

跨域协同增强的遥感影像小目标检测

张天扬 张向荣* 王冠淳 唐旭

(西安电子科技大学 人工智能学院 西安 710126)

摘要: 面对遥感影像中广泛分布的小目标, 现有遥感影像目标检测方法仍面临严峻挑战, 检测精度尚显不足。遥感影像小目标检测核心瓶颈在于小目标正样本匮乏与特征表示弱两大问题。因此, 本文提出一种跨域协同增强的遥感影像小目标检测方法(Cross-Domain Collaborative Enhancement Detector, CDCEDet), 通过在空间域优化小目标标签分配, 在频域提高小目标特征表示, 以提升小目标检测性能。具体而言, 本文首先设计尺度自适应锚框生成器, 依据目标空间分布与几何尺度特性动态生成匹配锚框集合, 有效缓解了锚框与小目标之间的尺度失配问题, 显著提升了小目标正样本数量。在此基础上, 提出分位数自适应标签分配机制, 通过建模每个目标与其对应锚框的交并比统计分布动态调整标签分配阈值, 有效缓解了固定阈值标签分配下的尺度偏差问题。此外, 针对小目标特征表示弱的问题, 从频域视角出发, 构建频率自适应融合模块, 利用自适应高通滤波器以增强小目标细节信息, 并结合自适应低通滤波器维持上采样后小目标内部特征的语义一致性, 从而提升小目标的特征表示能力。实验结果表明, 在AI-TODv2数据集上, 所提方法在AP₅₀和AP₅₀₋₉₅指标上较现有最优方法提升1.8%和0.7%; 在AI-TOD-R数据集上, 两项指标分别提升2.6%和0.7%, 展现出良好的小目标检测能力和泛化能力。

关键词: 小目标检测; 遥感影像; 标签分配; 特征融合; 频域学习

中图分类号: TP751.2

文献标识码: A

文章编号: 1009-5896(2026)00-0001-11

DOI: 10.11999/JEIT260317

CSTR: 32379.14.JEIT260317

1 引言

遥感影像目标检测旨在识别并定位感兴趣目标, 在国土安全、灾害监测及城市管理等领域具有广泛的应用价值^[1]。作为该领域的重要研究分支, 小目标检测聚焦像素面积小于 16×16 的目标, 受成像特性影响, 这类目标在影像中分布广泛。一方面, 遥感平台获取高度与传感器分辨率共同制约了影像中的目标尺度, 即使是建筑物、飞机等大型目标也常呈现为较小尺度。此外, 车辆、行人等本身面积较小的目标, 在成像过程中被进一步压缩。然而, 现有通用遥感影像目标检测算法在处理小目标时仍面临严重的性能瓶颈, 难以达到常规尺度的检测精度。因此, 突破小目标检测的技术难题, 已成为当前遥感智能解译领域极具价值的研究热点^[2]。

近期研究^[3]指出, 由于交并比(Intersection Over Union, IoU)度量准则对不同尺度目标存在显

著的敏感性偏差。对于小目标, 即使微小位置偏移也会导致IoU大幅下降; 而对于常规尺度目标, 相同幅度的偏移仅产生轻微影响。这种高度敏感性, 使得当前广泛采用的基于IoU度量准则的标签分配方法难以适配小目标, 造成小目标正样本匮乏问题, 已成为制约小目标检测性能的核心瓶颈。为此, 研究学者们提出了高斯概率度量^[4](Gaussian Probabilistic Metric, GPM)和感受野距离^[5](Receptive Field, RF)等代替方案, 并在遥感影像小目标检测任务中展现出良好性能。然而, 现有多数研究主要聚焦于IoU度量准则的尺度敏感性问题, 却忽略了锚框与小目标之间的尺度失配问题。事实上, 这种尺度失配问题同样是制约造成小目标正样本匮乏的关键因素, 尤其在IoU度量准则下, 其负面影响更为显著。如图1所示, 即使锚框与目标的中心位置、宽高比完全一致的前提下, IoU仍随两者尺度比例的减小而急剧下降。这表明, 即使几何形状完美匹配, 尺度失配本身即可导致IoU严重衰减, 从而阻碍小目标的有效检测。针对该问题一种直观的解决方法是生成能够自适应匹配对应目标尺度的锚框。

此外, 由于像素面积有限, 小目标普遍存在视觉外观质量差、特征表示能力弱的问题。特征金字塔网络(Feature Pyramid Network, FPN)^[6]作为经典的多尺度特征融合结构, 通过自顶向下的路径传递丰富语义信息, 有效增强了小目标的特征表示。

收稿日期: 2026-03-18; 改回日期: 2026-04-24; 网络出版: 2026-07-19

*通信作者: 张向荣 xrzhang@mail.xidian.edu.cn

基金项目: 国家自然科学基金(62501433, 62506285, 62571387), 中央高校基础科研业务费(QTZX25070), 陕西省自然科学基金基础研究计划青年项目(2025JC-YBQN-795)

Foundation Items: Natural Science Foundation of China (62501433, 62506285, 62571387), Fundamental Research Funds for the Central Universities under Grant QTZX25070, Natural Science Basic Research Plan in Shaanxi Province of China (2025JC-YBQN-795)

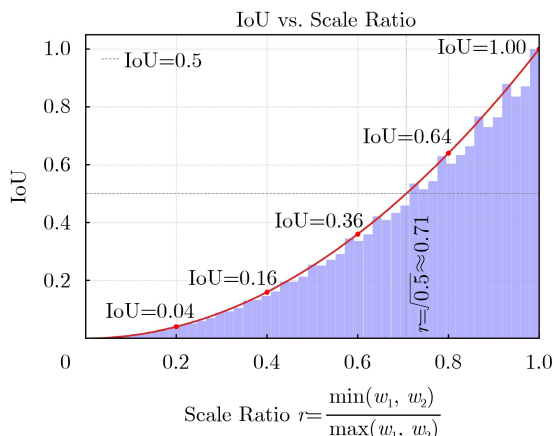


图1 锚框-目标尺度比例 r 与IoU的变化关系

刘等人^[7]提出了重校准特征融合网络,将小目标的浅层边界特征与深层语义特征结合,以更全面地捕捉目标的语义信息。王等人^[8]提出了自适应增强模块,以缓解不同尺度特征之间的直接融合导致的信息冲突。Wu 等人^[9]设计了特征感知对齐模块,通过学习空间变换关系实现多尺度特征融合对齐。尽管上述研究通过优化特征融合策略提升了检测性能,但FPN架构本身在遥感影像小目标检测中仍存在两个主要局限:首先,FPN输出特征的通道维度被压缩至256维,导致大量信息丢失。这对本身特征表示弱的小目标而言,损失更为显著;其次,在自顶向下的特征融合过程中,双线性插值上采样操作忽略了像素间的语义关联性,易造成上采样后目标内部特征语义不一致问题。近期研究表明^[10],图像频域特征对密集预测任务具有重要价值,其中高频分量能够捕捉特征图中的空间跳变。考虑到遥感影像小目标表现为相对于背景的局部脉冲信号,其在频域中具有显著的高频能量分布。

基于上述分析,本文提出了一种跨域协同增强的遥感影像小目标检测方法,在空间域构建动态标签分配机制以扩充小目标高质量正样本,在频域引

入自适应融合模块以增强目标特征表示能力,从而提升小目标检测性能。本文的主要贡献总结如下:

(1)提出了跨域协同增强的遥感影像小目标检测方法,通过在空间域提升小目标正样本数量、在频域增强其特征表示能力,实现了对遥感影像小目标的高精度检测,在两个公开遥感影像小目标检测数据集上的实验结果表明,本文方法表现出最优的检测性能。

(2)提出了尺度自适应锚框生成器,可根据不同目标的空间分布与几何尺度为其生成尺度匹配的锚框集合,有效缓解了固定锚框与小目标之间的尺度失配问题,并显著提升小目标正样本数量。

(3)设计了分位数自适应标签分配机制,通过建模每个目标与其对应锚框之间的IoU统计分布,动态设定标签分配阈值,消除了固定标签分配阈值带来的尺度偏差。

(4)提出了频率自适应融合模块,利用自适应高通滤波器突出小目标细节信息,结合自适应低通滤波器增强上采样后目标内部特征的语义一致性,从频域角度协同提升小目标的特征表示能力。

2 本文方法

2.1 总体框架

图2展示了本文提出的跨域协同增强的遥感影像小目标检测方法的整体架构。该方法以Faster R-CNN^[11]为基准框架,包含三个核心组件:尺度自适应锚框生成器、分位数自适应标签分配机制和频率自适应融合模块。其中,尺度自适应锚框生成器与分位数自适应标签分配机制被集成至候选区域生成网络,通过动态生成尺度匹配的锚框并自适应调整正负样本分配阈值,显著提升小目标正样本数量;频率自适应融合模块则被嵌入到FPN中,从频域角度自适应融合多尺度特征,有效增强小目标的特征表示能力。

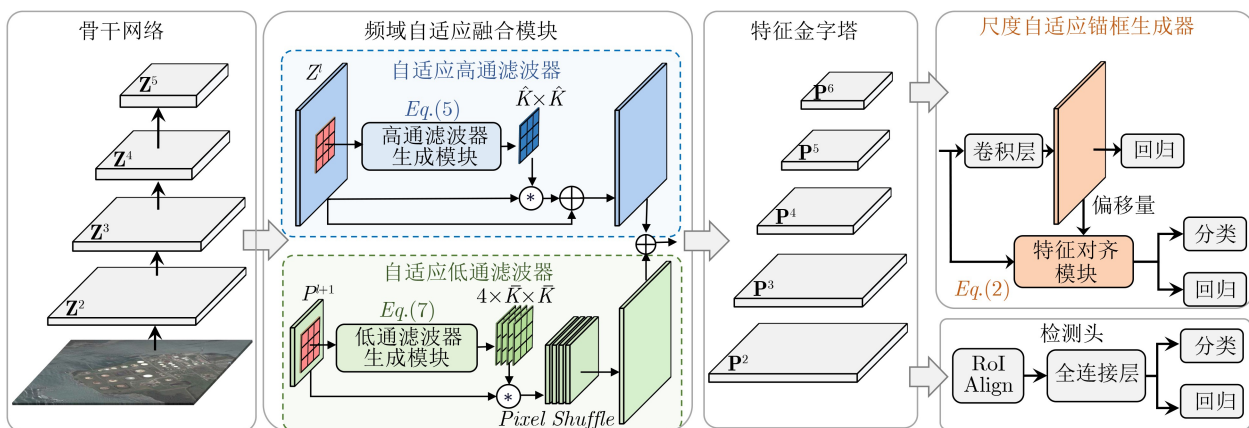


图2 跨域协同增强的遥感影像小目标检测方法整体结构图

2.2 尺度自适应锚框生成器

当前遥感影像小目标检测研究多聚焦于设计替代IoU的度量准则，以缓解其对小目标高敏感性导致的正样本分配不足问题。尽管此类方法取得了性能提升，却普遍忽略了锚框与目标之间的尺度失配问题。本文实验分析表明，这种尺度失配问题也是制约小目标正样本数量的又一个关键因素，因此需要重新审视小目标检测场景下的标签分配机制。

典型的目标检测框架通常采用均匀分布的锚框生成策略：在特征图上以固定步长预设一组宽高比与尺度固定的锚框。然而，这种静态均匀排布方式忽略了真实目标尺度的空间非均匀分布特性，导致大量预设锚框与目标之间出现严重的尺度不匹配现象，加剧了正样本分配不足的问题，尤其在小目标检测任务中更为突出。为此，本文提出尺度自适应锚框生成器，能够为每个真实目标动态生成与其尺度相匹配的锚框，为后续的标签分配奠定基础。

2.2.1 动态锚框生成模块

动态锚框生成模块依据图像中目标尺度特性，在特征图每个空间位置动态生成最优锚框配置。具体而言，对每个位置 p ，预测锚框中心坐标 (a_p^x, a_p^y) 与形状参数 (a_p^w, a_p^h) ，并通过最大化所生成锚框与其最近邻目标之间的IoU，来实现对不同目标的尺度匹配。为确保训练稳定性，本文采用式(1)所示的锚框解码变换，将动态锚框生成模块输出偏移量映射为锚框：

$$\begin{aligned} a_p^x &= p_x + d_x \times \sigma \times s, a_p^y = p_y + d_y \times \sigma \times s \\ a_p^w &= \sigma \times \exp(d_w), a_p^h = \sigma \times \exp(d_h) \end{aligned} \quad (1)$$

其中， d_* 为动态锚框生成模块输出偏移量， (p_x, p_y) 为特征图位置坐标， s 为特征图步长， $\sigma = 8$ 为尺度因子。本文采用IoU损失^[12]对该模块进行优化。与传统的均匀锚框生成策略不同，所提动态锚框生成模块在特征图的每个空间位置自适应地生成锚框，从而有效保证生成锚框与其对应目标的尺度一致性。

2.2.2 特征对齐模块

在传统均匀锚框生成策略下，特征图上所有空间位置均具有统一尺度与宽高比的锚框，有助于特征图学习到一致性的特征表示。然而，提出的动态锚框生成模块在每次训练迭代中动态生成锚框配置，导致所生成锚框与其对应位置特征之间存在空间错位问题。为此，本文设计了特征对齐模块，依据特征图各空间位置生成的锚框参数，通过可变形卷积操作实现相应区域特征的自适应对齐。给定输入特征图 $\mathbf{F}$ ，该模块计算生成锚框与其对应空间位置之间的偏移量，并作为可变形卷积的空间采样指导，其表示公式如下：

$$\tilde{\mathbf{F}}[p] = \sum_k w_k \cdot \mathbf{F}[p + o_p^k] \quad (2)$$

其中， $\tilde{\mathbf{F}}[p]$ 和 $\mathbf{F}[p]$ 分别为特征对齐模块在空间位置 p 的输出和输入特征图， w_k 为 3×3 卷积核参数， $o_p^k = o_p^{ctr} + o_p^{shp,k}$ 为空间位置 p 处的总偏移量，包含中心偏移 $o_p^{ctr} = (a_p^x - p_x, a_p^y - p_y)$ 和形状偏移 $o_p^{shp,k} \in \left\{ \left(-\frac{a_p^w}{2}, -\frac{a_p^h}{2} \right), \left(-\frac{a_p^w}{2}, 0 \right), \dots, \left(0, \frac{a_p^h}{2} \right), \left(\frac{a_p^w}{2}, \frac{a_p^h}{2} \right) \right\}$ 。

2.3 分位数自适应标签分配机制

在目标检测任务中，设置合理的正负样本分配阈值是确保标签分配质量的关键因素。传统目标检测算法通常采用固定IoU阈值，例如，通常将IoU大于0.7的候选框定义为正样本，而将IoU小于0.3的候选框定义为负样本。然而，由于IoU度量准则对不同尺度目标存在敏感性偏差，相同的位置偏移会导致小目标的IoU产生剧烈波动，而对常规尺度目标仅产生微小影响^[3]。这种由尺度差异引起的IoU敏感性偏差，使得固定阈值标签分配机制难以平衡不同尺度目标的匹配需求，在小目标检测任务中尤为突出。针对该问题，本文提出分位数自适应标签分配机制，通过动态分析目标所匹配锚框集合的IoU分布，自适应确定其正负样本分配阈值。

给定目标 G_i 及锚框集合 $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$ ，首先计算每个锚框 a_j 与目标 G_i 之间的IoU，并构建IoU分布 $\mathbf{D}_i = \{\text{IoU}_{i1}, \text{IoU}_{i2}, \dots, \text{IoU}_{in}\}$ 。随后从该分布中选取特定分位数 α ，作为该目标的正负样本分配阈值 T_i ：

$$T_i = \text{Quantile}(\{\text{IoU}_{ij} \in \mathbf{D}_i \mid \text{IoU}_{ij} \geq \theta\}, \alpha) \quad (3)$$

其中， θ 表示IoU过滤阈值，用于剔除低质量匹配锚框。对于锚框 a_j ，若其与真实目标 G_i 的IoU大于等于分位数阈值，则将其视为正样本；若IoU低于分位数阈值，则视为负样本；其余锚框视为忽略样本，不参与训练。

此外，考虑到模型在训练过程中学习状态的动态演变特性，本文进一步设计了动态分位数阈值机制。在训练初期，由于生成锚框质量尚未优化，匹配质量较低，此时采用较高的分位数阈值以过滤低质量匹配，仅保留高质量锚框；随着训练逐步收敛，所提尺度自适应锚框生成器能够生成更多与目标尺度匹配的高质量锚框，此时通过降低分位数阈值，可为目标分配更多数量的正样本，促进模型对目标的充分学习。所提动态分位数阈值可表示为：

$$a_t = a_0 - a_1 \cdot \frac{t}{T} \quad (4)$$

其中， $a_0 = 0.85$ 和 $a_1 = 0.1$ 分别表示初始分位数阈值和其变化范围， t 表示当前训练轮次， T 表示总训

训练轮次。通过动态分位数阈值机制,使得模型在训练初期能够严格筛选高质量锚框以抑制噪声干扰;而在训练后期,则适度放宽分位数阈值,增加小目标的正样本数量。

2.4 频率自适应融合模块

多尺度特征融合是增强小目标特征表示的有效手段,FPN作为最具代表性的方法,通过自顶向下的路径将高层语义信息传递至低层,以丰富低层特征的语义表示。然而,FPN在遥感影像小目标检测中存在两个主要局限:一方面,通道压缩操作易导致小目标的关键细节信息丢失;另一方面,双线性插值操作可能在目标上采样过程中引入语义噪声干扰。针对上述问题,本文从频域视角重新审视多尺度特征融合,提出频率自适应融合模块,其主要包括自适应高通滤波器和自适应低通滤波器。

2.4.1 自适应高通滤波器

FPN通过卷积操作将骨干网络提取的高维特征统一映射至256维,这种维度约简不可避免地造成信息丢失,对特征表示能力弱的小目标产生的负面影响尤为显著。图像的高频分量能够敏锐捕捉特征图中空间跳变,与小目标在空间域中表现出的局部脉冲特性高度契合。受此启发,本文设计了自适应高通滤波器,通过动态强化小目标的高频分量,有效弥补了FPN中通道压缩导致的信息损失。

具体而言,自适应高通滤波器以骨干网络的输出特征图 $\mathbf{Z}^l, l \in \{2, 3, 4\}$ 为输入,首先通过一个 3×3 卷积层来生成初始滤波器核参数 $\hat{\mathbf{V}}^l$ 。随后,利用Softmax操作将初始滤波器核参数转化为归一化的权重分布,构成低通滤波器^[13]。最后,通过从单位核 $\mathbf{E}$ 中减去该低通分量,得到对应的高通滤波器核。整个过程可形式化表示为:

$$\begin{aligned} \hat{\mathbf{V}}^l &= \text{Conv}_{3 \times 3}(\mathbf{Z}^l), \\ \hat{\mathbf{W}}_{i,j}^{l,p,q} &= \mathbf{E} - \text{Softmax}(\hat{\mathbf{V}}_{i,j}^l) \\ &= \mathbf{E}^{p,q} - \frac{\exp(\hat{\mathbf{V}}_{i,j}^{l,p,q})}{\sum_{p,q \in \Omega} \exp(\hat{\mathbf{V}}_{i,j}^{l,p,q})} \end{aligned} \quad (5)$$

其中, $\hat{\mathbf{V}}^l \in R^{\hat{K}^2 \times H \times W}$ 为初始滤波器核参数, $\hat{K}$ 为高通滤波器的核尺寸。 $\hat{\mathbf{W}}_{i,j}^{l,p,q}$ 为位置 (i, j) 处的高通滤波器核参数, $\mathbf{E}$ 为单位核, $\Omega = \hat{K} \times \hat{K}$ 表示高通滤波器所覆盖区域。基于上述自适应高通滤波器,本文进一步引入残差连接机制,将原始特征与滤波后的高频分量相加,生成高频信息增强的特征图 $\tilde{\mathbf{Z}}^l$:

$$\tilde{\mathbf{Z}}_{i,j}^l = \mathbf{Z}_{i,j}^l + \sum_{p,q \in \Omega} \hat{\mathbf{W}}_{i,j}^{l,p,q} \cdot \mathbf{Z}_{i+p,j+q}^l \quad (6)$$

2.4.2 自适应低通滤波器

FPN中双线性插值上采样操作仅依赖空间位置关系而忽略了相邻像素间的语义关联,容易在上采样过程中引入语义噪声,导致上采样后目标内部区域语义不一致问题。依据频域分析原理,特征图中的低频分量表征了全局语义结构与空间一致性信息。因此,本文在特征图上采样过程中引入自适应低通滤波器,自适应地保留高层级特征图中占主导的低频成分,并抑制上采样过程中引入的高频噪声干扰,从而提升上采样后特征的语义一致性。

所提自适应低通滤波器由一个 3×3 卷积层和Softmax操作构成,其中,卷积层输出通道数为 $u^2 \bar{K}^2$, $\bar{K}$ 表示低通滤波器的核尺寸, $u = 2$ 为上采样因子。自适应低通滤波器以高层级特征图 $\mathbf{P}^{l+1}, l \in \{2, 3, 4\}$ 为输入,首先通过 3×3 卷积层获取初始化滤波器核参数。随后,在空间和通道维度上对其重塑,得到四组滤波器核参数 $\bar{\mathbf{V}}^l \in R^{4 \times \bar{K}^2 \times H \times W}$ 。最后,对每组滤波器核参数执行Softmax操作,确保其满足非负性约束和归一化条件,从而得到最终的低通滤波器。自适应低通滤波器的生成过程可形式化表示为:

$$\begin{aligned} \bar{\mathbf{V}}^{l+1} &= \text{Conv}_{3 \times 3}(\mathbf{P}^{l+1}) \\ \bar{\mathbf{W}}_{i,j}^{l+1,g,p,q} &= \text{Softmax}(\bar{\mathbf{V}}_{i,j}^{l+1,g}) \\ &= \frac{\exp(\bar{\mathbf{V}}_{i,j}^{l+1,g,p,q})}{\sum_{p,q \in \Omega} \exp(\bar{\mathbf{V}}_{i,j}^{l+1,g,p,q})} \end{aligned} \quad (7)$$

其中, $\bar{\mathbf{W}}_{i,j}^{l+1,g,p,q}$ 代表位置 (i, j) 处第 g 组低通滤波器核参数, $\Omega = \bar{K} \times \bar{K}$ 表示低通滤波器所覆盖区域。

在获得低通滤波器后,本文将每组低通滤波器 $\bar{\mathbf{W}}^{l+1,g}$, 以空间自适应加权求和的方式作用于高层级特征图 $\mathbf{P}^{l+1}$, 以增强其低频响应。随后,利用像素重组(Pixel Shuffle)操作对加权后的特征图进行重组,实现2倍上采样。上述过程可表示为:

$$\begin{aligned} \tilde{\mathbf{P}}_{i,j}^{l+1,g} &= \sum_{p,q \in \Omega} \bar{\mathbf{W}}_{i,j}^{l+1,g,p,q} \cdot \mathbf{P}_{i+p,j+q}^{l+1}, \\ \tilde{\mathbf{P}}^{l+1} &= \text{PixelShuffle}(\tilde{\mathbf{P}}^{l+1,1}, \tilde{\mathbf{P}}^{l+1,2}, \tilde{\mathbf{P}}^{l+1,3}, \tilde{\mathbf{P}}^{l+1,4}) \end{aligned} \quad (8)$$

所提FAF模块仍遵循FPN的多尺度特征融合方式^[6],通过逐元素相加将增强后的特征图 $\tilde{\mathbf{Z}}^l$ 与上采样特征图 $\tilde{\mathbf{P}}^{l+1}$ 融合,生成具有丰富语义与空间细节的多尺度特征金字塔。由于小目标普遍存在于低层级特征,因此FAF模块应用于 $l \in \{2, 3, 4\}$ 层级,对于 $\mathbf{P}^5$ 和 $\mathbf{P}^6$ 层级特征图,仍沿用原有生成方式。

3 实验结果与分析

3.1 数据集

(1)AI-TODv2数据集^[3]：该数据集是针对遥感影像小目标检测设计的基准数据集，共包含 28,036 张遥感影像和 752,745 个标注目标。数据集目标平均尺度仅为12.7像素，其中超过85%的目标面积小于 16×16 像素。整个数据集按40%、10%和50%的比例划分为训练集、验证集和测试集。在本文实验中，均使用训练集和验证集进行网络训练，并在测试集上评估模型性能。

(2)AI-TOD-R数据集^[14]：该数据集是基于AI-TODv2数据集扩展构建，旨在支持遥感影像旋转小目标检测任务。目标的旋转角度定义为其最小外接矩形边界框与水平轴方向之间的夹角，取值范围限定于 $(0^\circ, 90^\circ]$ 。数据集中目标的平均尺度仅为10.6像素，对小目标检测算法提出了更严峻挑战。

3.2 实验细节和评价指标

本文采用在ImageNet数据集上预训练的ResNet-50^[15]作为骨干网络初始化权重，其余网络参数采用随机初始化。训练阶段使用带动量的随机梯度下降(Stochastic Gradient Descent, SGD)优化算法，批训练大小为2，共训练12个轮次。初始学习率设为0.005，在第8和第10轮次后分别以0.1倍衰减策略调整学习率。数据预处理方面，所有输入遥感影像统一缩放至 800×800 像素，且仅采用随机翻转作为数据增强手段。推理阶段，类别置信度阈值设为0.05，针对水平边界框和旋转边界框，非极大值抑制(Non-maximum Suppression NMS)的IoU 阈值分别设置为0.5和0.1。

评价指标方面，本文以平均精度(Average Precision, AP)为核心指标，从整体目标检测性能和各尺度目标检测性能进行评价，具体包括七项指标： AP_{50-95} 、 AP_{50} 、 AP_{75} 、 AP_{vt} 、 AP_t 、 AP_s 和 AP_m 。其中， AP_{50-95} 定义为IoU 阈值在 $[0.5, 0.95]$ 区间内以步长0.05采样所得10个AP值的平均； AP_{50} 和 AP_{75} 分别对应IoU阈值为0.5和0.75时的AP值，用于评估算法模型在不同定位精度要求下的整体表现； AP_{vt} 、 AP_t 、 AP_s 和 AP_m 进一步细化尺度维度评价，将目标按像素面积划分为四个尺度^[3]：极小目标(very tiny) $\in [2^2, 8^2]$ ；微小目标(tiny) $\in (8^2, 16^2]$ ；小目标(small) $\in (16^2, 32^2]$ ；中等目标(medium) $\in (32^2, 64^2]$ 。此外，选取参数量、计算量(GFLOPs)和每秒帧数(FPS)作为模型复杂度评价指标，参数量反映模型的存储规模；GFLOPs表示模型单次前向推理所需的浮点运算量；FPS用于评估模型的推理速度。

3.3 对比实验结果

3.3.1 AI-TODv2数据集实验结果

为验证所提方法的有效性，本文选取RFLA^[5]、MENet^[17]、PGDP^[20]等多个代表性遥感影像小目标检测算法进行对比。针对部分方法未在AI-TODv2数据集上报告官方结果，本文基于其开源代码进行了复现。表1展示了详细的对比实验结果。可以看出，本文方法在整体检测精度 AP_{50-95} 上达到24.6%，较近期提出的DCEDet^[18]、DCNet^[19]和PGDP^[20]，分别取得了1.1%、1.0%和0.7%的性能提升。尤其是在极小目标检测精度

AP_{vt} 上，本文方法较对比算法平均提升超过2%，充分证明了本文方法在遥感影像小目标检测任务中的有效性。由于IoU对小目标位置偏差高度敏感，其在小目标检测中存在固有局限，因此，部分研究学者提出基于分布度量的替代准则。然而，本文方法在采用IoU度量准则下，整体检测精度 AP_{50-95} 较NWD^[3]、ADAS-GPM^[3]和RFLA^[4]分别提升3.2%、2.3%和1.7%。这表明，在合理的模型设计下，IoU度量准则在遥感影像小目标检测任务中仍具备有效性与应用潜力。此外，现有小目标检测方法虽在小尺度目标检测性能(AP_{vt} 和 AP_t)上取得显著增益，但往往以牺牲中等目标检测精度(AP_m)为代价。如表1所示，LTDNet^[16]和MENet^[17]相较于与基线方法Faster R-CNN，其中尺度目标检测精度 AP_m 上分别下降了2.7%和3.6%，这种性能失衡现象制约了模型在实际复杂场景中的应用。相比之下，本文方法在所有尺度目标检测性能上展现出优异的均衡性，这主要得益于所提尺度自适应锚框生成器能够动态生成与不同目标尺度相匹配的锚框集合，从而促进模型对各尺度目标的均衡学习。

表 1 AI-TODv2数据集上不同算法的对比实验结果

方法	AP_{50-95}	AP_{50}	AP_{75}	AP_{vt}	AP_t	AP_s	AP_m
Faster R-CNN* ^[11]	12.8	29.9	9.4	0.0	9.2	24.6	37.0
FSANet* ^[9]	17.6	45.0	10.5	5.4	15.8	22.9	33.8
LTDNet ^[16]	18.9	46.7	12.1	5.3	17.7	24.2	34.3
NWD ^[3]	21.4	53.2	12.5	7.7	20.7	26.8	35.2
MENet* ^[17]	21.5	52.7	13.6	8.4	21.6	25.4	33.4
ADAS-GPM ^[3]	22.3	53.7	13.5	7.1	21.9	27.5	35.1
RFLA* ^[4]	22.9	55.7	14.7	8.5	22.4	28.7	36.4
DCEDet ^[18]	23.5	53.9	16.8	8.5	24.1	28.2	37.1
DCNet ^[19]	23.6	56.3	15.7	7.6	23.6	28.6	36.4
PGDP ^[20]	23.9	55.8	16.6	7.3	23.5	29.0	-
CDCEDet(本文方法)	24.6	58.1	16.4	10.5	24.3	29.4	38.1

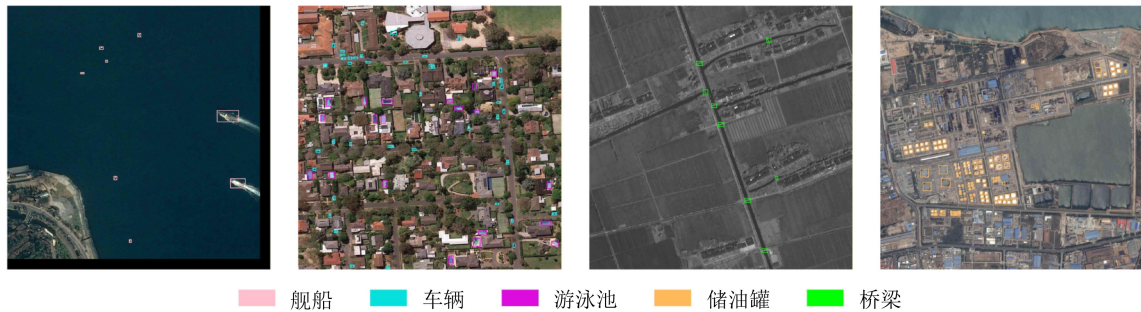
注：*表示基于官方代码复现结果，粗体表示最优结果，下划线表示次优结果

图3(a)展示了本文方法在AI-TODv2数据集上的可视化检测结果。定性分析表明,无论在目标密集和稀疏分布场景,本文方法均能准确、稳定地检测出小目标,验证了其在复杂遥感场景中的鲁棒性。

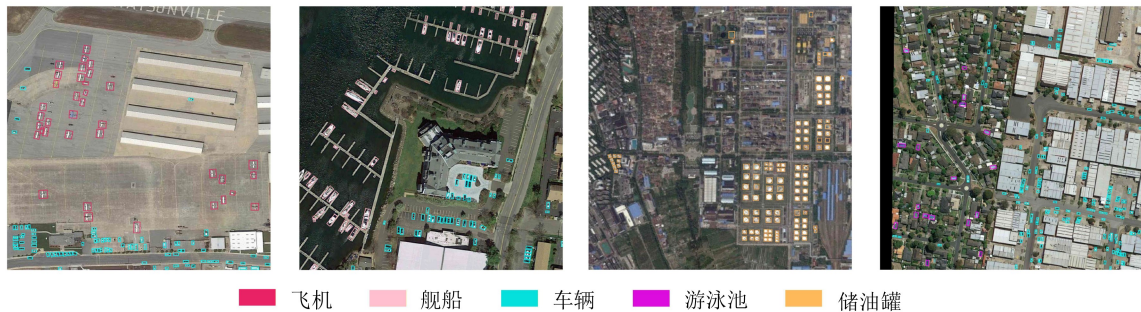
3.3.2 AI-TOD-R数据集实验结果

为进一步验证所提方法在遥感影像旋转小目标检测场景下的泛化能力,本文在AI-TOD-R数据集上与多种代表性方法进行了对比,实验结果如表2所示。由于AI-TOD-R数据集中的目标采用旋转标注格式,本文以RoI-Transformer^[21]作为基线模型来集成所提出各个模块进行对比。由表2第1行(基线方法)与第9行(本文方法)结果可知,本文方法在整体检测精度 AP_{50-95} 上达到最优的16.7%,并在各

尺度目标检测指标上相较于基线模型取得显著提升,其中,极小目标检测精度 AP_{vt} 提升5.5%,微小目标检测精度 AP_t 提升5.1%,小目标检测精度 AP_s 提升1.9%,中等目标检测精度 AP_m 提升0.8%。与典型的旋转目标检测方法Oriented R-CNN^[22]、ReDet^[23]和ARS-DETR^[25]相比,本文方法在整体检测精度 AP_{50-95} 上分别提升5.5%、5.1%和2.4%。与最近提出的遥感影像旋转小目标检测方法DCFL^[13]相比,本文方法在整体检测精度 AP_{50-95} 上取得了1.0%的性能提升,且在各尺度检测精度实现了全面的性能超越。此外,本文也将NWD^[3]和RFLA^[5]扩展至旋转小目标检测场景作为对比方法,实验结果显示,本文方法在整体检测精度 AP_{50-95} 上分别较



(a) AI-TODv2数据集上可视化检测结果



(b) AI-TOD-R数据集上可视化检测结果

图3 本文方法在AI-TODv2和AI-TOD-R数据集上可视化检测结果

表2 AI-TOD-R数据集上不同算法的对比实验结果

方法	AP_{50-95}	AP_{50}	AP_{75}	AP_{vt}	AP_t	AP_s	AP_m
RoI Transformer ^{*[21]}	12.3	35.3	5.5	1.3	10.6	19.7	24.5
Oriented R-CNN ^[22]	11.2	33.2	4.3	0.6	9.1	19.5	23.2
ReDet ^[23]	11.6	32.8	4.8	1.4	9.5	19.4	23.2
Oriented RepPoints ^[24]	13.0	40.3	4.2	5.2	12.2	16.8	21.4
ARS-DETR ^[25]	14.3	41.1	5.8	6.3	14.5	17.6	18.7
NWD ^{*[3]}	15.4	46.9	5.4	5.9	14.2	20.4	24.4
RFLA ^{*[4]}	16.0	48.2	5.6	5.9	14.8	20.1	24.1
DCFL ^[13]	15.7	47.0	5.8	6.3	14.8	19.6	22.4
CDCEDet(本文方法)	16.7	50.8	5.7	6.8	15.7	21.6	25.3

注: *表示基于官方代码复现结果, 粗体表示最优结果, 下划线表示次优结果

NWD^[3]和RFLA^[5]提升了 1.3%和 0.7%，同时在 AP_{50} 指标上取得了 3.9% 和 2.6% 的性能增益。上述实验结果充分证明了本文方法在更具挑战性的遥感影像旋转小目标检测任务中的有效性和泛化性。图3(b)可视化了本文方法在AI-TOD-R数据集上的检测结果，其在复杂背景与目标多方向分布下仍能实现高精度的小目标检测。

3.4 消融实验

本节在AI-TODv2数据集上对所提各个模块开展系统性消融实验分析，以全面评估其有效性。

3.4.1 尺度自适应锚框生成器有效性

针对均匀分布锚框与目标间的尺度失配问题，本文提出尺度自适应锚框生成器(Scale Adaptive Anchor Generator, SAAG)，可为每个目标动态生成与其尺度相匹配的锚框集合。如表3第2行所示，相较于均匀分布锚框生成策略，所提SAAG在整体检测精度 AP_{50-95} 上提升10.0%，且在极小目标和微小目标的检测性能增益尤为显著， AP_{vt} 和 AP_t 指标分别提升7.4%和13.9%。图4直观对比了均匀分布锚框与所提SAAG生成锚框在目标正样本分配上的差异。可以看出，SAAG生成的锚框在尺度分布上与目标高度一致，并能够为小目标分配

更多正样本锚框，从而有效缓解因尺度失配导致的小目标正样本分配不足问题。

3.4.2 分位数自适应标签分配机制有效性

针对固定IoU阈值标签分配机制存在的尺度偏差问题，本文提出分位数自适应标签分配机制(Quantile-based Adaptive Label Assignment, QALA)，通过分析每个目标与其匹配锚框间IoU的统计分布，为不同尺度目标动态生成专属标签分配阈值。对比表3第2行与第3行实验结果可知，将固定IoU阈值标签分配机制替换为所提QALA机制，模型在极小目标检测精度 AP_{vt} 和微小目标检测精度 AP_t 上分别提升1.9%和0.8%，验证了该机制对小目标正样本分配的有效增强。此外，表4展示了QALA机制中初始分位数阈值参数 a_0 的敏感性分析结果。实验分析表明，当处于0.75至0.95区间时，模型表现出良好的鲁棒性，整体检测精度 AP_{50-95} 维

持在23.0%左右，并在 $a_0 = 0.85$ 时达到最优值23.2%。然而，随着逐渐减小(如 $a_0 = 0.1$)，放宽的阈值降低了正样本的筛选标准，导致大量背景区域被误分为正样本，使得模型性能显著下降，甚至低于未引入QALA机制的基线模型。因此，本文选取 $a_0 = 0.85$ 为初始分位数阈值。

此外，考虑到所提SAAG在训练阶段会逐步提升锚框质量，本文在QALA机制中引入了时序动态调整策略。表5展示了分位数变化范围参数的敏感性分析结果。实验结果表明，适当的时序动态调整策略可有效增强检测性能。在 $a_1 = 0.1$ 时，整体检测精度 AP_{50-95} 提升至23.6%，且在不同目标尺度上均表现出性能增益，验证了该策略的有效性。然而，过大的分位数变化范围会引起严重的阈值波动，导致背景噪声被误分类为正样本并降低检测精度。因此，本文设置分位数变化范围 $a_1 = 0.1$ 以实现最佳检测性能。

3.4.3 频率自适应融合模块有效性

为增强小目标的特征表示能力，本文提出频率自适应融合(Frequency Adaptive Fusion, FAF)模块。如表6最后两行的实验结果所示，在不同模型配置下引入FAF模块后，检测性能均呈现一致的提升趋势，充分验证了所提FAF模块在小目标检测任务中的有效性。此外，进一步分析了FAF模块在不同参数配置下的性能表现。实验结果表明，当自适应高通滤波器核尺寸 $\hat{K} = 3$ ，自适应低通滤波器核尺寸 $\bar{K} = 5$ 时，模型达到最优检测性能平衡。这主要是因为高通滤波器核尺寸过大会引入额外噪声，而低通滤波器核尺寸过小则会限制语义信息的完整性，二者均不利于小目标特征表示。

3.4.4 整体框架有效性分析

将所提出的SAAG、QALA和FAF同时集成至基准模型，构成了本文提出的跨域协同增强的遥感影像小目标检测方法。如表3最后一行实验结果所示，本文方法在所有评估指标上均取得了最优性能，验证了所提各模块之间的协同互补性。此外，图5展示了本文方法与基线模型、RFLA^[5]和DCNet^[19]在AI-TODv2数据集上的检测结果可视化

表 3 本文提出各个模块的消融实验结果

SAAG	QALA	FAF	AP_{50-95}	AP_{50}	AP_{75}	AP_{vt}	AP_t	AP_s	AP_m	参数量	GFLOPs	FPS
			12.6	29.5	8.8	0.0	8.4	25.6	37.1	41.16M	144.09	54.8
✓			22.6	53.9	15.3	7.4	22.3	28.1	37.7	41.75M	144.18	45.0
✓	✓		23.6	56.1	15.8	9.3	23.1	29.0	37.0	41.75M	144.18	44.6
		✓	14.0	32.5	9.8	0.0	10.6	26.4	37.8	47.31M	155.34	48.4
✓	✓	✓	24.6	58.1	16.4	10.5	24.3	29.4	38.1	47.90M	155.43	40.5

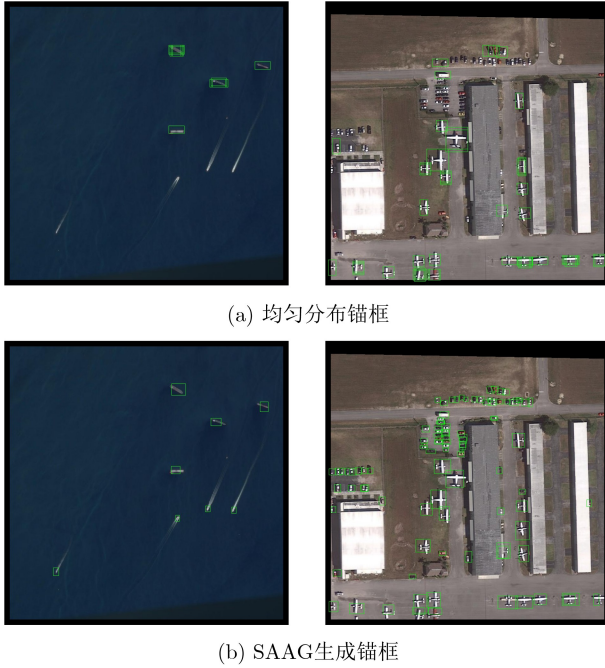


图 4 SAAG与均匀分布锚框的目标正样本分配可视化对比

表 4 QALA机制中初始分位数阈值参数 a_0 敏感性分析

a_0	AP_{50-95}	AP_{vt}	AP_t	AP_s	AP_m
-	22.6	7.4	22.3	28.1	37.7
0.95	23.0	8.9	22.9	28.3	36.9
0.85	23.2	8.9	22.6	28.7	36.7
0.75	23.1	8.8	22.5	28.4	36.9
0.5	22.8	8.3	22.6	28.0	36.3
0.3	22.5	7.3	22.5	28.1	35.9
0.1	22.1	6.9	22.1	27.2	35.4

对比, 其中绿色边界框和红色边界框表示正确检测和漏检目标。可以看出, 基线模型存在严重的漏检现象, 反映出其在小目标检测任务中的局限性。近期提出的代表性方法RFLA^[5]和DCNet^[19]虽在一定程度上改善了检测性能, 但仍存在明显的漏检问题。相比之下, 本文方法显著提升了小目标的召回率, 展现出更强的小目标检测能力与鲁棒性。

3.4.5 模型复杂度分析

表3最后三列对所提方法的模型复杂度进行了分析。SAAG通过引入额外的卷积层进行锚框偏移量预测, 增加了0.59M的参数量; 同时, 特征对齐

表 5 QALA机制中分位数变化范围参数 a_1 敏感性分析

a_1	AP_{50-95}	AP_{vt}	AP_t	AP_s	AP_m
0	23.2	8.9	22.6	28.7	36.7
0.1	23.6	9.3	23.1	29.0	37.0
0.3	23.0	8.4	22.9	27.9	36.6
0.5	22.8	8.3	22.8	28.8	36.0
0.7	22.8	8.5	22.4	28.4	35.6

表 6 不同滤波器核尺寸配置下的FAF模块消融实验结果

滤波器核尺寸		AP_{50-95}	AP_{vt}	AP_t	AP_s	AP_m
高通	低通					
$\hat{K} = 3$	$\bar{K} = 3$	24.3	9.1	24.1	29.2	37.6
$\hat{K} = 5$	$\bar{K} = 3$	24.4	9.5	24.6	29.1	37.7
$\hat{K} = 5$	$\bar{K} = 5$	24.2	9.0	24.6	28.7	37.3
$\hat{K} = 3$	$\bar{K} = 5$	24.6	10.5	24.3	29.4	38.1

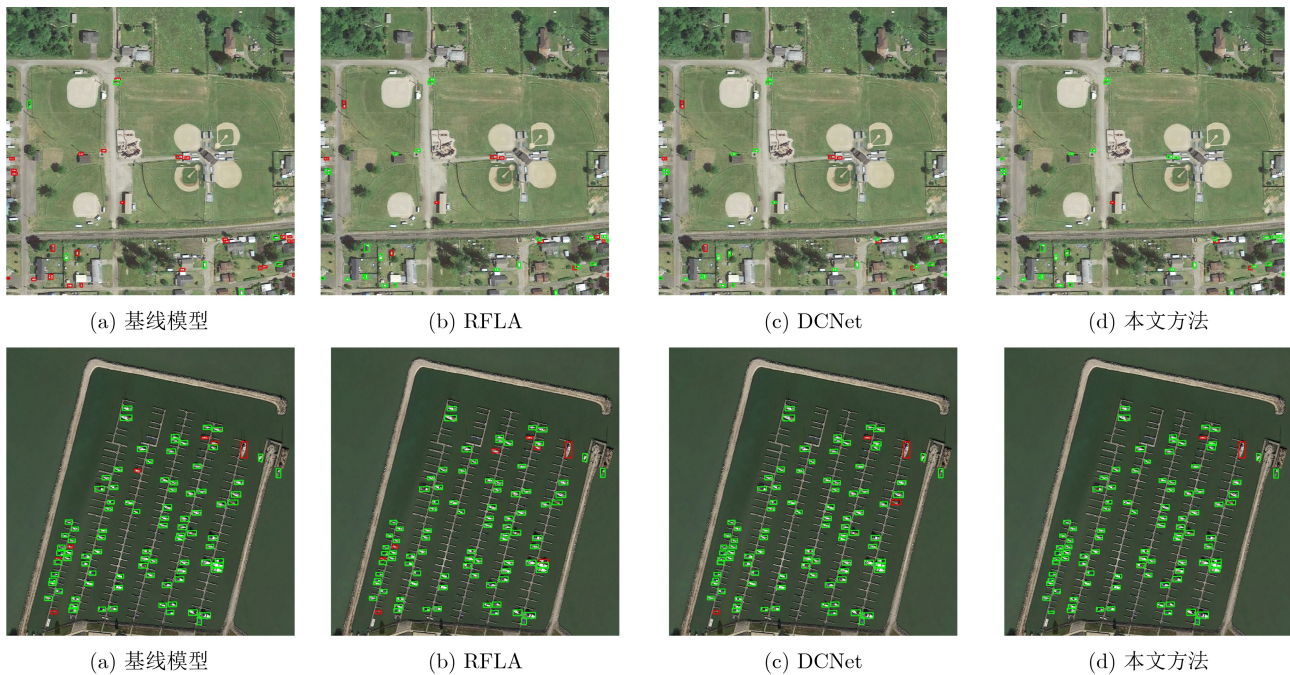


图 5 本文方法与其他方法的检测结果可视化对比

过程中采用了可变形卷积操作,对速度推理产生了一定影响,导致FPS下降9.8。QALA机制仅在训练阶段引入,在推理阶段可直接丢弃,因此不会增加模型复杂度。所提出的FAF模块由于作用在多个特征融合阶段,导致模型参数量增加6.15M,运算量增加11.25GFLOPs。综合分析表明,相较于基线模型,本文方法参数量增加6.74M,运算量增加11.34GFLOPs, FPS由54.8下降至40.5,但在小目标检测精度获得了显著提升,其中 AP_{vt} 由0.0%提升至10.5%, AP_t 由8.4%提升至24.3%。

5 结论

本文提出了一种跨域协同增强的小目标检测方法,从空间域标签分配与频域特征表示两个关键维度协同优化检测性能。在标签分配方面,提出尺度自适应锚框生成器,依据目标空间分布与几何尺度特性动态生成匹配的锚框集合,显著提升小目标正样本数量;在此基础上,进一步设计分位数自适应标签分配机制,通过建模目标与匹配锚框间IoU的统计分布,动态调整标签分配阈值,有效缓解了固定阈值分配策略下的尺度偏差问题。在目标特征表示方面,从频域分析视角出发,提出频率自适应融合模块,利用自适应高通滤波器和低通滤波器来增强小目标的特征表示。在AI-TODv2和AI-TOD-R数据集上的实验表明,本文方法展现出优越的小目标检测精度和泛化能力。未来工作将聚焦于模型轻量化与实时性优化,推动该方法在实际遥感应用中的部署落地。

参 考 文 献

- [1] 邹旻瑞,李宇轩,戴一冕,等. UMM-Det: 面向异构多模态遥感影像的一体化目标检测框架[J]. 电子与信息学报, 2025, 47(12): 4704–4713. doi: [10.11999/JEIT250933](#).
ZOU Minrui, LI Yuxuan, DAI Yimian, et al. UMM-Det: A unified object detection framework for heterogeneous multimodal remote sensing imagery[J]. *Journal of Electronics & Information Technology*, 2025, 47(12): 4704–4713. doi: [10.11999/JEIT250933](#).
- [2] ZHANG Xiangrong, ZHANG Tianyang, WANG Guanchun, et al. Remote sensing object detection meets deep learning: A metareview of challenges and advances[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2023, 11(4): 8–44. doi: [10.1109/MGRS.2023.3312347](#).
- [3] XU Chang, WANG Jinwang, YANG Wen, et al. Detecting tiny objects in aerial images: A normalized Wasserstein distance and a new benchmark[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 190: 79–93. doi: [10.1016/j.isprsjprs.2022.06.002](#).
- [4] FU Ronghao, CHEN Chengcheng, YAN Shuang, et al. Gaussian similarity-based adaptive dynamic label assignment for tiny object detection[J]. *Neurocomputing*, 2023, 543: 126285. doi: [10.1016/J.NEUCOM.2023.126285](#).
- [5] XU Chang, WANG Jinwang, YANG Wen, et al. RFLA: Gaussian receptive field based label assignment for tiny object detection[C]. 17th European Conference on Computer Vision-ECCV 2022, Tel Aviv, Israel, 2022: 526–543.
- [6] 刘杰,刘书豪,田明,等. 复杂环境下无人机航拍小目标检测算法[J]. 电子与信息学报, 2026, 48(4): 1763–1773.
LIN T Y, DOLLÁR P, GIRSHICK R B, et al. Feature pyramid networks for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 936–944. LIU Jie, LIU Shuhao, TIAN Ming, et al. Small object detection algorithm for UAV aerial images in complex environments[J]. *Journal of Electronics & Information Technology*, 2026, 48(4): 1763–1773. doi: [10.11999/JEIT251126](#). (本条文献为两条文献,请确认).
- [7] 王坤,丁麒龙. 利用自适应融合和混合锚检测器的遥感图像小目标检测算法[J]. 电子与信息学报, 2024, 46(7): 2942–2951. doi: [10.11999/JEIT230966](#).
WANG Kun and DING Qilong. Remote sensing images small object detection algorithm with adaptive fusion and hybrid anchor detector[J]. *Journal of Electronics & Information Technology*, 2024, 46(7): 2942–2951. doi: [10.11999/JEIT230966](#).
- [8] WU Jixiang, PAN Zongxu, LEI Bin, et al. FSANet: Feature-and-spatial-aligned network for tiny object detection in remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5630717. doi: [10.1109/TGRS.2022.3205052](#).
- [9] CHEN Linwei, FU Ying, GU Lin, et al. Frequency-aware feature fusion for dense image prediction[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(12): 10763–10780. doi: [10.1109/TPAMI.2024.3449959](#).
- [10] REN Shaoqing, HE Kaiming, 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. doi: [10.1109/TPAMI.2016.2577031](#).
- [11] WANG Jiaqi, CHEN Kai, YANG Shuo, et al. Region proposal by guided anchoring[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, USA, 2019: 2960–2969. doi: [10.1109/CVPR.2019.00308](#).
- [12] MAGID S A, ZHANG Yulun, WEI Donglai, et al. Dynamic high-pass filtering and multi-spectral attention for image super-resolution[C]. 2021 IEEE/CVF International

- Conference on Computer Vision (ICCV), Montreal, Canada, 2021: 4288–4297. doi: [10.1109/ICCV48922.2021.00425](https://doi.org/10.1109/ICCV48922.2021.00425).
- [13] XU Chang, ZHANG Ruixiang, YANG Wen, *et al.* Oriented tiny object detection: A dataset, benchmark, and dynamic unbiased learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026, 48(3): 3167–3184. doi: [10.1109/TPAMI.2025.3634161](https://doi.org/10.1109/TPAMI.2025.3634161).
- [14] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [15] LIU Dongyang, ZHANG Junping, QI Yunxiao, *et al.* Exploring lightweight structures for tiny object detection in remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 5623215. doi: [10.1109/TGRS.2025.3567345](https://doi.org/10.1109/TGRS.2025.3567345).
- [16] ZHANG Tianyang, ZHANG Xiangrong, ZHU Xiaoqian, *et al.* Multistage enhancement network for tiny object detection in remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5611512. doi: [10.1109/TGRS.2024.3363614](https://doi.org/10.1109/TGRS.2024.3363614).
- [17] HU Xinkai, REN Zhida, BHATTI U A, *et al.* DCEDet: Tiny object detection in remote sensing images based on dual-contrast feature enhancement and dynamic distance measurement[J]. *Remote Sensing*, 2025, 17(16): 2876. doi: [10.3390/rs17162876](https://doi.org/10.3390/rs17162876).
- [18] ZHANG Tianyang, ZHANG Xiangrong, HUA Chaozhuo, *et al.* Tiny object detection based on dynamic scale-awareness label assignment and contextual enhancement[J]. *Pattern Recognition*, 2026, 172: 112449. doi: [10.1016/j.patcog.2025.112449](https://doi.org/10.1016/j.patcog.2025.112449).
- [19] BIAN Jinghao, FENG Mingtao, DONG Weisheng, *et al.* Feature information driven position Gaussian distribution estimation for tiny object detection[C]. 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, USA, 2025: 30376–30386. doi: [10.1109/CVPR52734.2025.02828](https://doi.org/10.1109/CVPR52734.2025.02828).
- [20] DING Jian, XUE Nan, LONG Yang, *et al.* Learning RoI transformer for oriented object Detection in aerial images[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, USA, 2019: 2849–2858. doi: [10.1109/CVPR.2019.00296](https://doi.org/10.1109/CVPR.2019.00296).
- [21] XIE Xingxing, CHENG Gong, WANG Jiabao, *et al.* Oriented R-CNN for object detection[C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021: 3520–3529. doi: [10.1109/ICCV48922.2021.00350](https://doi.org/10.1109/ICCV48922.2021.00350).
- [22] HAN Jiaming, DING Jian, XUE Nan, *et al.* ReDet: A rotation-equivariant detector for aerial object detection[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, USA, 2021: 2785–2794. doi: [10.1109/CVPR46437.2021.00281](https://doi.org/10.1109/CVPR46437.2021.00281).
- [23] LI Wenting, CHEN Yijie, HU Xuankai, *et al.* Oriented RepPoints for aerial object detection[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, USA, 2022: 1819–1828. doi: [10.1109/CVPR52688.2022.00187](https://doi.org/10.1109/CVPR52688.2022.00187).
- [24] ZENG Ying, CHEN Yushi, YANG Xue, *et al.* ARS-DETR: Aspect ratio-sensitive detection transformer for aerial oriented object detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5610315. doi: [10.1109/TGRS.2024.3364713](https://doi.org/10.1109/TGRS.2024.3364713).
- 张天扬: 男, 讲师, 研究方向为遥感影像目标检测等, zhang-tianyang@xidian.edu.cn.
- 张向荣: 女, 教授, 研究方向为人工智能、遥感影像解译等, xrzhang@mail.xidian.edu.cn.
- 王冠淳: 男, 讲师, 研究方向为遥感影像目标检测等.
- 唐 旭: 男, 教授, 研究方向为人工智能、计算机视觉等.

Cross-Domain Collaborative Enhancement for Tiny Object Detection in Remote Sensing Images

ZHANG Tianyang ZHANG Xiangrong WANG Guanchun TANG Xu

(School of Artificial Intelligence, Xidian University, Xi'an 710126, China)

Abstract:

Objective The rapid development of deep learning has significantly advanced object detection in remote sensing images (RSIs). Due to constraints imposed by imaging conditions and object size, a large number of tiny objects (with a pixel area of less than 16×16) are widely distributed in RSIs. However, compared with satisfactory detection performance on normal-scaled objects, current object detection methods exhibit a significant performance gap for tiny objects. This is primarily attributed to the two critical limitations:

insufficient positive label assignment and weak feature representations. To address these issues, a novel Cross-Domain Collaborative Enhancement Detector (CDCEDet) is proposed. The CDCEDet jointly optimizes the label assignment in the spatial domain and strengthens the feature representations in the frequency domain for tiny objects, thereby establishing an accurate and robust framework for remote sensing tiny object detection.

Methods The overall framework of the proposed CDCEDet is depicted in Fig. 2 and comprises three key components. First, a Scale-Adaptive Anchor Generator (SAAG) is devised to dynamically generate anchors matched with the scales of ground-truth (GT) objects, effectively mitigating the scale mismatch issue neglected in previous works. Compared to conventional uniform-scale anchor generators, the proposed SAAG significantly increases positive samples for tiny objects, even under the IoU-threshold-based label assignment. Second, a Quantile-based Adaptive Label Assignment (QALA) mechanism is proposed to replace the fixed IoU threshold label assignment. The QALA exploits the IoU distribution between each GT object and its matched anchors to dynamically generate an adaptive threshold tailored to each GT object, thereby further boosting positive samples for tiny objects. Third, a Frequency-Adaptive Fusion (FAF) module is constructed to enhance feature representations from a frequency-domain perspective. On the one hand, adaptive high-pass filters are employed to highlight high-frequency information, which mitigates the information loss during channel compression. On the other hand, adaptive low-pass filters are applied to smooth and up-sample high-level features, thereby alleviating the feature inconsistency within the up-sampled objects.

Results and Discussions Extensive experiments are conducted on two publicly available remote sensing tiny object detection datasets, AI-TODv2 and AI-TOD-R, by comparing the proposed method with several latest approaches, including RFLA, DCNet, and DFCL. On the AI-TODv2 dataset (Table 1), the proposed CDCEDet achieves performance gains of 1.8% and 0.7% in terms of AP_{50} and AP_{50-95} , respectively, compared to the latest method. On the AI-TOD-R dataset (Table 2), AP_{50} and AP_{50-95} are improved by 2.6% and 0.7%, respectively. These results demonstrate that the proposed CDCEDet possesses superior performance and robust generalization capabilities for tiny object detection in RSIs. Ablation studies and configuration analyses of the CDCEDet core modules (Table 3, Table 4, Table 5, and Table 6) further verify the effectiveness of each proposed component and their complementary nature. Qualitative evaluations on both datasets (Fig. 3) confirm that the proposed method accurately detects tiny objects in both sparse and dense distributions scenarios. Fig. 4 confirms that the proposed SAAG generates scale-matched anchors for each object and assigns more positive samples to tiny objects than widely used uniformly distributed anchor generator. Furthermore, in comparative visualizations with RFLA and DCNet (Fig. 5), CDCEDet exhibits superior detection accuracy and demonstrates a more effective reduction in missed detections.

Conclusions This paper proposes a novel CDCEDet to address the insufficient label assignment and weak feature representations in remote sensing tiny object detection. Specifically, a SAAG is proposed to generate scale-matched anchors tailored to each GT object, significantly increasing the number of positive samples for tiny objects. Additionally, a QALA mechanism is devised that exploits the IoU distribution between GT objects and their matched anchors to dynamically adjust the assignment threshold, thereby effectively mitigating the scale bias induced by the fixed threshold. Finally, a FAF module is constructed from a frequency-domain perspective, which adopts the adaptive high-pass filters and low-pass filters to strengthen the feature representations for tiny objects. Extensive experiments on two tiny object detection benchmarks demonstrate the superior performance and robust generalization of the proposed CDCEDet. Future research will focus on enhancing model efficiency and real-time performance to facilitate the practical deployment of the proposed approach in remote sensing scenarios.

Key words: Tiny object detection; Remote sensing images; Label assignment; Feature fusion; Frequency learning