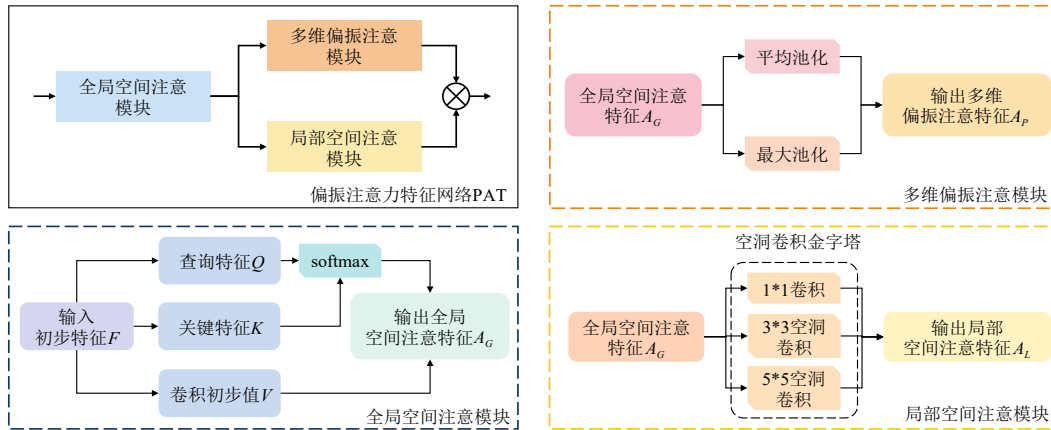


图5 本文提出的PAT网络结构示意图

由于PDMDN网络将多维偏振信息保存到图像的不同通道中,因此多维偏振注意网络采用通道注意架构设计,通过对输入的全局注意信息不同通道特征分别进行平均池化与最大池化,并将池化结果累加以获取多维偏振注意特征 A_p :

$$A_p = \sum [\text{avg}(A_G^i) + \text{max}(A_G^i)] \quad (14)$$

其中, A_G^i 表示全局注意信息的第*i*个通道.

全局空间注意获取目标的大致位置后,局部空间注意模块将利用注意机制细化目标位置.局部空间注意模块采用空洞卷积金字塔,利用1*1卷积、3*3空洞卷积与5*5空洞卷积从全局注意特征中获取局部空间注意特征 A_L :

$$A_L = \text{conv} \left[\sum_{i=1,3,5} \text{holoconv}_i(A_G) \right] \quad (15)$$

PAT与其子网络结构如图5所示.全局空间注意网络的架构基于Transformer网络^[33]的Self-attention结构改进而来,通过三组不同的卷积操作,全局空间注意网络会从输入的初步特征中获取包含目标大致位置信息的查询特征 Q 、包含目标纹理、形状、语义等不同层级信息的关键特征 K 与包含目标的初步注意信息的卷积初步值 V :

$$Q = \text{convblocks1}(F) \quad (10)$$

$$K = \text{convblocks2}(F) \quad (11)$$

$$V = \text{convblocks3}(F) \quad (12)$$

其中,convblocks1、convblocks2、convblocks3分别表示获取查询特征、关键特征与卷积初步值的卷积模块,将查询特征将与关键特征融合并进行softmax处理后与卷积初步值加权融合,即可得到图像的全局注意信息 A_G :

$$A_G = \lambda_1 \text{softmax}(Q, K) \times \lambda_2 V \quad (13)$$

其中, λ_1 、 λ_2 表示权值.

其中, holoconv_i 表示尺寸为*i*的空洞卷积.

4 实验验证与结果分析

4.1 分焦平面红外偏振无人机数据集

由于缺少公开数据集,因此本文使用自研分焦平面红外偏振相机^[29]采集了大量红外偏振无人机数据,并对数据进行逐帧标注构建了一套红外偏振无人机目标检测数据集.采集数据所使用的非制冷长波红外偏振相机工作波段为8~14 μm ,能获取包含14位有效信息的红外图像,并将数据以低14位的形式存储到16位图像中.非楼体背景数据采集于新疆某无人机展,在距离展示区约200~300 m处使用50 mm红外镜头拍摄;楼体背景数据采集于西北工业大学,使用尺寸为130 mm \times 157 mm \times 48.5 mm,型号为DJI Neo的无人机于150 m外使用50 mm红外镜头拍摄.采集过程中通

过观测拍摄数据手动调节相机固定台跟踪拍摄不同无人机目标. 数据采集完成后选取包含典型干扰场景的无人机数据, 通过逐帧标注无人机位置框构建无人机检测数据集. 采集、构建的数据集共包含超 12 000 张标注后的无人机数据, 涵盖小型穿梭机, 多旋翼无人机及固

定翼无人机三类目标. 本实验从数据集中选取 7 756 张图像, 并分别取其中 6 295 与 1 461 张图像作为训练数据与测试数据. 数据集部分图像如图 6 所示, 其中红色框为目标, 绿色框为图像中的无人机抛投物、飞鸟等干扰物体.

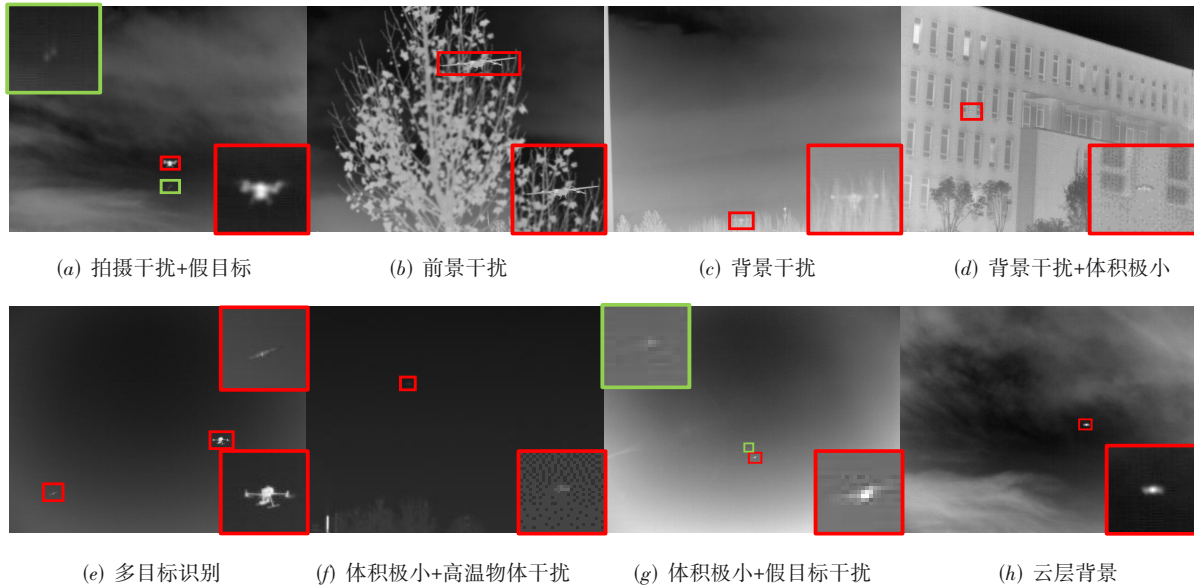


图6 数据集部分图像展示

为验证算法有效性, 数据集包含大量干扰, 具体可分为目标层面的干扰与环境层面的干扰. 其中目标层面的干扰包括与无人机外观相似、移动较快的飞鸟及无人机携带、投放的灭火弹等抛投物. 环境层面的干扰包括: 导致图像动态性能下降, 减弱目标显著性的高温物体干扰; 复杂背景如云层、楼房、植被等对无人机目标的干扰; 前景树木、行人或其他目标遮挡在镜头与无人机之间引发的前景遮挡干扰; 拍摄过程中由于相机对焦不准、镜头抖动等情况产生的拍摄干扰. 此外, 数据集中包含大量具有挑战性的检测场景, 其中包括目标尺寸小于 9×9 个像素的小目标检测; 场景中同时存在多个不同目标的多目标检测以及目标与背景温差较小、显著性较低的弱目标检测.

4.2 实验设置

本文所有实验均在 PyCharm 2024.1.1 以 Python3.9 环境在搭载 Tesla P40 的服务器上运行, 服务器内存为 23 GB. 用于对比的算法均使用默认参数与默认训练策略与训练 epoch. 由于骨干网络基于 Yolo 架构构建, 因此本文采用 Yolo 默认 epoch 数进行训练.

为验证本文算法的有效性, 本文将所提出的 ANDD 检测算法与现有最先进的目标检测算法进行比较, 其中包括 Yolo v11^[34]、Yolo v8^[34]、Yolo-FIRI^[35]、CVPR 2024 提出的 Cross_KD^[36]、Saliency_DETR^[37]、CVPR

2023 提出的 Q-DETR^[38]. 对比网络均采用默认网络参数配置与损失函数. 由于大部分去马赛克方法如双三次插值、牛顿插值多项式^[30]或本文的教师网络方法^[32]的单张图像处理时间较长, 若使用此类方法进行训练会使网络单次训练时间超过 1 个月, 故实验中不对比其他去马赛克方法对检测结果的影响.

参考对比算法, 本文使用以下评价指标: 平均召回率 (Average Recall, AR)、平均精度 (Average Precision, AP) 与 F1 分数 (F1 Score, F1), 同时为进一步比较算法计算量与实时性, 本文同时对比了不同算法参数量 (单位: M) 与单张图像的检测时间 (单位: ms). AR 表示所有类别召回率的平均值, 单个目标类别的召回率表示如下:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (16)$$

其中, TP 表示网络正确检测的真实目标数量; FN 表示模型没有预测到的真实目标数量, 通过对所有类别的召回率取均值即可得到网络 AP 值; AR 表示准确率-召回率曲线 (Precision-Recall Curve, PR Curve) 下的面积, 其中准确率与 AP 的计算分别表示为

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (17)$$

$$\text{AP} = \int_0^1 \text{Precision}(\text{Recall})d(\text{Recall}) \quad (18)$$

其中,FP表示网络错误检测的假目标数量.实验中除总AP外还会使用 AP_{50} 与 AP_{75} ,即预测框与真值框交并比为50%与75%时的AP值; $F1$ 分数为综合考虑准确率与召回率的评价指标,该指标表示如下:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (19)$$

由于Cross_KD、Saliency_DETR与Q-DETR未提供检测结果可视化代码,故本文不提供可视化结果进行定性对比,仅使用以上指标评估不同算法检测结果.

4.3 结果对比与分析

本文提出的ANDD方法与其他对比方法的检测结果对比如表1所示.同时为直观展示各方法的检测性

能,图7(a)与图7(e)~(j)展示了本文方法与其他对比方法在交并比为50%时的PR曲线图结果对比,其中Small UAV、Normal UAV、Fixed UAV分别表示小型穿梭机、多旋翼无人机与固定翼无人机.

从表1可以看到,本文提出的ANDD算法在AP、AR、 $F1$ 分数与 AP_{50} 上均取得了最优结果.与同样使用Yolo backbone的Yolo-FIRI、Yolo v11、Yolo v8相比,本文提出的ANDD方法在AP、AR、 $F1$ 分数上分别提升了[14.66%, 25.47%, 12.50%]、[12.71%, 13.54%, 12.50%]、[7.61%, 1.53%, 2.27%].同时,本文提出的ANDD算法单张图像检测时间仅15.4 ms,能以每秒64张图像的速率对无人机目标进行实时检测.

表1 红外偏振马赛克数据集上本文算法与其他目标检测方法结果对比

算法名称	AP	AR	$F1$	AP_{50}	AP_{75}	参数量/M	时间/ms
CrossKD	0.594	0.713	0.77	0.843	0.675	33.840	391.0
Saliency_DETR	<u>0.645</u>	0.771	0.82	0.821	0.771	55.858	609.1
Q-DETR	0.297	0.405	0.48	0.580	0.285	36.947	498.5
Yolo-FIRI	0.580	0.742	0.80	<u>0.876</u>	0.737	6.585	6.8
Yolo v11	0.590	0.820	0.80	0.830	0.639	<u>9.428</u>	<u>13.6</u>
Yolo v8	0.618	<u>0.917</u>	<u>0.88</u>	0.862	0.670	11.137	14.7
ANDD	0.665	0.931	0.90	0.905	<u>0.751</u>	17.439	15.4

注:算法名称列加粗部分为本文算法;AP、AR等指标列的加粗数值表示最优结果,下划线数值表示次优结果.

与针对红外图像目标检测构建的Yolo-FIRI相比,本文方法能够在红外偏振马赛克图像上获取反映目标表面理化特性的偏振信息,有效地提升目标检测性能.同时从PR曲线图可以看到本文方法相比Yolo-FIRI方法在检测各类型目标时具有更稳定的检测结果;与Yolo v11相比,尽管Yolo v11采用了近年最新的Yolo backbone,然而文献[39]的研究表明,Yolo v11的网络架构不适用于缺少纹理、结构信息的红外图像,且由于该算法缺少针对红外偏振马赛克图像的特征设计,难以利用偏振信息提升检测性能,致使该方法在偏振马赛克图像上的检测结果相比ANDD方法较差,甚至所有指标均劣于Yolo v8.从图7(f)可以看到,尽管在检测多旋翼与固定翼无人机时性能优秀,但Yolo v11在复杂干扰环境下检测小型穿梭机时性能很差,无法有效识别受复杂背景、镜头抖动等干扰影响的小型无人机目标;与此同时,尽管性能优于Yolo v11,但由图7(e)可知Yolo v8同样难以识别受复杂干扰的小型穿梭机,大幅降低了算法的检测性能.与Yolo v8相比,本文提出的ANDD方法能够利用PDMDN网络充分获取红外偏振马赛克图像中的偏振信息,并利用PAT网络发挥偏振信息的检测优势,提升算法在强干扰与复杂场景下对无人机目标的检测性能.

从计算速度与算法运行效率方面可以看到,PDMDN与PAT子网络的引入导致ANDD网络相比

Yolo-FIRI、Yolo v11与Yolo v8具有显著更高的参数量与相对较长的单张图像检测时间.与三类Yolo backbone网络相比,由于添加了PDMDN网络与PAT网络,ANDD网络的参数量相比三种网络分别增加了[10.854 M, 8.011 M, 6.302 M],即分别提升了[164.83%, 107.67%, 56.59%],然而得益于PDMDN网络的预训练过程,ANDD网络仍旧具有较高的实时性,能以单张图像15.4 ms,即每秒64张图像的速率实现对无人机目标的实时检测识别.

与使用其他backbone的CrossKD、Saliency_DETR、Q-DETR相比,本文提出的ANDD方法在三类分数上分别提升了[11.95%, 30.58%, 16.88%]、[3.10%, 20.75%, 10.29%]、[123.91%, 129.88%, 87.50%].

CrossKD是一种基于知识蒸馏(Knowledge Distillation, KD)的目标检测方法,通过知识蒸馏从教师网络中学习目标特征.然而不论是该算法的教师网络还是训练后的蒸馏网络均缺少针对红外偏振马赛克图像的设计,因此性能远劣于本文提出的ANDD方法;Saliency_DETR是一种基于Transformer的目标检测算法,该算法利用显著性信息过滤平衡了计算效率与精度,有效避免了尺度偏差.然而在红外偏振马赛克图像中,许多受地物背景干扰的目标与体积较小的无人机目标的红外辐射信息显著性较低,致使Saliency_DETR无法利用目标显著性信息提升此类目标的计算精度.尽管

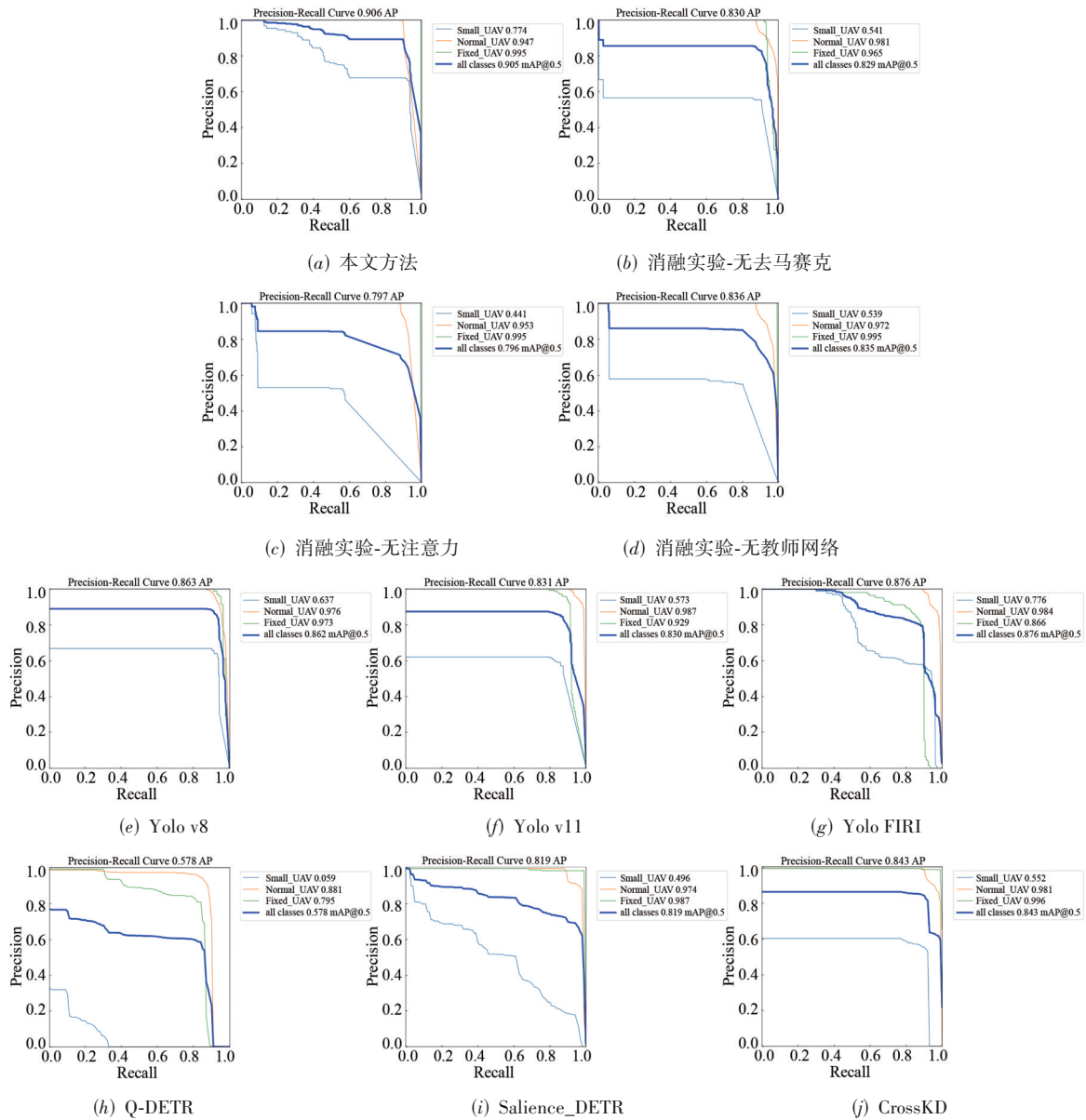


图7 本文方法与其他目标检测方法在交并比为50%时的PR曲线图结果

该方法具有最好的 AP_{75} 指标,证明方法在检测到目标的时候能够更为准确地标注目标位置,但较低的 AP 、 AR 与 $F1$ 分数表明该方法相比本文方法易受图像噪声与复杂场景的干扰,在检测受复杂干扰影响的无人机目标时准确率较低;Q-DETR 是一种基于分布式精细蒸馏的轻量化 Transformer 目标检测算法,利用双层优化问题最大化目标的信息熵以提升网络检测性能.然而与 Saliency_DETR 相似,该网络无法利用偏振信息的特性导致算法难以从红外偏振马赛克图像中获取更为有效的目标信息,进而使算法难以提升目标信息熵提升网络检测性能,致使 Q-DETR 的检测性能远劣于其他用于对比的目标检测算法.由折线图可以看到,CrossKD、Saliency_DETR 与 Q-DETR 方法得益于其优秀的设计使

其在检测多旋翼无人机与固定翼无人机时具有优秀的检测结果,但检测受复杂干扰影响的小无人机时这些算法无法利用偏振信息提升目标检测性能,大幅降低了算法有效性.

相比 CrossKD、Saliency_DETR 与 Q-DETR,本文方法得益于 Yolo backbone 而具有显著更少的参数量与更快的处理速度.与三类使用其他 backbone 的算法相比,ANDD 网络的参数量分别减少了 [16.401 M, 38.419 M, 19.508 M],即分别降低了 [48.17%, 68.78%, 52.80%].进一步分析网络结构可知,CrossKD、Saliency_DETR 与 Q-DETR 均使用 ResNet 作为网络 backbone,其网络参数量远高于基于 Yolo 骨干的 ANDD 网络的同时算法计算效率较低导致这三类网络的单张图像处理时间较长,

无法实现对无人机目标的实时检测、跟踪。

4.4 消融实验

为进一步验证本文提出的 ANDD 方法的有效性, 本文分别对比了完整 ANDD 算法、去除去马赛克 PDMDN 的算法、去除注意力特征 PAT 的算法、不使用教师网络训练 PDMDN 的算法与仅使用骨干网络(即 Yolo v8)的目标检测结果, 如表 2 与图 7(a)~(d)所示。同时, 图 8、图 9 展示了消融实验中网络中间层的输出结果, 以分析不同子网络及教师网络在检测中起到的作用, 两图目标位置真值如图 6(b)、图 6(h)所示。与去除了偏振解码网络、去除偏振注意力特征网络和不使用教师网络训练的方法相比, 完整 ANDD 算法在各评价指标上均取得了最佳的实验结果, 证明了本文算法的有效性。

从消融实验结果可以看到, 当单独使用具有去马赛克能力的偏振解码网络 PDMDN 或基于注意力机制的 PAT 网络时, 网络的总体性能相较于骨干网络反而

表 2 消融实验结果对比

算法名称	AP	AR	F1	AP ₅₀	AP ₇₅
ANDD	0.665	0.931	0.90	0.905	0.751
无去马赛克	0.611	0.750	0.76	0.829	0.681
无注意力	0.639	0.797	0.80	0.796	<u>0.722</u>
无教师网络	<u>0.644</u>	0.834	0.83	0.744	0.644
仅骨干(Yolo v8)	0.618	<u>0.917</u>	<u>0.88</u>	<u>0.862</u>	0.670

注: 算法名称列加粗部分为原始算法; AP、AR 等指标列的加粗数值表示最优结果, 下划线数值表示次优结果。

略有下降: 当去除具有去马赛克能力的偏振解码网络 PDMDN 后, 算法无法获取图像偏振信息, 只能靠红外偏振马赛克图像中的强度信息检测目标, 此时受背景干扰、高温物体干扰或本身辐射信息较弱的小型穿梭机目标显著性较低致使检测性能下降; 当去除注意力特征 PAT 时, 如图 8、图 9 所示, ANDD 网络能够利用偏振解码网络获取的偏振信息提升部分目标如固定翼无人机、小型无人机的检测准确性, 并提升算法的总 AP

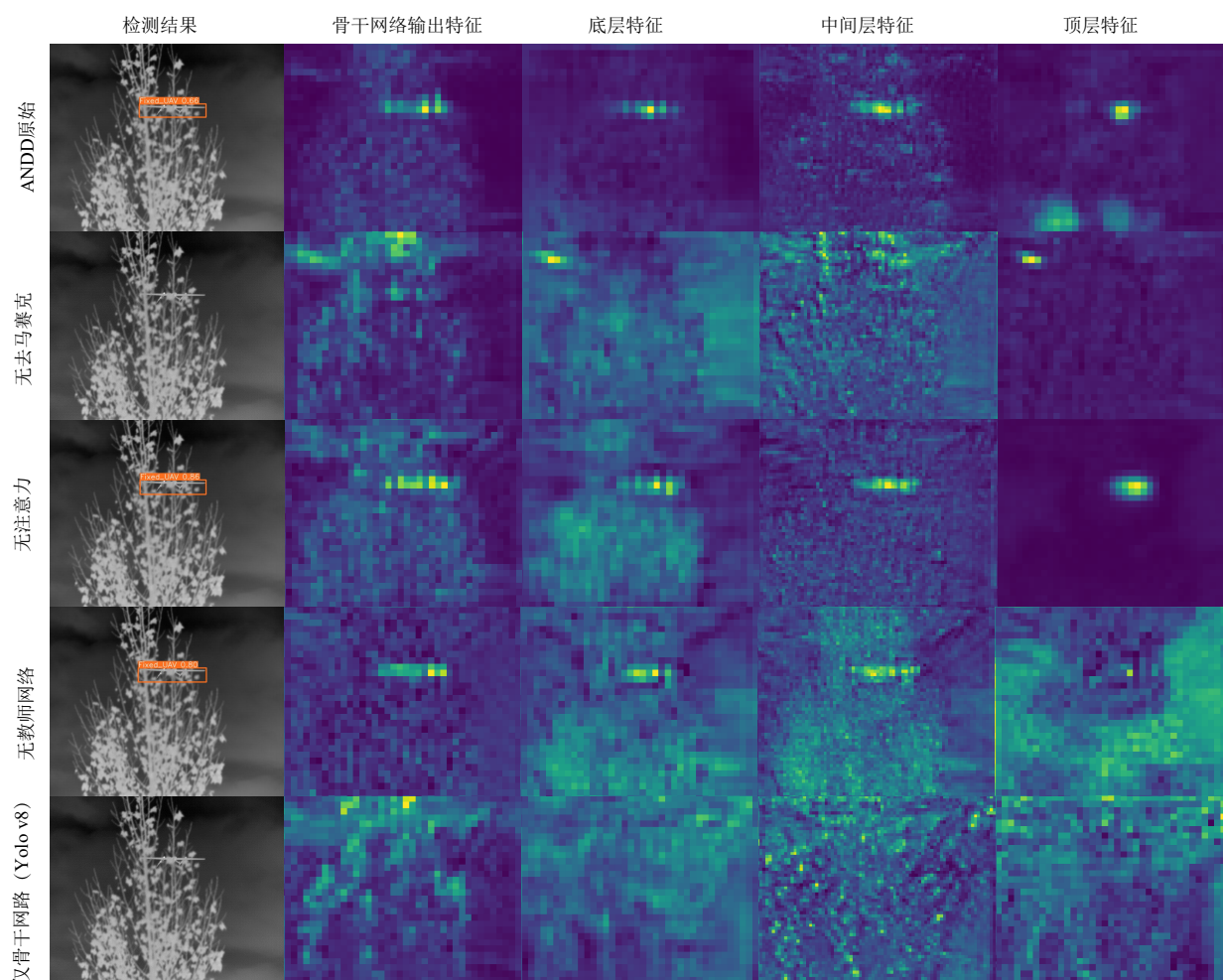


图 8 消融实验网络中间层输出

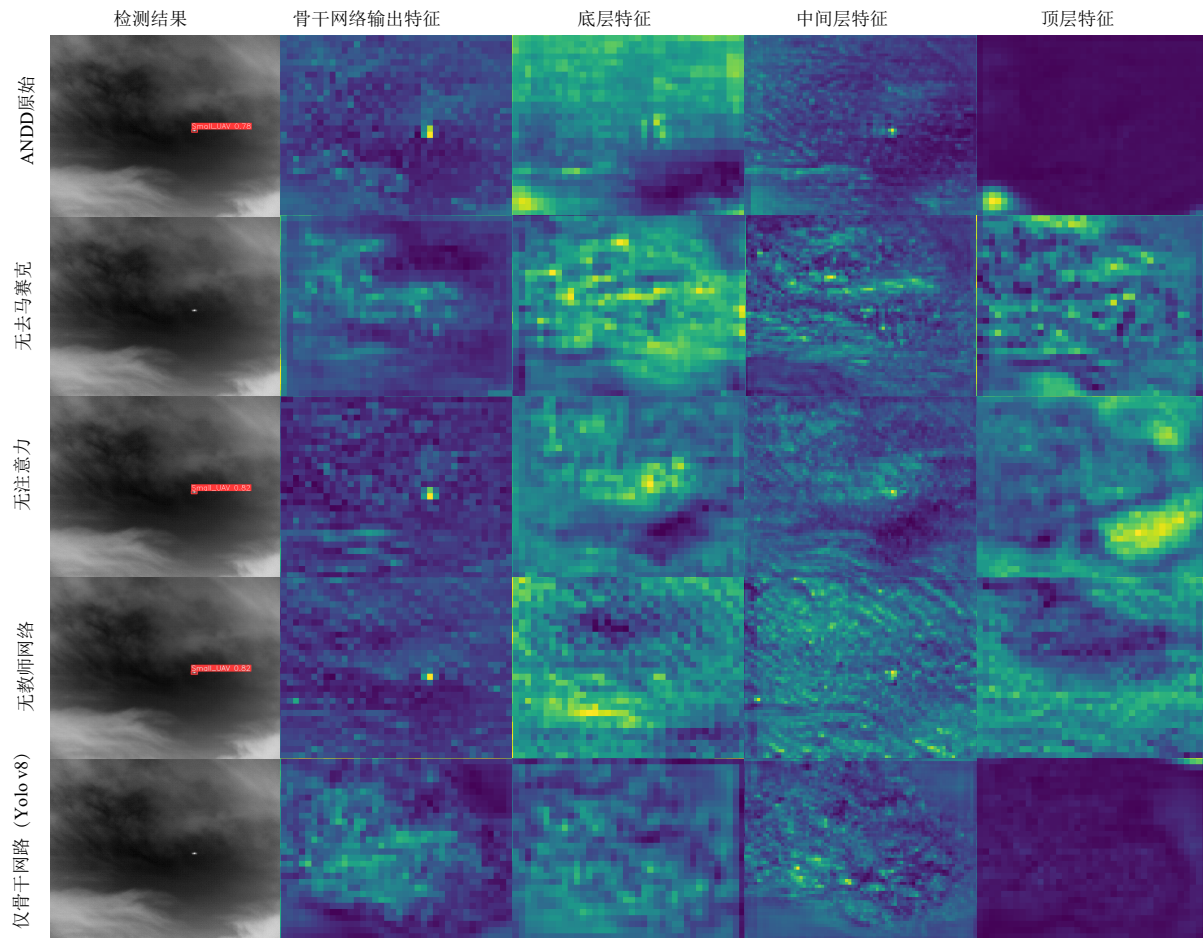


图9 消融实验网络中间层输出

与 AP_{75} 指标,然而由于骨干网络无法充分利用图像中的偏振信息抑制 PDMDN 解算后残余的噪声,致使算法的误检率 (FA 率) 提升,降低了算法的整体性能;而当网络不使用教师网络训练 PDMDN 时,联合训练得到的网络能够获取图像的部分偏振信息,使其 AP、AR 与 $F1$ 分数均优于缺少了部分子网络的版本,如图 8、9 所示,联合训练后得到的网络能够利用获取的部分偏振信息提升目标在中间层的特征显著性,进而提升检测准确性。然而由于缺少教师网络的训练,联合训练网络获取的偏振信息有效性较低,如图 8 所示,相较使用教师网络训练的原始 ANDD 网络,联合训练后的网络获取的目标特征更容易受到各类干扰的影响,降低了算法的准确性。

相比缺少去马赛克结构、缺少注意力特征与缺少教师训练的网络,完整的 ANDD 算法在使用 PDMDN 网络获取图像偏振信息、提升目标显著性的同时还能够利用注意力信息进一步提升检测准确性,抑制去马赛克残留的噪声对检测准确性的影响。从图 8 中可以看到,得益于注意力特征,被遮挡的无人机目标在 ANDD 网络的底层特征与中间层特征中具有显著更高的权

重;而在检测图 9 中的小型无人机目标时,ANDD 网络的注意力特征使其获取的底层与中间层特征包含有更多的无人机信息,有效提升了算法的检测性能。

5 结论

本文提出了一种基于红外偏振成像的目标检测方法 ANDD,该方法能利用图像中的偏振信息提升无人机目标的检测性能,实现对无人机目标的实时检测识别。算法的输入为分焦平面红外偏振成像系统获取的红外偏振马赛克图像,为获取图像中的偏振信息提升目标检测性能,本文构建了偏振解码蒸馏网络 PDMDN 并使用教师网络与偏振去马赛克数据训练,使该网络能快速获取图像准确的偏振信息,并利用偏振信息提升目标与背景之间的对比度。之后,为进一步发挥偏振信息在目标检测中的优势,本文设计了偏振注意力特征网络 PAT,在利用偏振信息增强后的图像上利用空间注意力与偏振注意力信息提升网络的目标检测性能。为验证本文提出的 ANDD 方法性能,使用分焦平面红外偏振成像系统构建了包含多种复杂干扰与不同类型无人机的无人机目标检测数据集,并在该数据集上与现

有最先进的几种目标检测算法进行了对比实验,验证了本文提出的 ANDD 方法的有效性.

参考文献

- [1] LIN N, TANG H L, ZHAO L, et al. A PDDQNLP algorithm for energy efficient computation offloading in UAV-assisted MEC[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(12): 8876-8890.
- [2] SHEN Y, PAN Z W, LIU N, et al. Performance analysis of legitimate UAV surveillance system with suspicious relay and anti-surveillance technology[J]. *Digital Communications and Networks*, 2022, 8(5): 853-863.
- [3] ZHANG Y F, ZHANG H X, GAO X R, et al. UAV target detection for IoT via enhancing ERP component by brain-computer interface system[J]. *IEEE Internet of Things Journal*, 2023, 10(19): 17243-17253.
- [4] NAJIHAH MUHAMAD ZAMRI F, GUNAWAN T S, HAJAR YUSOFF S, et al. Enhanced small drone detection using optimized YOLOv8 with attention mechanisms[J]. *IEEE Access*, 2024, 12: 90629-90643.
- [5] ELSAYED M, REDA M, MASHALY A S, et al. LERF-Net: An enlarged effective receptive field backbone network for enhancing visual drone detection[J]. *The Visual Computer*, 2025, 41(4): 2219-2232.
- [6] LI Y H, FU M Z, SUN H X, et al. Radar-based UAV swarm surveillance based on a two-stage wave path difference estimation method[J]. *IEEE Sensors Journal*, 2022, 22(5): 4268-4280.
- [7] CAI Z X, WANG Y, JIANG Q, et al. Toward intelligent lightweight and efficient UAV identification with RF fingerprinting[J]. *IEEE Internet of Things Journal*, 2024, 11(15): 26329-26339.
- [8] RAHMAN M H, SEJAN M A S, AZIZ M A, et al. A comprehensive survey of unmanned aerial vehicles detection and classification using machine learning approach: Challenges, solutions, and future directions[J]. *Remote Sensing*, 2024, 16(5): 879.
- [9] DEWANGAN V, SAXENA A, THAKUR R, et al. Application of image processing techniques for UAV detection using deep learning and distance-wise analysis[J]. *Drones*, 2023, 7(3): 174.
- [10] JIANG W, PAN H X, WANG Y P, et al. A multi-level cross-attention image registration method for visible and infrared small unmanned aerial vehicle targets via image style transfer[J]. *Remote Sensing*, 2024, 16(16): 2880.
- [11] HUANG Z R, ZHENG Y Q, LI J H, et al. High-resolution metalens imaging polarimetry[J]. *Nano Letters*, 2023, 23(23): 10991-10997.
- [12] RUBIN N A, SHI Z J, CAPASSO F. Polarization in diffractive optics and metasurfaces[J]. *Advances in Optics and Photonics*, 2021, 13(4): 836-970.
- [13] HE C, HE H H, CHANG J T, et al. Polarisation optics for biomedical and clinical applications: A review[J]. *Light: Science & Applications*, 2021, 10: 194.
- [14] ZHANG Y X, PU M B, JIN J J, et al. Crosstalk-free achromatic full Stokes imaging polarimetry metasurface enabled by polarization-dependent phase optimization[J]. *Opto-Electronic Advances*, 2022, 5(11): 220058.
- [15] LI X B, YAN L, QI P F, et al. Polarimetric imaging via deep learning: A review[J]. *Remote Sensing*, 2023, 15(6): 1540.
- [16] LI X B, HAN Y L, WANG H Y, et al. Polarimetric imaging through scattering media: A review[J]. *Frontiers in Physics*, 2022, 10: 815296.
- [17] WANG J L, DUAN J, XIE G F, et al. Sea surface target image enhancement method based on separable reflection and refraction[J]. *Optics & Laser Technology*, 2025, 181: 112012.
- [18] WANG X, DING J J, ZHANG Z, et al. IPNet: Polarization-based camouflaged object detection via dual-flow network[J]. *Engineering Applications of Artificial Intelligence*, 2024, 127: 107303.
- [19] 赵永强, 李宁, 潘泉. 分焦平面红外偏振摄像技术[M]. 北京: 科学出版社, 2022.
ZHAO Y Q, LI N, PAN Q. Division of Focal Plane Infrared Polarization Photography[M]. Beijing: Science Press, 2022. (in Chinese)
- [20] WEI M, ZHANG Y, SHI Z G, et al. Analysis of visible/infrared polarization characteristics of small UAV with complex background of buildings[C]//2020 International Conference on Image, Video Processing and Artificial Intelligence. Washington: SPIE, 2020: 77.
- [21] WANG B S, LI Q, MAO Q C, et al. A survey on vision-based anti unmanned aerial vehicles methods[J]. *Drones*, 2024, 8(9): 518.
- [22] MA X R, ZHAO H, ZENG Y Z, et al. Stokes parameters measurement system designed by symmetric division of amplitude with double quarter-wave plate[J]. *Optics Communications*, 2023, 549: 129888.
- [23] JIA W T, LIU K, JIANG K, et al. Rotating dual-retarders to correct polarization measurement error for division-of-amplitude polarimeter in full field of view[J]. *Optics and Lasers in Engineering*, 2024, 181: 108360.
- [24] PEÑA-GUTIÉRREZ S, ROYO S. Polarization calibra-

- tion assessment for a broadband imaging polarimeter based on a division of aperture architecture[J]. *Optics Express*, 2023, 31(3): 3839-3856.
- [25] ZHANG R, DUAN B, WANG R F, et al. Black phosphorus-based absorbing bionic stacked structured linear polarization detector[J]. *Advanced Optical Materials*, 2023, 11(7): 2201828.
- [26] MA F Y, ZHOU Y F, DOU W T, et al. Systematical and universal calibration scheme for division-of-aperture polarimetric camera[J]. *Optics and Lasers in Engineering*, 2025, 184: 108589.
- [27] ZHANG Y, LI H S, SUN J H, et al. Spectral polarization image reconstruction using compressed sensing method[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5530213.
- [28] 赵永强, 乔新博, 李宁, 等. 偏振视觉[J]. *中国科学: 信息科学*, 2024, 54(7): 1620-1645.
ZHAO Y Q, QIAO X B, LI N, et al. Polarization vision[J]. *Scientia Sinica (Informationis)*, 2024, 54(7): 1620-1645. (in Chinese)
- [29] QIAO X B, ZHAO Y Q, CHEN L L, et al. Mosaic gradient histogram for object tracking in DoFP infrared polarization imaging[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 194: 108-118.
- [30] LI N, ZHAO Y Q, PAN Q, et al. Demosaicking DoFP images using Newton's polynomial interpolation and polarization difference model[J]. *Optics Express*, 2019, 27(2): 1376-1391.
- [31] LIU J, DUAN J, HAO Y F, et al. Polarization image demosaicking and RGB image enhancement for a color polarization sparse focal plane array[J]. *Optics Express*, 2023, 31(14): 23475-23490.
- [32] LI N, WANG B L, GOUDAIL F, et al. Joint denoising-demosaicking network for long-wave infrared division-of-focal-plane polarization images with mixed noise level estimation[J]. *IEEE Transactions on Image Processing*, 2023, 32: 5961-5976.
- [33] RAMACHANDRAN P, PARMAR N, VASWANI A, et al. Stand-alone self-attention in vision models[J]. 2019: 68-80.
- [34] JOCHER G, DERRENGER P, JK4E, et al. Ultralytics YOLO vision[EB/OL]. (2023-11-12)[2024-09-27]. <https://docs.ultralytics.com/>.
- [35] LI S S, LI Y J, LI Y, et al. YOLO-FIRI: Improved YOLOv5 for infrared image object detection[J]. *IEEE Access*, 2021, 9: 141861-141875.
- [36] WANG J B, CHEN Y M, ZHENG Z H, et al. CrossKD: Cross-head knowledge distillation for object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 16520-16530.
- [37] HOU X Q, LIU M Q, ZHANG S L, et al. Saliency DETR: Enhancing detection transformer with hierarchical saliency filtering refinement[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 17574-17583.
- [38] XU S, LI Y J, LIN M B, et al. Q-DETR: An efficient low-bit quantized detection transformer[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 3842-3851.
- [39] YUCESOY Y F, SAHIN C. Object detection in infrared images with different spectra[C]//2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications. Piscataway: IEEE, 2024: 1-6.

作者简介



乔新博 男,1999年2月出生于黑龙江省哈尔滨市.现为西北工业大学博士研究生.主要研究方向为红外偏振成像及其应用.
E-mail: qxinbo@foxmail.com



郭阳 男,1990年1月出生于宁夏回族自治区银川市.现为西北工业大学博士研究生.主要研究方向为图像处理与传感器设计.
E-mail: guoyang665@mail.nwpu.edu.cn



赵永强 男,1976年7月出生于陕西省西安市.现为西北工业大学自动化学院教授、博士生导师.主要研究方向为信息融合、图像处理、偏振成像等.获中国专利奖、省部级科技进步奖等奖项10项.在国内外发表学术论文100余篇.
E-mail: zhaoyq@nwpu.edu.cn



刘亮 男,1981年2月出生于湖北省黄石市.现为海军航空大学副教授.主要研究方向为激光技术与红外技术.
E-mail: liul513@126.com