



· 人工智能图像处理 ·

【主持人语】随着科技的不断发展与创新,人工智能正逐渐渗透到人类生活的各个领域,改变着我们的生活方式和社会结构。图像处理作为一门跨学科的前沿技术领域,在计算机视觉、遥感监测、医学影像分析等多个学科领域取得了显著的应用与发展。在图像处理领域,人工智能的应用也越来越广泛,为图像处理带来了更高效、更准确的解决方案。本专题聚焦于人工智能深度学习算法在图像处理领域的应用研究,共收录了 4 篇研究论文,涉及图像检测、图像识别、图像评价等方面。其中:《多尺度特征增强与交互融合的遥感小目标检测》针对遥感图像小目标检测任务中存在的诸多问题,提出了一种基于多尺度特征增强与交互融合的遥感小目标检测算法 MFE-FIF,显著提升了遥感目标的检测效果;《联合多视角特征的口罩遮挡人脸识别算法》以 BoTNet 为主干网络,嵌入注意力增强模块聚焦于眼部区域,并引入注意力机制将关注点集中在口罩遮挡外的人脸区域,提取有效特征,以提高佩戴口罩的人脸识别性能;《多阶段渐进处理的图像去雨方法》提出一种多阶段渐进式处理的图像去雨算法,使上下阶段的特征融合,大大提高了去雨算法的性能;《尺度因子感知对比学习的无参考型超分辨率图像质量评价》通过设计由尺度因子识别分支和质量分数预测分支组成的双分支网络结构,并结合轻量倒残差块设计,显著提升无参考图像质量评价与主观感知的一致性。期待本专题能够引发大家探讨人工智能深度学习算法在图像处理领域更多可能的研究方向和应用方法,为从事相关交叉领域的研究人员提供一些新的思路。

【主持人】李云红,西安工程大学电子信息学院教授,硕士生导师。

多尺度特征增强与交互融合的遥感小目标检测

李云红,魏小双,苏雪平,李丽敏,田谷丰,郝特吉,冯准若,李仕博

(西安工程大学 电子信息学院,陕西 西安 710048)

摘要 针对遥感图像小目标检测任务中,存在目标细节纹理信息模糊导致特征提取与融合不佳、小目标漏检等问题,提出了一种基于多尺度特征增强与交互融合的遥感小目标检测算法。首先,采用跨层多分支连接结构的多尺度特征增强(multiscale feature enhancement, MFE)模块,利用 Split 分流操作丰富和增强不同梯度获取的纹理特征信息,同时引入轻量级特征幻影模块 Ghost 进行通道线性变换,生成更多有效的特征细节信息流,以增强对图像中局部细节特征信息的关注;其次,构建特征交互融合(feature interaction fusion, FIF)模块,引入多分支串并行的卷积块与自适应机制的池化块,交互输入特征的通道语义信息和空间特征变换,捕获全局上下文信息,精确小目标的关键位置信息,加强特征信息之间的相关性,实现细粒度特征的多维度交互融合。使用公开的光学遥感数据集 DIOR 验证所提算法,改进后的网络模型平均精度值为 87.6%,与 NPMR-Det、YOLOv7、YOLOv5 等其他 7 种优秀算法相比均有提高,改进后的遥感图像小目标检测算法取得了

收稿日期:2024-10-05

基金项目:国家自然科学基金(62203344);陕西省自然科学基金基础研究重点项目(2022JZ-35);陕西高校青年创新团队。

第一作者:李云红,女,博士,教授,从事人工智能图像处理等研究,hitliyunhong@163.com。

更好的检测精度。

关键词 遥感图像小目标检测;多尺度特征增强;Split 分流;自适应机制;细节特征交互融合
中图分类号:TP391.41 **DOI**:10.16152/j.cnki.xdxbr.2025-02-005

Remote sensing small target detection based on feature multi-scale enhancement and interaction fusion

LI Yunhong, WEI Xiaoshuang, SU Xueping, LI Limin, TIAN Gufeng,
HAO Teji, FENG Zhunruo, LI Shibo

(School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract In order to achieve remote sensing image small object detection task, there are problems such as poor feature extraction and fusion due to the blurring of target detail texture information, and leakage of small targets, a remote sensing objects detection algorithm that leverages multiscale feature enhancement and interactive fusion is proposed. Firstly, the multiscale feature enhancement (MFE) module with cross-layer and multi-branch connection structure is used to enrich and enhance the texture feature information obtained from different gradients by using the Split shunt operation, and at the same time, the lightweight feature phantom module Ghost is introduced to perform the linear transformation of the channel, generating more effective feature detail information flow to enhance attention to local detail feature information in the image. Secondly, the feature interaction fusion (FIF) module is constructed, which introduces a multi-branch serial parallel convolution block and an adaptive mechanism pooling block to interact with the channel semantic information and spatial transformation of the input features, capture the global context information, and accurately locate the key position of the small targets. information, enhance the correlation between feature information, and achieve multi-dimensional interactive fusion of fine-grained features. The proposed algorithm is validated with DIOR, which is a remote sensing dataset. The optimized network model achieves a mean accuracy precision of 87.6%, which is higher than the other seven excellent algorithms including NPMMR-Det, YOLOv7, and YOLOv5. The improved small target detection algorithm for remote sensing images achieves better detection accuracy.

Keywords remote sensing image small target detection; multiscale feature enhancement; Split triage; adaptive mechanism; interactive fusion of detailed features

近年来,随着遥感技术不断发展,遥感目标检测已取得了一定突破,在国防军事、目标监视、灾害救援领域有广泛的应用前景。然而,由于遥感图像具有幅面大、目标小,且成像细节信息模糊等特点,因此,遥感目标检测存在细粒度特征信息提取和融合困难、漏检、误检等问题。目前, Fast-RCNN、Mask-RCNN 等双阶段算法与单阶段网络 YOLO、SSD 等均被广泛研究,在深度学习领域,这些方法在大中型目标检测任务中效果较佳,但在遥感图像目标检测过程中存在一定局限性。

为提升小目标检测精度,Guo 提出了通过高级语义特征图与包含丰富纹理细节的浅层特征图相融合提出 FPN 结构^[1],然而该方法忽略了对小目标的精准定位问题。基于上述工作,Qu 等人提出了一种将空洞卷积与特征融合的算法,以增强

浅层特征图的上下文细节信息^[2],然而,由于感受野的变化,导致目标特征的细微模糊化。Wang 等人通过调节感受野与目标尺度的融合策略解决该问题,但多次尺度变换和网络前向传播增加了计算复杂度^[3]。随后其他小目标检测算法^[4-8]陆续被提出。Pang 等人通过引入注意力模型降低了虚警率,但忽略了复杂背景对遥感目标检测的影响^[9]。为解决遥感目标尺度变化等问题,Zhang 等人将提取出的上下文信息融入空间感知注意力模块,在全局和局部都可以使网络关注信息更丰富的位置^[10],然而,这导致网络复杂度的增加,容易出现过拟合或性能下降的问题。李云红等人采用残差的旁路直连方式缓解网络梯度消失、复杂度高的问题^[11]。另外,Li 等人采用了多尺度的视觉注意力机制,但不同尺度的注意力分布在小目

标检测任务中可能会表现出不同的效果,因此对目标尺度的敏感性可能随着尺度的改变而变化^[12]。李婕等人采用平行层的特征共享结构最大化目标表征,在不同层级中,该结构对目标的感受野和特征表达有所变化^[13],然而,这可能导致特征提取尺度变化较大的小目标造成性能下降。李云红等人提出 2 种特征提取通道,对不变区域和易变区域进行特征提取,提升算法的鲁棒性^[14]。随后,Gong 等人在特征融合权重中引入 FPN 结构进行研究,但在 FPN 结构中,特征图融合过程中可能存在信息冗余的问题,从而导致不必要的计算和存储开销^[15]。虽然这些检测算法在通用场景图像中具有出色的检测能力,但对遥感目标检测效果欠佳。

针对以上遥感小目标细粒度信息不清晰、特征提取困难、特征融合不佳和目标漏检等问题,本文提出一种基于多尺度特征增强与交互融合的遥感目标检测算法。首先,提出 MFE 模块,利用 Split 分流操作丰富和增强不同梯度获取的目标纹理特征信息,同时引入轻量级特征幻影模块 Ghost 进行通道线性变换,生成更多有效的特征细节信息流,以增强对图像中局部细节特征信息的关注,提升对遥感小目标特征提取和表征的能力;其次,提出了 FIF 模块,构建多分支串并行的卷积

块与自适应机制池化块,交互输入特征信息的通道语义和空间变换,目的是捕获上下文信息,精确小目标的关键位置信息,加强特征信息之间的相关性,从而实现细粒度信息的有效特征交互融合。

1 改进遥感图像目标检测算法

算法主要分为 3 个部分,即主干 Backbone、特征融合层 Neck 和检测头 Head,整体网络结构 MFE-FIF 如图 1 所示。Backbone 主要通过卷积神经网络进行下采样提取特征,本文将原来特征信息损失较大的 CSP 结构替换为了 MFE 模块,有效实现了对上下文语义信息和通道特征的关注,增强细节纹理特征的表征能力,在该模块中间层设计了轻量化 Ghost 模块,从原始特征出发,通过线性变换尽可能减少计算量,结合多层级信息提取更加细粒度语义特征,进一步提高模型充分捕获局部感受野内的细节特征,增强对目标中不清晰细节特征的表征和提取。Neck 部分重新设计了 FIF 模块,在提取方式上对正负样本权重分布更加合理,突出目标的显著性,能在串并行分支卷积块和自适应池化块中,通过交互融合的方式融合来自不同尺度的语义信息和位置信息,提高对关键信息的感知能力,减少信息损失。

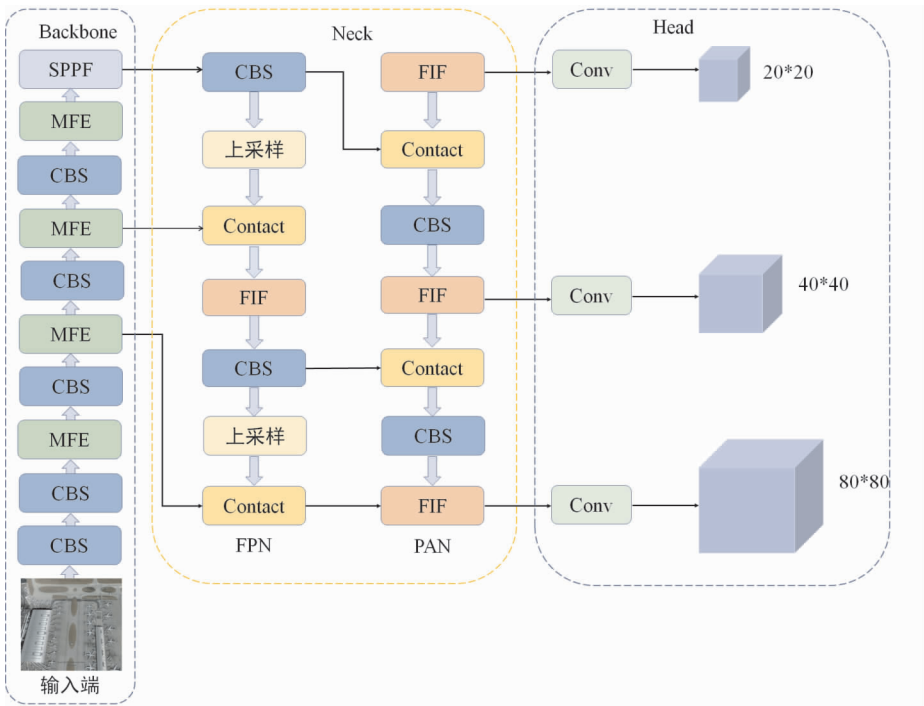


图 1 MFE-FIF 遥感目标检测网络总图
Fig. 1 Overview of MFE-FIF remote sensing target detection network

1.1 多尺度特征增强网络模块 MFE

为解决 Backbone 中多个阶段下采样后图像分辨率降低、目标细节纹理信息损失较大的问题,提出了多尺度特征增强(multiscale feature enhancement, MFE)模块,其详细结构如图 2 所示。

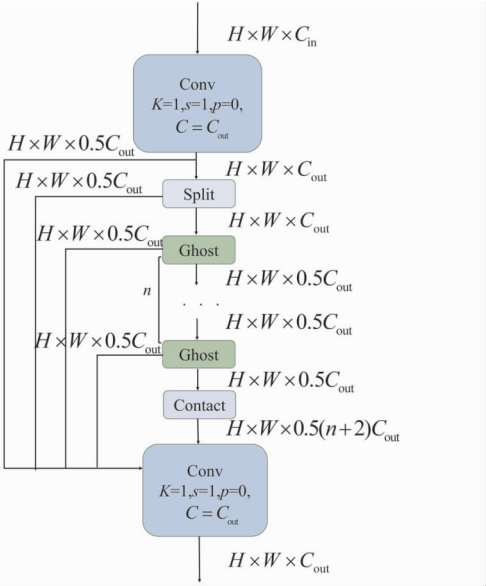


图 2 多尺度特征增强模块 MFE

Fig. 2 Multiscale feature enhancement network structure

MFE 模块能将精确的位置信息和语义信息嵌入到通道中,增强对小目标细节特征的表达能 力,同时更好地处理不同尺寸和低对比度的小目

标,增强对重要细节特征的关注和对图像中关键 信息的提取和表征能力。

设输入遥感图像 X 的高、通道数和宽分别为 H 、 C 、 W , n 代表模块的重复个数, k 为卷积核大小, p 为补丁。为了增强复杂背景下遥感图像小目标的细节信息,并对不同尺度的小目标有更深入的理解,将 X 经过卷积,得到输出 $H \times W \times C_{out}$, 提取到的特征被分为 2 个分支。其中一个分支采用残差连接的方式进行多尺度特征提取,将捕获到的经过卷积的空间信息和纹理信息传送到 Contact 操作层,而另一个分支通过 Split 操作将通道数量减半,并将提取到的细节特征信息送入轻量化 Ghost 模块,进行参数调整和线性变换,并在通道之间共享相似性,以此可将多尺度特征图合并到一个固定尺寸特征图中拼接,经卷积核大小 1×1 , Padding 为 1 的卷积层卷积后输出, 1×1 内核使模型能够捕捉局部跨通道交互特征,增强目标的细节特征。

通过图 2 可以看到在 MFE 模块网络中间层设计的轻量级幻影 Ghost 模块,该模块通过通道分割和混合操作降低了参数量,提高模型在复杂背景下的鲁棒性,能够捕捉目标的空间形态和上下文信息,有助于更精确地定位遥感小目标的位置, Ghost 模块如图 3 所示。

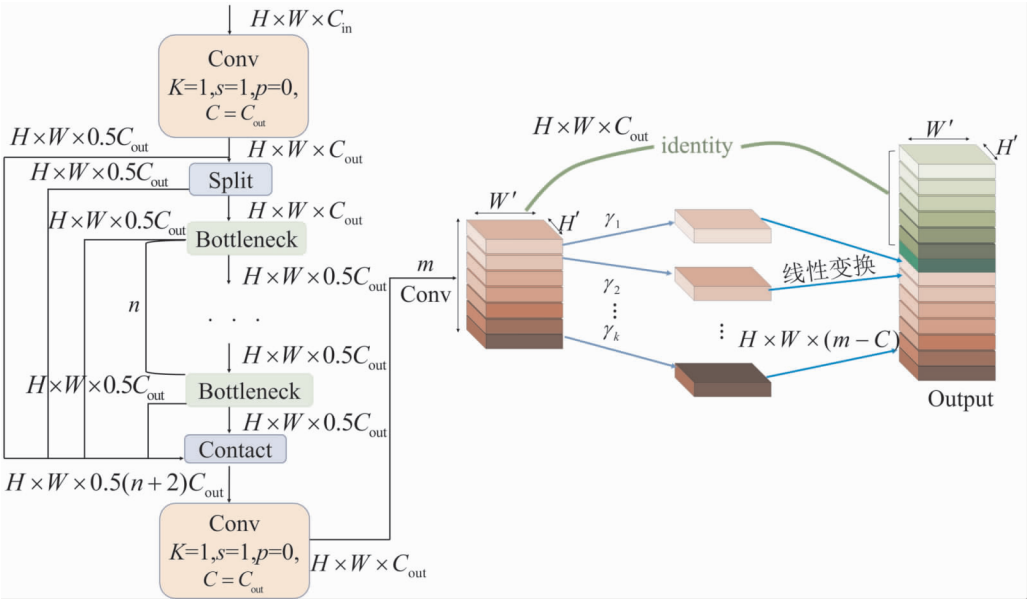


图 3 幻影模块 Ghost

Fig. 3 Ghost network structure diagram

从图3可以看出,通过 Split 分流,拼接跨层多尺度特征得到的丰富的梯度流信息 $h \times w \times c_{out}$ 。将其送入右边通道,首先,通过常规卷积提取本征特征图(intrinsic feature maps) Y' ,聚合通道间的关键细节信息。然后,利用系数 γ 对 Y' 对各通道特征图进行线性变换操作,生成 Ghost 幻影操作后的特征图 $Y_{i,j}$ 。这些特征图包含共同的特征信息,一部分通过卷积生成固有特征,另一部分通过线性运算增加特征和通道得到 Ghost 图,其计算公式如(1)所示。

$$Y_{i,j}^{ghost} = \gamma_j \times Y'_{i,k} + \sum_{m=1, m \neq j}^k \gamma_m \times Y'_{i,m} \quad (1)$$

式中: γ_j 表示对第 j 个通道进行线性变换的系数; $\sum_{m=1, m \neq j}^k \gamma_m \times Y'_{i,m}$ 表示对 Ghost 特征图的第 j 个通道特征映射,将其余的 $k-1$ 个通道进行线性操作后相加。最后,将本征特征图和 Ghost 特征图拼接为 Output 图,在 Ghost 模块中通过线性变换对迭代,以保留固有特征映射。若需输出 n 层,则 c ($c < n$) 层通过卷积得到,剩余 $n-c$ 层经过线性变换生成特征图,计算公式如(2)所示。

$$Y = [Y'_1, Y_1^{ghost}, Y'_2, Y_2^{ghost}, \dots, Y'_c, Y_c^{ghost}, Y_{c+1}^{ghost}, \dots, Y_n^{ghost}] \quad (2)$$

若输入为 $X \in \mathbf{R}^{h_0 \times w_0 \times c_0}$, 输出为 $Y \in \mathbf{R}^{h_1 \times w_1 \times c_1}$, $s \times s$ 为卷积核大小,则普通卷积计算量为 $h_0 \times w_0 \times c_1 \times s \times s \times c_0$ 。假设在 Ghost 模块中的输入为 $X \in \mathbf{R}^{h_1 \times w_1 \times c_1}$, 输出为 $Y \in \mathbf{R}^{h_0 \times w_0 \times c_1}$, 卷积核大小为 $s \times s$, 则卷积计算量表示为 $h_0 \times w_0 \times c_1 \times s \times s$ 。当满足 $c \gg \mathbf{R}$ 时,普通卷积的计算量近似为 Ghost 模块中卷积的 $\mathbf{R}$ 倍。降低网络模型中的计算量,可以有效避免因参数数量过多导致的训练时间过长和过拟合风险,抑制遥感小目标复杂背景的干扰因素,从而提升小目标检测精度。

1.2 特征交互融合模块 FIF

由于原网络 Neck 中的 CSP 结构在捕捉尺寸不一的遥感小目标细节特征方面存在不足,表现为响应不一致、感受野大小灵活性不佳,导致模型对遥感图像中密集小目标细节纹理特征融合不理想,从而影响检测的精度。因此,提出特征交互融合 FIF 模块,其详细网络结构如图4所示。

FIF 模块采用了深度可分离卷积 DSConv 与空间通道重建卷积 SCRConv 结合的卷积块进行特征提取。SCRConv 通过有效地处理特征交互,共享语义信息,实现了信息的丰富流动,且易于嵌

入到现有架构中。同时,能更好地理解目标与环境的关系,从而提高算法对复杂环境的适应性,提高遥感图像小目标检测的准确性。在处理遥感图像中目标子类内部与子类之间微小差异的情况时,该模块将通道间的特征交互与空间上的特征变换分开处理,以更好地融合特征并感知关键细节信息,实现了全局空间信息交互。

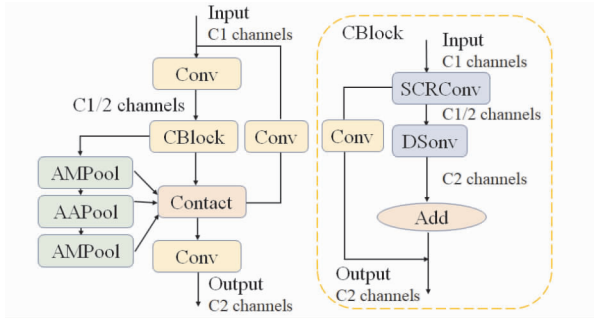


图4 特征交互融合模块 FIF 网络结构图

Fig.4 Network structure diagram of feature interaction fusion module FIF

此外,在此模块中设计了自适应最大池化(adaptive max pool, AMP)与平均池化(adaptive avg pool, AAP)串行的结构,在空间上捕获长距离的相互作用的同时,处理主干网络中的特征映射,减少冗余信息,提高特征表达能力。AMP 能集中注意力在小目标区域,突出目标细节信息的显著性,并有助于精确地定位目标位置,从而提高目标的辨识度。AAP 则可感知整体信息,提供目标区域的平均特征,并平滑背景干扰。

在 FIF 模块中,设输入遥感图像的高度为 H_{in} , 宽度为 W_{in} , 深度为 D_{in} , 输出图像的高度为 H_{out} , 宽度为 W_{out} , 深度为 D_{out} , 道数为 C_1 。一个分支的特征图送入卷积核大小为 5×5 、步长 stride 为 1、边缘填充 padding 为 2 的卷积提取特征后,与多尺度特征层拼接。另一个分支输入到 CBlock(conv block) 卷积块中,计算公式如(3)所示。首先,将输入特征图 X 经过卷积块函数 f_{CBlock} 处理得到 Y_2 ; 然后,对 Y_2 进行自适应最大池化和自适应平均池化操作得到 P_1 和 P_2 , 以增强特征表达的多样性和丰富性,突出小目标的关键特征。池化后的特征图与并行通道拼接实现特征融合,再经卷积核大小为 3×3 、步长 stride 为 1、padding 为 1 的卷积层进行交互特征与信息共享,通过卷积操作 W_{conv2} 得到输出特征图 O , 计算公式如(4)所示。

$$Y_2 = f_{CBlock}(X) \quad (3)$$

$$O = W_{\text{conv2}} \otimes \text{concat} \begin{bmatrix} \text{AMP}(Y_2) \\ \text{AAP}(Y_2) \end{bmatrix} \quad (4)$$

式中:AMP 为自适应最大池化;AAP 为自适应全局池化; $\otimes$ 代表卷积操作。

2 实验过程与结果分析

实验环境配置:CPU 为 E5-2686 v4 @ 2.30 GHz,GPU 为 NVIDIA GeForce RTX 3090 24 GiB 显存,Windows10 操作系统,深度学习框架为 Pytorch2.0.1。为公平比较,实验采用统一的训练参数,输入图像大小为 640×640 pixel,初始学习率设置为 0.01,设置实验训练轮数为 epoch = 300,每个 batch 包含 16 个样本,IOU 阈值分别设置为 0.5 和 0.5 ~ 0.95。

2.1 实验数据

实验采用大型遥感 DIOR 数据集,其涵盖图像 23 463 张和目标实例 192 472 个,有 20 种类别:高速公路服务区(expressway service area)、桥梁(bridge)、网球场(tennis court)、烟囱(chimney)、高速公路收费站(expressway toll station)、飞机(airplane)、体育场(stadium)、棒球场(baseball field)、天桥(overpass)、大坝(dam)、火车站(train station)、海港(harbor)、高尔夫球场(golf field)、船舶(ship)、储罐(storage tank)、篮球场(basketball court)、车辆(vehicle)、地面田径场(ground track field)、机场(airport)、风车(wind mill)。数据集中图像尺寸为 800×800 pixel,空间分辨率范围为 0.5 ~ 30 m,涵盖了不同时间和天气状况下的遥感场景。与其他遥感数据集对比,DIOR 数据集中小目标对象实例类内和类别之间相似度更高,从而也会在遥感小目标检测中面临更大的挑战。因此,选用 DIOR 数据集训练和评估本文所

提出的遥感图像小目标检测算法,其中训练集与测试集分别包括 18 770 张、4 693 张图像。

2.2 评价指标

采用多类别平均精度值(mAP)、准确率(Precision)、召回率(Recall)、参数量作为评价指标。在不同的 IOU 阈值 0.5 和 0.5 ~ 0.95 条件下,通过计算 P-R (precision-recall) 曲线下的面积得到平均精度 AP,阈值的间隔为 0.05。准确率(式中简记 P)为算法预测的结果与实际结果相符的程度,其高低反应了算法的正确率高低。召回率(式中简记 R)指算法能正确检测出实际存在物体的概率,召回率越高,表示算法能更好地检测到实际存在的物体。具体公式如下所示。

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (8)$$

式中:TP(true positive)表示正确识别的小目标实例个数;FP(false positive)指错误标记为正确的小目标样本数量;FN(false negative)代表算法错误预测实际小目标样本为错误样本的个数; N 是小目标总的类别数,在本文中设定其值为 20; AP_i 表示第 i 个目标类别的平均精度值;mAP 是对每个目标类别的 AP_i 求平均值得到。

2.3 消融实验

为验证论文所创新的 2 个网络模块对遥感图像小目标检测过程中性能的影响,在 DIOR 数据集上进行了多组相应的消融实验,包括 YOLOv5s、MFE 模块、FIF 模块、整体改进网络结构后的 YOLOv5s,消融实验结果如表 1 所示。

表 1 本文算法在 DIOR 数据集上的消融实验结果
Tab. 1 Ablation experiment results of our algorithm on DIOR database

模型	mAP50/%	mAP50-95/%	P/%	R/%	模型内存占用量/MB
YOLOv5s	85.6	58.1	88.4	79.3	14.5
YOLOv5s + FIF	86.4	63.0	88.7	82.4	14.3
YOLOv5s + MFE	86.1	62.3	88.6	82.7	12.7
YOLOv5s + FIF + MFE	87.6	63.4	88.6	83.3	127.3

根据表 1 第 2 行,引入特征交互融合 FIF 模块后,mAP50 提高了 0.8%,mAP50-95 提高了 4.9%,参数量降低了 0.2 MB。这是由于将 Neck 中的 CSP 替换为更高效的卷积块和池化层,对不

同通道细节信息交互融合更精准。 P 和 R 分别提高了 0.3%、3.1%。第 3 行中,将 Backbone 中的 CSP 结构替换为多尺度特征增强模块 MFE 模块,且在网络中间层引入轻量化 Ghost 模块后,mAP50 提高了 0.5%,mAP50-95 提高了 4.2%,模型参数量减少了 1.8 MB。本文提出的算法融合了 MFE 和 FIF 模块,在大部分评价指标上都具有优势,mAP50 提高了 2%,mAP50-95 提高了 5.3%, P 与 R 分别提升了 0.2%、4.0%,证明改进的模型整体性能有所提高,有效提升了遥感小目标检测精度。

2.4 对比实验

使用 DIOR 数据集与其他多种主流的经典算法进行对比实验,如 CSFF^[16]、SCRDet^[17]、YOLOv4^[18]、SSD^[19]、CornerNet^[20]、NPMMR-Det^[21]、YOLOv7^[22]

等,验证本文所提改进算法在遥感数据集上的有效性。对比实验结果如表 2 所示,可以看出,本文所提算法与 CornerNet、NPMMR-Det、原模型 YOLOv5 相比,mAP50 分别提高了 22.7%、10.2%、2.0%,且所提方法部分目标类别的 AP 值达到了更佳的检测效果。这得益于本文利用骨干网络丰富的梯度信息流融合,增强遥感小目标的细节特征,提高对其细粒度信息的提取。同时,交互融合特征尽量避免特征信息丢失,利用轻量级模块线性变换减少无效信息冗余,有效改善了遥感图像中小目标的检测效果。实验结果表明,在检测对象过小、目标纹理不清晰时,本文提出的算法仍能有效利用小目标细节信息,在遥感图像数据集的目标检测中具有显著的优势。

表 2 不同算法对比实验结果
Tab. 2 Experimental result comparison of different algorithms

单类 AP 值	网络模型								
	CSFF ^[16]	SCRDet ^[17]	YOLOv4 ^[18]	SSD ^[19]	CornerNet ^[20]	NPMMR-Det ^[21]	YOLOv5	YOLOv7 ^[22]	Ours
Airplane	57.2	80.8	96.0	59.5	58.5	88.1	97.2	98.0	96.7
Airport	79.6	87.7	87.9	72.7	84.2	87.1	87.8	93.1	90.5
Baseball field	70.1	80.6	94.7	72.4	72.0	82.0	96.0	95.7	96.8
Basketball court	87.4	89.8	91.9	75.7	80.8	92.6	93.4	96.6	91.5
Bridge	46.1	57.8	60.0	29.7	46.4	49.6	86.7	66.6	66.6
Chimney	76.6	80.9	90.8	65.8	75.3	80.6	81.1	93.1	92.7
Dam	62.7	75.2	69.9	56.6	64.3	73.4	95.7	85.4	79.0
Expressway service area	82.6	90.0	92.1	63.5	81.6	83.7	88.0	96.9	97.4
Expressway toll station	73.2	82.9	87.5	53.1	76.3	70.5	80.8	90.4	94.0
Golf field	78.2	84.5	87.6	65.3	79.5	85.0	91.2	92.9	83.8
Ground track field	81.6	83.6	83.7	68.6	79.5	84.8	74.1	93.6	92.5
Harbor	50.7	63.2	55.6	49.4	26.1	64.8	72.8	75.5	77.4
Overpass	59.5	67.3	68.7	48.1	60.6	63.3	64.4	76.2	75.0
Ship	73.3	72.6	94.6	59.2	37.6	91.6	96.0	96.3	95.9
Stadium	63.4	79.2	93.8	61.0	70.7	71.3	93.9	97.2	93.6
Storage tank	58.5	70.4	88.4	46.6	45.2	76.1	86.1	90.1	90.5
Tennis court	85.9	90.0	95.7	76.3	84.0	92.0	96.1	97.9	97.5
Train station	61.9	70.7	44.4	55.1	57.1	69.1	70.7	76.4	76.5
Vehicle	42.9	58.8	62.1	27.4	43.0	58.6	67.3	77.4	68.8
Wind mill	86.9	90.3	90.4	65.7	75.9	83.6	92.0	88.7	94.8
mAP50/%	68.0	77.8	81.3	58.6	64.9	77.4	85.6	84.7	87.6

2.5 改进算法效果与分析

为进一步验证本文所提出的改进算法在遥感数据集 DIOR 上的效果,选取测试集中具有挑战

性的遥感图像进行目标检测,并将检测结果对比可视化,结果如图 5 所示。

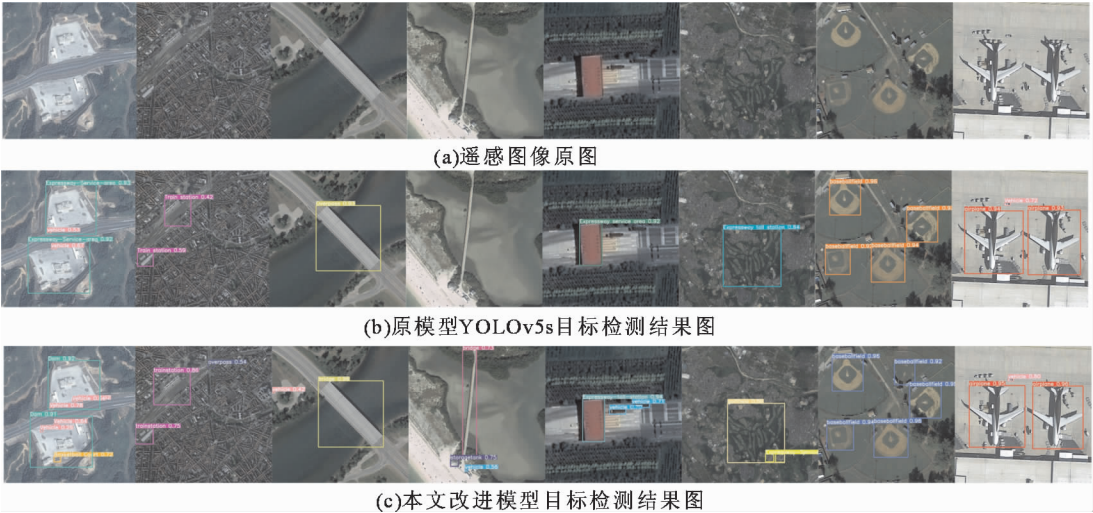


图 5 改进模型与 YOLOv5s 目标检测结果对比图

Fig. 5 Comparison of target detection results between the improved model and YOLOv5s

图 5 展示了数据集 DIOR 在 YOLOv5 原模型和本文改进后算法的检测结果对比,可以看出,本文所提模型增强了对细节纹理方面的检测,检测到的不同类别小目标物体数量与相同小目标物体置信度都有较大提升。图 5(c) 中的第 1 幅图相比原模型的检测结果,可以明显看出检测到的车和篮球场数量增加。对比图 5(b) 和 5(c) 检测结果,第 2 幅图中火车站的置信度分别提升 44% 和 16%,第 3 幅图中原模型未检测到小目标,但本文所提出的方法检测出更多小目标物体且有较高的置信度。从其他结果来看,本文算法对车辆、天桥、飞机、高速公路收费站、棒球场等检测到的目标数量更多、精度值也更高,实现了对遥感图像中小目标细节纹理信息的增强,以及对通道语义、空间信息等多尺度特征融合更佳。综合以上分析,本文所提出的改进算法在处理遥感图像小目标检测任务时,展现出了显著的优势,有助于实际生产与生活中问题的解决。

3 结语

本文设计了一种基于多尺度特征增强与交互融合的遥感目标检测方法。构建模块 MFE 增强小目标的多尺度细节特征,细化空间和渠道维度,引入 Ghost 模块进行线性变换降低模型参数量。同时在 Neck 层,提出特征交互融合模块 FIF,利用浅层语义信息与深层定位特征融合,并采用串并行结合的卷积块与自适应机制的池化块,有效减少上下文语义和空间位置信息的丢失,共享与

融合不同通道层的关键细节信息。通过设置与其他经典网络模型多组对比实验结果,验证了本文所提改进算法在遥感图像小目标检测方面的有效性与较高的精度值,具有较强实用性和应用性。但是,本文方法仍然有一些局限性,例如,尽管本文在所提出的 MFE 模块中引入了轻量级 Ghost 模块,但因网络深度与计算复杂度的增加,相比原模型轻量化效果不明显。将 MFE 和 FIF 模块都引入模型后的参数量并未减少,模型计算复杂度较高,检测速度表现不佳,训练过程耗时。因此,未来的研究工作将重点集中在优化算法方面,包括参数量、检测速度和模型大小,比如数据增强、采用剪枝、稀疏训练、引入多特征检测器等。同时,将结合遥感小目标背景复杂、类间与类内目标尺度变换大等特点,尝试其他模型压缩和加速方法,在保证高效检测小目标的同时,提升检测效果,使算法能够低成本地适用于不同的实际环境与硬件设备,帮助相关领域在处理无人机捕获的场景时获得更好的体验。

参考文献

[1] GUO K. RDSSD: Refinement deconvolutional single shot detector[J]. Journal of Physics: Conference Series, 2020, 1651(1): 012140.

[2] QU J S, SU C, ZHANG Z W, et al. Dilated convolution and feature fusion SSD network for small object detection in remote sensing images[J]. IEEE Access, 2020, 8: 82832-82843.

[3] WANG J L, CUI Z Y, JIANG T, et al. Lightweight deep neural networks for ship target detection in SAR

- imagery[J]. IEEE Transactions on Image Processing, 2022, 32: 565-579.
- [4] WU S X, WANG X H, GUO C C. Application of feature pyramid network and feature fusion single shot multibox detector for real-time prostate capsule detection[J]. Electronics, 2023, 12(4): 1060.
- [5] HAO Z, HONG X G, ZHU L. Detecting small objects in thermal images using single-shot detector[J]. Automatic Control and Computer Sciences, 2021, 55(2): 202-211.
- [6] DENG C F, WANG M M, LIU L, et al. Extended feature pyramid network for small object detection[J]. IEEE Transactions on Multimedia, 2021, 24: 1968-1979.
- [7] YIN Q J, YANG W Z, RAN M Y, et al. FD-SSD: An improved SSD object detection algorithm based on feature fusion and dilated convolution[J]. Signal Processing: Image Communication, 2021, 98: 116402.
- [8] 徐丹青, 吴一全. 光学遥感图像目标检测的深度学习算法研究进展[J]. 遥感学报, 2024, 28(12): 3045-3073.
- XU D Q, WU Y Q. Research progress on deep learning algorithms for object detection in optical remote sensing images[J]. Journal of Remote Sensing, 2024, 28(12): 3045-3073.
- [9] PANG J M, LI C, SHI J P, et al. R2-CNN: Fast tiny object detection in large-scale remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5512-5524.
- [10] ZHANG G J, LU S J, ZHANG W. CAD-Net: A context-aware detection network for objects in remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12): 10015-10024.
- [11] 李云红, 刘杏瑞, 谢蓉蓉, 等. 基于超分辨率重建和公共特征子空间的低分辨率人脸识别[J]. 西北大学学报(自然科学版), 2023, 53(2): 241-247.
- LI Y H, LIU X R, XIE R R, et al. Low resolution facial recognition based on super-resolution reconstruction and common feature subspace[J]. Journal of Northwest University (Natural Science Edition), 2023, 53(2): 241-247.
- [12] ZHANG Y L, YUAN Y, FENG Y C, et al. Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5535-5548.
- [13] 李婕, 周顺, 朱鑫潮, 等. 结合多通道注意力的遥感图像飞机目标检测[J]. 计算机工程与应用, 2022, 58(1): 209-217.
- LI J, ZHOU S, ZHU X C, et al. Aircraft target detection in remote sensing images using multi-channel attention[J]. Computer Engineering and Applications, 2022, 58(1): 209-217.
- [14] 李云红, 聂梦瑄, 苏雪平, 等. 分区域特征提取的人脸识别算法[J]. 西北大学学报(自然科学版), 2020, 50(5): 811-818.
- LI Y H, NIE M X, SU X P, et al. Facial recognition algorithm based on regional feature extraction[J]. Journal of Northwest University (Natural Science Edition), 2020, 50(5): 811-818.
- [15] GONG Y Q, YU X H, DING Y, et al. Effective fusion factor in FPN for tiny object detection[C]//2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2021: 1159-1167.
- [16] CHENG G, SI Y J, HONG H L, et al. Cross-scale feature fusion for object detection in optical remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(3): 431-435.
- [17] YANG X, YANG J R, YAN J C, et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 8231-8240.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2024-09-25]. <https://arxiv.org/abs/2004.10934v1>.
- [19] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[M]//Computer Vision-ECV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [20] LAW H, DENG J. CornerNnet: Detecting objects as paired keypoints[J]. Computer Vision, 2020, 128(3): 642-656.
- [21] HUANG Z C, LI W, XIA X G, et al. A novel nonlocal-aware pyramid and multiscale multitask refinement detector for object detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 5601920.
- [22] 余俊宇, 刘孙俊, 许桃. 融合注意力机制的YOLOv7遥感小目标检测算法研究[J]. 计算机工程与应用, 2023, 59(20): 167-175.
- YU J Y, LIU S J, XU T. Research on YOLOv7 remote sensing small target detection algorithm incorporating attention mechanism[J]. Computer Engineering and Applications, 2023, 59(20): 167-175.