

基于改进RT-DETR的空海无人机目标检测模型研究

赵嘉懿, 肖馨悦, 董欣乐, 李自新

云南大学信息学院, 云南 昆明

收稿日期: 2025年10月17日; 录用日期: 2025年11月17日; 发布日期: 2025年11月26日

摘要

针对空海无人机目标检测中目标尺度变化大、动态背景干扰强, 以及传统模型泛化能力不足、多尺度特征处理效率低的问题, 本文提出一种融合频域解耦与视觉Transformer的RT-DETR改进模型(FD-ViT-RT-DETR)。该模型以RT-DETR为基础框架, 前端设计频域解耦模块, 通过快速傅里叶变换分离图像域不变特征(目标核心特征)与域特定特征(环境干扰特征), 并结合实例级对比损失强化特征区分度; 中间优化视觉Transformer架构, 新增空间压缩网络降低30%计算成本以适配边缘设备, 同时引入位置注意力偏置提升长距离依赖捕捉能力; 后端搭配不确定性最小查询选择方案优化检测精度。在含508张空海无人机图像的数据集上实验, 结果显示: 该模型精度(P)达86.4%、召回率(R)达87.4%、mAP50达88.1%, 均优于YOLOv5、YOLOv8及原始RT-DETR, 且在强逆光、海浪反光等复杂场景中鲁棒性更强, 可为空海搜救、安防监控等实际应用提供高效检测方案。

关键词

空海无人机目标检测, RT-DETR, 频域解耦, 视觉Transformer, 检测鲁棒性

Research on Air-Sea UAV Target Detection Model Based on Improved RT-DETR

Jiayi Zhao, Xinyue Xiao, Xinle Dong, Zixin Li

School of Information, Yunnan University, Kunming Yunnan

Received: October 17, 2025; accepted: November 17, 2025; published: November 26, 2025

Abstract

Aiming at the problems in air-sea UAV target detection, such as large target scale variation, strong dynamic background interference, insufficient generalization ability of traditional models, and low

文章引用: 赵嘉懿, 肖馨悦, 董欣乐, 李自新. 基于改进 RT-DETR 的空海无人机目标检测模型研究[J]. 计算机科学与应用, 2025, 15(11): 247-256. DOI: 10.12677/csa.2025.1511301

efficiency in multi-scale feature processing, this paper proposes an improved RT-DETR model integrating frequency domain decoupling and Vision Transformer (FD-ViT-RT-DETR). Based on the RT-DETR framework, the model is designed with a frequency domain decoupling module at the front end, which separates image domain-invariant features (core target features) from domain-specific features (environmental interference features) through fast Fourier transform, and enhances feature distinguishability by combining instance-level contrastive loss. In the middle layer, the Vision Transformer architecture is optimized by adding a spatial compression network to reduce 30% of computational costs for adapting to edge devices, while introducing positional attention bias to improve the ability of capturing long-range dependencies. At the back end, an uncertainty-minimized query selection scheme is adopted to optimize detection accuracy. Experiments on a dataset containing 508 air-sea UAV images show that the model achieves a precision (P) of 86.4%, a recall (R) of 87.4%, and an mAP50 of 88.1%, all outperforming YOLOv5, YOLOv8, and the original RT-DETR. It also exhibits stronger robustness in complex scenarios such as strong backlight and sea wave reflection, providing an efficient detection solution for practical applications like air-sea search and rescue, and security monitoring.

Keywords

Air-Sea UAV Target Detection, RT-DETR, Frequency Domain Decoupling, Vision Transformer, Detection Robustness

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

无人机技术的蓬勃发展为空海目标检测带来全新机遇，其在海上搜救、环境监测、安防监控等领域的应用价值日益凸显。然而，空海环境的特殊性给目标检测任务带来严峻挑战：一方面，无人机高空或海上飞行获取的图像存在目标尺度变化大、分辨率低的问题；另一方面，海浪、云层、光照反射等动态背景干扰，易导致目标特征被掩盖，传统检测模型难以稳定提取有效特征[1]。

早期空海目标检测依赖传统图像处理技术，如基于颜色、纹理的特征分类与边缘检测，但这类方法泛化能力弱，难以适应复杂环境变化。随着深度学习发展，基于 CNN 的目标检测算法(如 YOLO 系列、Faster R-CNN)逐渐应用于该领域，但 CNN 的局部感受野限制了其对空海场景长距离依赖关系的捕捉能力，在复杂背景下检测效果有限[2]。DETR 系列算法通过 Transformer 自注意力机制突破局部感受野局限，其中 RT-DETR 作为实时端到端框架，在速度与精度上优于传统 YOLO 系列，但在空海环境中仍面临泛化能力不足、多尺度特征处理效率低的问题[3]。

频域解耦技术通过将图像从时域转换到频域，分离不同频率成分，为提升模型泛化能力提供新思路；视觉 Transformer 则凭借强大的全局建模能力，可有效捕捉空海场景中的长距离关联[4]。基于此，本文将频域解耦与视觉 Transformer 融合到 RT-DETR 框架中，设计高效混合编码器与不确定性最小查询选择方案，旨在解决空海无人机目标检测的核心痛点，为实际应用提供更优解决方案。

2. 模型改进

本文提出的基于频域解耦与视觉 Transformer 融合的 RT-DETR 模型，核心改进包括频域解耦模块、优化视觉 Transformer 架构两部分，整体架构如图 1 所示。该架构以 RT-DETR 为基础，前端引入频域解耦模块处理空海图像的复杂干扰，中间通过优化的视觉 Transformer 强化特征提取能力，后端结合高效混

合编码器与不确定性最小查询选择方案, 实现检测精度与实时性的平衡, 适配无人机边缘计算设备的算力需求[5]。

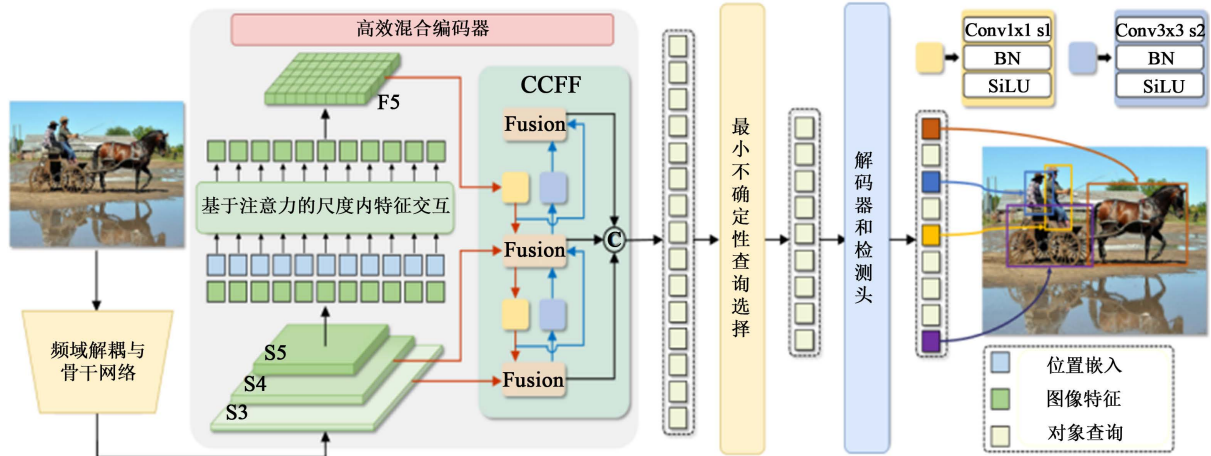


Figure 1. The improved network diagram

图 1. 改进后的网络图

2.1. 频域解耦

为克服空海环境中光照变化、天气干扰导致的模型泛化能力下降问题, 本文设计频域解耦模块, 通过分离图像的域不变特征与域特定特征, 提升模型对复杂环境的适应性。该模块结构如图 2~4 所示, 其中 F 表示快速傅里叶变换(FFT), 骨干网络分为 B1(特征提取前段)和 B2(特征提取后段), H 表示检测头, ROI-Align 表示感兴趣区域对齐操作, P 表示对比投影头[6]。

2.1.1. 频域分离与特征提取

首先, 对输入的空海无人机图像执行快速傅里叶变换(FFT)

$$F_{ff} = \text{FFT}(F_s) \quad (1)$$

将时域信号转换为频率域信号, 得到图像的振幅谱与相位谱。

$$\text{Amp}, \text{Phase} = \text{Decompose}(F_{ff}) \quad (2)$$

振幅谱包含图像的强度信息, 相位谱包含图像的结构信息。针对空海场景中光照、海浪等干扰主要影响振幅谱的特点, 设计两个可学习滤波器 Ψ_{si} 和 Ψ_{ss} , 分别用于提取对泛化有积极贡献的域不变谱和有消极贡献的域特定谱:

$$\text{Amp}_{si} = \Psi_{si} \quad (3)$$

$$\text{Amp}_{ss} = \Psi_{ss} \quad (4)$$

将提取的域不变振幅谱、域特定振幅谱分别与原始相位谱结合, 形成两个独立的频率域特征信号, 完成频域分离过程[7], 具体流程如图 2 所示。

2.1.2. 逆变换与特征优化

对两个频率域特征信号分别执行逆快速傅里叶变换(IFFT), 将其转换回时域, 生成域不变分量(包含目标核心特征)与域特定分量(包含环境干扰特征),

$$F_{si} = \text{IFFT}(\text{Amp}_{si}, \text{Phase}) \quad (5)$$

$$F_{ss} = \text{IFFT}(Amp_{ss}, Phase) \quad (6)$$

流程如图 3 所示。随后, 将两个分量分别输入骨干网络的 B1 部分, 提取对应的初始特征图[8] (记为 X_{si} 和 X_{ss})。

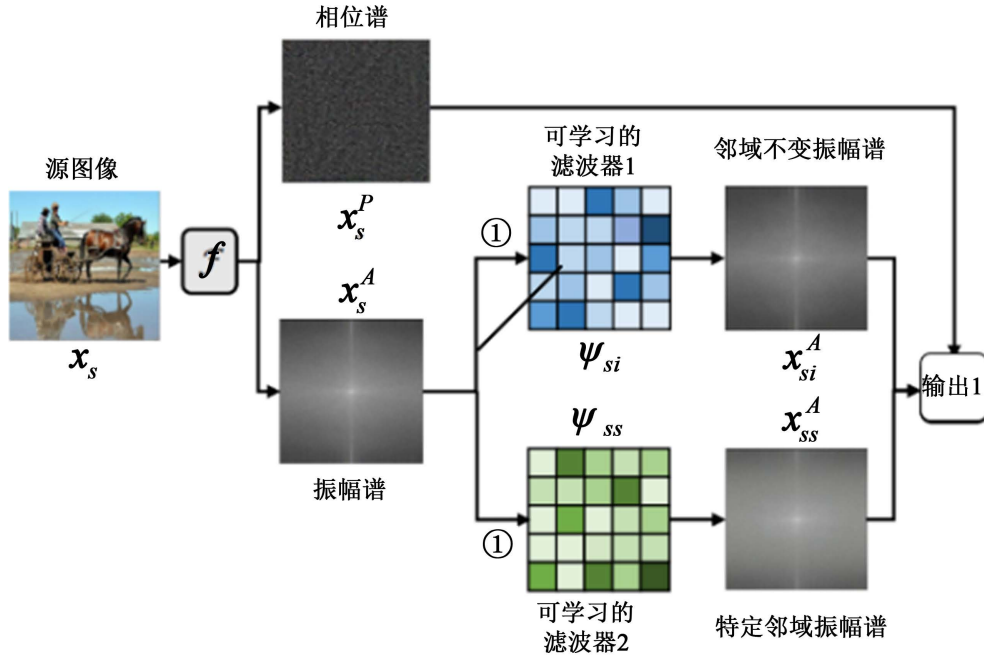


Figure 2. Fourier transform of the original image
图 2. 图像进行傅里叶变换

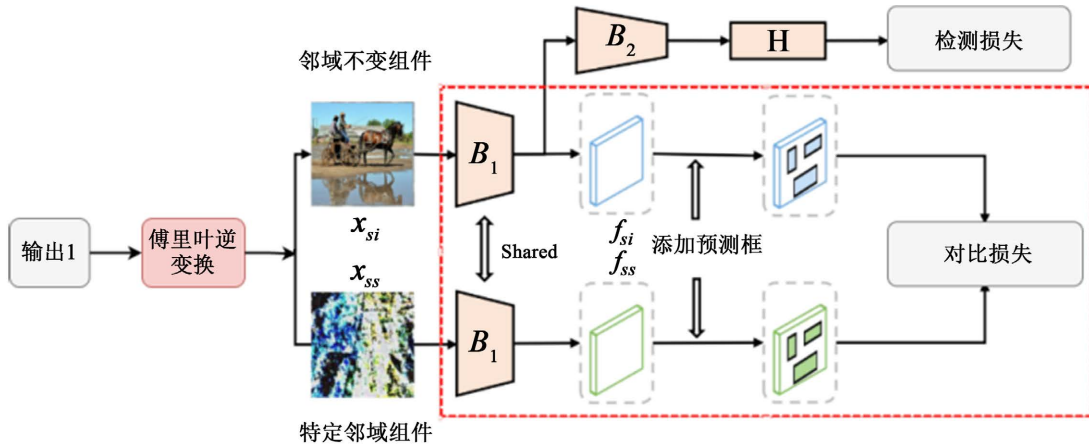


Figure 3. Generating domain-invariant components and domain-specific components via Inverse Fast Fourier Transform (IFFT)

图 3. 进行逆快速傅里叶变换(IFFT)生成领域不变分量和领域特定分量

根据图像中的目标定位标签与特征图维度比例, 从 X_{si} 和 X_{ss} 中裁剪出包含目标的实例级特征区域, 并通过 ROI-Align 操作对齐两者的空间尺寸, 确保后续特征处理的一致性。为强化域不变特征与域特定特征的区分度, 引入对比投影头 P, 将实例级特征映射到统一的特征空间, 并设计实例级对比损失:

$$f_{si} = \text{ROI Align}(F_{si}, bbox) \quad (7)$$

$$f_{ss} = \text{ROI Align}(F_{ss}, \text{bbox}) \quad (8)$$

其中, bbox 为目标边界框。

训练过程中采用交替优化策略: 先固定骨干网络与检测头参数, 优化可学习滤波器与对比投影头, 确保频域解耦效果; 再固定频域解耦模块参数, 优化骨干网络与检测头, 提升目标检测精度。通过迭代优化, 实现频率域解耦与网络训练的协同推进, 详细结构如图 4 所示。

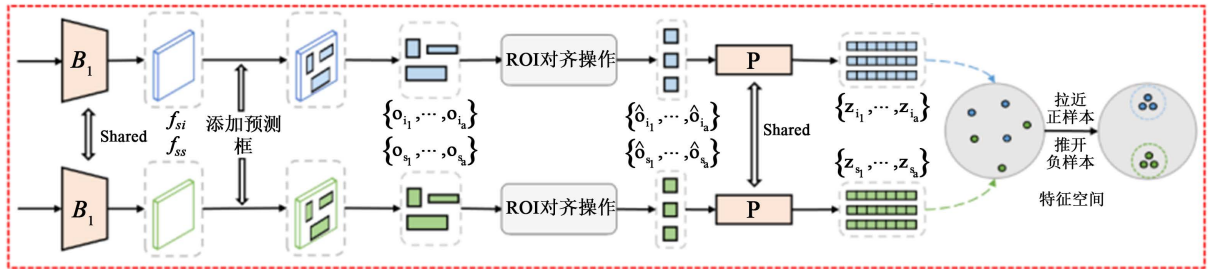


Figure 4. The detailed structure of generating domain-invariant components and domain-specific components via Inverse Fast Fourier Transform (IFFT)

图 4. 进行逆快速傅里叶变换(IFFT)生成领域不变分量和领域特定分量的详细结构

2.2. 视觉 Transformer 结构

为解决传统视觉 Transformer (ViT) 在空海场景中计算成本高、局部细节捕捉能力弱[9]的问题, 本文设计优化的视觉 Transformer 模块, 由三个级联的 Transformer 单元组成, 结构如图 5 所示。该模块在 ViT 编码模块(层标准化、多头注意力、多层感知机)基础上, 新增空间压缩网络与位置注意力偏置, 兼顾全局特征建模与轻量化需求[8]。

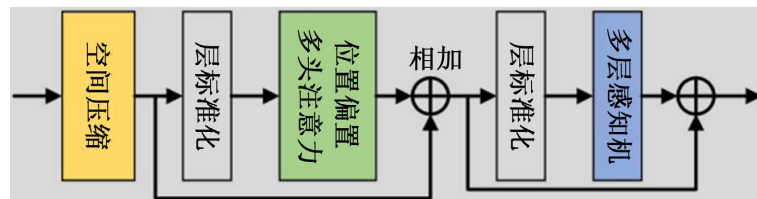


Figure 5. Visual Transformer structure

图 5. 视觉 Transformer 结构

2.2.1. 空间压缩网络

针对无人机边缘设备算力有限的特点, 在视觉 Transformer 的输入端引入空间压缩网络, 通过 2 倍下采样操作减少特征图的空间维度(如将 128×128 的特征图压缩至 64×64)。该操作在降低 Transformer 注意力计算的 Token 数量、减少 30% 计算成本的同时, 通过最大池化保留特征图中的关键结构信息(如小目标的轮廓特征), 避免因下采样导致的细节丢失。

2.2.2. 位置注意力偏置

空海场景中目标与背景的空间关联(如船舶与海平面的相对位置、救生艇与海浪的分布关系)对检测精度至关重要。传统多头注意力仅依赖特征相似度计算权重, 易受空海场景中大面积动态背景(如海浪、云层)干扰, 导致目标与背景特征混淆。为此, 在多头注意力计算中引入位置注意力偏置, 通过学习空海场景的空间先验信息, 动态调整不同位置 Token 的注意力权重, 其数学公式与在注意力计算中的作用方式具体如下:

位置注意力偏置的计算分为“偏置矩阵生成”与“带偏置的多头注意力计算”两步, 确保偏置能精

准适配空海场景的空间特征分布[10]。首先是位置注意力偏置矩阵的生成，设视觉 Transformer 输入特征图的空间维度为 $H \times W$ (实验中适配无人机图像分辨率，设 $H = W = 64$)，Token 总数 $N = H \times W$ ，Token 特征维度为 d (设为 256)。第一步通过正弦位置编码嵌入 Token 的空间位置信息，该编码能量化 Token 在特征图上的空间坐标，避免不同位置编码重叠，其公式为：

$$PE(i) = \left[\sin\left(\frac{p_{x,i}}{10000^{\frac{2k}{d}}}\right), \cos\left(\frac{p_{x,i}}{10000^{\frac{2k}{d}}}\right), \sin\left(\frac{p_{y,i}}{10000^{\frac{2k}{d}}}\right), \cos\left(\frac{p_{y,i}}{10000^{\frac{2k}{d}}}\right) \right]_{k=0}^{\frac{d}{4}-1} \quad (9)$$

其中 $p_{x,i}$ 、 $p_{y,i}$ 是第 i 个 Token 在特征图上的横、纵坐标， k 为维度索引 $d = 256$ 时， $k = 0, 1, \dots, 63$ ， $PE(i)$ 维度为 d ，确保与 Token 特征维度一致以直接参与后续计算。

第二步通过多层感知机(MLP)学习空海场景的空间依赖关系，生成位置注意力偏置矩阵 $B \in R^{M \times N}$ ，公式为：

$$B_{i,j} = \text{MLP}(PE(i) - PE(j)) \quad (10)$$

这里 $PE(i) - PE(j)$ 计算第 i 个查询 Token 与第 j 个键 Token 的位置差异向量，捕捉两者的空间相对关系，MLP 为两层感知机(输入维度 d 、隐藏维度 $d/2 = 128$ 、输出维度 1)，通过空海场景数据训练可学习“海平面附近 Token 需增强注意力”“云层内部 Token 需抑制注意力”等先验知识， $B_{i,j}$ 作为第 i 个查询 Token 对第 j 个键 Token 的偏置值，取值范围为 $[-1, 1]$ ，正值表示增强注意力权重，负值表示抑制注意力权重。

完成偏置矩阵生成后，将其融入传统多头注意力计算，通过偏置调整原始注意力分数，实现对空海场景空间关联的精准捕捉，带位置注意力偏置的多头注意力计算公式为：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right) \quad (11)$$

其中 Q 、 K 、 V 分别为查询矩阵、键矩阵、值矩阵， $\frac{QK^T}{\sqrt{d}}$ 是原始注意力分数，用于量化 Token 间的特征相似度， B 与原始注意力分数逐元素相加，共同决定最终注意力权重，Softmax 则通过归一化操作确保注意力权重总和为 1，便于后续特征加权融合。

2.3. 最小不确定性查询选择

RT-DETR 的解码器初始化依赖查询选择，传统方案基于分类分数挑选查询，忽略定位精度，导致所选特征存在不确定性(如“分类为船舶但定位边界框偏差大”)，影响解码器性能。针对该问题，本文提出不确定性最小查询选择方案，通过量化定位与分类的一致性、优化不确定性损失，为解码器提供高质量初始查询，其原理、算法流程与实现细节如下：

2.3.1. 核心原理

特征不确定性定义为“定位预测分布与分类预测分布的差异”：若一个查询的定位边界框与真实框重合度高、分类置信度高，则定位与分类预测分布一致，不确定性小，该查询更适合作为解码器初始查询；反之，若定位偏差大或分类置信度低，则不确定性大，需排除该查询。

通过构建不确定性损失函数，将不确定性量化指标融入模型训练，引导模型选择“定位准确且分类可靠”的查询，提升解码器初始化质量，进而优化检测精度。

2.3.2. 算法流程

不确定性最小查询选择的算法流程分为 4 步，具体如下：

1) 编码器特征提取:

将频域解耦模块输出的域不变特征 X_{si} 输入高效混合编码器(由 AIFI 与 CCFF 模块组成), 得到编码器特征图 $E \in R^{H \times W \times d}$, 并将其展平为 Token 序列 $E_{flat} \in R^{N \times d}$ ($N = H \times W$)。

2) 初步查询候选生成:

对每个 Token $e_i \in E_{flat}$, 通过分类头(cls_head)与回归头(reg_head)分别预测分类置信度 c_i 与边界框 b_i , 筛选出分类置信度 $c_i \geq 0.5$ 的 Token 作为初步查询候选集 $C = \{e_i | c_i \geq 0.5\}$ 。

3) 不确定性计算:

对候选集 C 中的每个查询 e_i , 计算其不确定性 U_i , 公式如下:

$$U_i = \|P(e_i) - C(e_i)\|_2 \quad (12)$$

其中: $P(e_i)$ 为定位预测分布, 由边界框 b_i 与真实框 b_{gt} 的交并比(IoU)归一化得到,

$$P(e_i) = \frac{\text{IoU}(b_i, b_{gt})}{\max_C \text{IoU}(b_j, b_{gt})}; \quad C(e_i) \text{ 为分类预测分布, 由分类置信度 } c_i \text{ 归一化得到, } C(e_i) = \frac{c_i}{\max_C c_j}; \quad \|\cdot\|_2$$

为 L2 范数, 量化两个分布的差异, U_i 越小表示查询的不确定性越低。

4) 最优查询选择

从候选集 C 中选择不确定性最小的 K 个查询(K 设为 100, 与 RT-DETR 解码器查询数量一致)作为解码器的初始查询 $Q = \{e_i | U_i \leq \text{top} - K(U_C)\}$ 。

2.3.3. 算法流程

1) 不确定性损失函数: 将不确定性指标融入模型训练损失, 总损失函数如下:

$$L_{total} = L_{det} + \lambda \cdot L_{uncert} \quad (13)$$

2) 高效计算优化: 为避免不确定性计算导致的 latency 增加, 采用“分批次计算”策略: 将候选集 C 按批次(batch size = 32)输入不确定性计算模块, 利用 GPU 并行计算模块 U_i , 确保查询选择过程的耗时控制在 1 ms 以内, 不影响模型实时性。

3) 动态调整机制: 根据空海场景的复杂度(如强逆光、海浪反光场景的复杂度高于晴天场景)动态调整分类置信度阈值(0.5~0.7): 场景复杂度高时降低阈值(如 0.5), 增加候选集数量以避免漏选有效查询; 场景复杂度低时提高阈值(如 0.7), 减少候选集数量以提升计算效率。场景复杂度通过图像的亮度方差 ($\sigma_{\text{brightness}}$) 量化, $\sigma_{\text{brightness}} \geq 0.3$ 为高复杂度场景, $\sigma_{\text{brightness}} < 0.3$ 为低复杂度场景。

3. 实验结果分析

3.1. 硬件条件及参数设置

实验数据集来自 Kaggle 平台, 采用 COCO 格式注释, 共 508 张空海无人机目标图像, 按训练集:测试集:验证集 = 12:1:3 的比例划分。模型训练参数设置如下: 批量大小为 64, 迭代次数为 100 次, 初始学习率为 0.001, 权重衰减系数为 0.0005。

Table 1. Operating environment

表 1. 运行环境

CPU	RAM	GPU	操作系统	Cuda	Pytorch	Python
R9 7945HX	16GB	RTX4060	Windows 11	11.6	28.0	3.8.0

3.2. 评价指标

为全面、客观地评估该模型在空海无人机目标检测任务中的性能, 本文从模型效率与检测精度两大维度, 精度(Precision, P)、召回率(Recall, R)、平均精度(Average Precision, AP)及平均精度均值(mean Average Precision, mAP)共 4 项核心评价指标, 形成多维度评估体系。

精度(Precision, P): 衡量模型预测为“目标”的结果中真实目标的比例, 反映模型抑制假阳性(如将浪花误判为目标)的能力, 计算公式为:

$$P = \frac{TP}{TP + FP} \quad (14)$$

其中, TP (True Positive)表示“预测为目标且实际为目标”的有效预测框数量; FP (False Positive)表示“预测为目标但实际为背景”的无效预测框数量。空海场景中, 该指标重点评估模型对“海浪、云层”等背景干扰的抗干扰能力。

召回率(Recall, R): 衡量数据集中所有真实目标被模型成功检测的比例, 反映模型捕捉真阳性(如小目标漏检)的能力, 计算公式为:

$$R = \frac{TP}{TP + FN} \quad (15)$$

平均精度(Average Precision, AP): 针对单一目标类别, 计算不同置信度阈值(0.05~0.95, 步长 0.05)下精度 - 召回率曲线(P - R 曲线)下的面积, 综合反映模型对该类目标的检测性能, 单位为百分比(%). 空海场景中重点关注“救生艇”“浮标”等小目标类别的 AP 值。

$$AP = \int_0^1 P(R) dR \quad (16)$$

平均精度均值(mean Average Precision, mAP): 计算数据集中所有目标类别的 AP 值的算术平均值, 是衡量模型整体检测性能的核心指标, 单位为百分比(%), 计算公式为:

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c \quad (17)$$

3.3. 实验结果

3.3.1. 频域解耦模块对比实验

为验证本文 FD 模块“双可学习滤波器 + 实例级对比损失”设计的有效性, 引入固定高斯滤波、小波变换两种传统频域处理方法作为对照, 在原始 RT-DETR 基础上单独嵌入不同频域模块, 对比核心指标变化, 结果如表 2 所示。

Table 2. Comparative experiments of frequency domain modules

表 2. 频域模块的对比实验

Models	P	R	mAP50	mAP50-90
base	86.0	86.1	85.4	64.6
+固定高斯滤波	86.2	86.1	87.5	64.1
+小波变换	86.0	87.0	86.1	64.7
+FD (ours)	86.4	87.4	88.1	65.3

由表 2 可知, 本文提出的频域解耦模块在各项指标上均优于传统频域处理方法, P 值提升 0.4%, R 值提升 1.5%, mAP50 及 mAP50-90 均最高, 证明其能有效提升模型在空海环境下的泛化能力。

3.3.2. 消融实验

解耦模块、优化视觉 Transformer、与不确定性最小查询选择方案, 结果如表 3 所示。

Table 3. Ablation experiments

表 3. 消融实验

experiments	FD	ViT	UMQS	P	R	mAP50	mAP50-90
0				86.0	86.1	85.4	64.6
1	√			83.6	86.5	86.4	65.0
2		√		84.9	84.5	87.9	65.4
3			√	85.0	86.7	88.3	64.3
4	√	√		85.2	87.0	88.0	65.1
5	√	√	√	86.4	87.4	88.1	65.3

3.3.3. 主流算法横向对比

为验证本文提出的频域解耦与视觉 Transformer 融合的 RT-DETR 模型在空海无人机目标检测的有效性, 本文选取目前目标检测领域的一些主流检测算法进行对比。如: YOLOv5、YOLOv8、RT-DETR。表 4 为各算法在本文自制数据集下训练的客观性能指标:

Table 4. Comparative experiment of object detection algorithm

表 4. 目标检测算法对比实验

Models	P	R	mAP50	mAP50-90
YOLOv5	85.4	81.1	85.4	64.6
YOLOv8	85.7	82.2	85.6	65.8
RT-DETR	86.0	86.1	85.4	64.6
ours	86.4	87.4	88.1	65.3

由表可知, 本文提出的改进模型相较于现目前的主流检测模型, 在复杂空海环境的检测上表现最佳, mAP50 值、mAP50-95 值分别为 88.1%、65.3%, 在对比实验中均为最高, 相较于改进前各项指标均有提升, 改进后模型性能提升显著。

3.3.4. 检测效果与分析

选取 RT-DETR 与改进后的模型检测进行对比, 检测效果如图 6 所示。初始 RT-DETR 模型在复杂空海环境干扰测试, 极端场景(如强逆光、海浪反光、云层遮挡)因素影响下存在漏检误检、准确性下降、小目标检测效果不佳的缺陷, 改进后的模型能够准确检测图像中的所有目标, 与初始 YOLOv8 模型相比, 准确率也有所提升, 成功地解决了漏检误检、小目标检测效果不佳的问题。

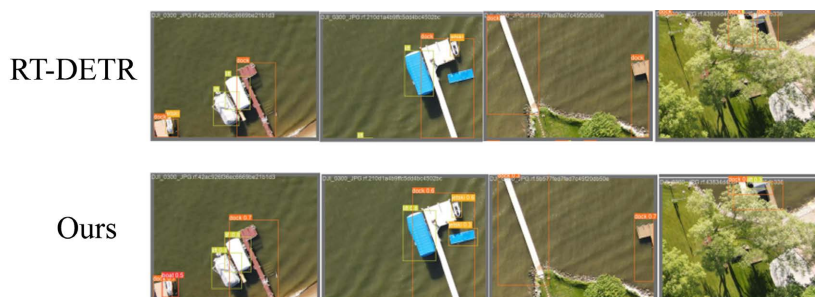


Figure 6. Comparison diagram of model detection effects

图 6. 模型检测效果对比图

4. 结论

本文提出的 FD-ViT-RT-DETR 模型针对空海无人机目标检测场景的核心挑战,通过三级改进机制实现了检测性能的显著提升。频域解耦模块借助双可学习滤波器与实例级对比损失,有效分离目标核心特征与海浪、光照等动态背景干扰,显著增强了模型对复杂环境的抗干扰能力;优化视觉 Transformer 架构通过空间压缩网络降低 30% 计算成本以适配边缘设备,同时结合位置注意力偏置强化长距离特征关联,平衡了检测精度与实时性需求;不确定性最小查询选择方案进一步提升了目标捕捉的完整性。实验结果显示,该模型在空海无人机数据集上的精度(86.4%)、召回率(87.4%)及 mAP50 (88.1%)均全面优于 YOLOv5、YOLOv8 及原始 RT-DETR,尤其在强逆光、云层遮挡等极端场景中,检测结果的稳定性与鲁棒性表现突出。综上,FD-ViT-RT-DETR 为无人机空海监测、搜救安防等实际应用提供了兼具高精度与场景适应性的技术支撑,验证了频域-空域特征融合在复杂环境目标检测中的有效性。

参考文献

- [1] 尹秋燕, 丁婧, 聂志刚. YOLO-AirPose: 无人机航拍视角下的人体姿态估计算法[J/OL]. 计算机应用, 1-10. <https://link.cnki.net/urlid/51.1307.TP.20250929.1700.014>, 2025-10-19.
- [2] 邱略能, 郑志祥, 孟德威. 基于改进 YOLO 的高压配电网污闪绝缘子无人机检测方法[J]. 计算技术与自动化, 2025, 44(3): 123-127.
- [3] 童浩然, 金涵. 面向无人机航拍图像的端到端检测算法[J]. 现代电子技术, 2025, 48(19): 103-109.
- [4] 陈星星, 李生林, 周香伶. 改进 RT-DETR 的水下目标检测算法[J/OL]. 激光与光电子学进展, 1-15. <https://link.cnki.net/urlid/31.1690.TN.20250923.1413.108>, 2025-10-19.
- [5] 张学锋, 秦继洋, 龙红明, 等. 基于改进 RTDETR 的边缘多尺度焦粉检测[J/OL]. 钢铁, 1-14. <https://doi.org/10.13228/j.boyuan.issn0449-749x.20250388>, 2025-10-19.
- [6] 王泽玄, 雷雪梅. 基于 RT-DETR 的轻量化交通标志检测算法[J]. 现代电子技术, 2025, 48(18): 57-64.
- [7] 刘臣杰, 刘巍, 杨雯迪, 等. DEPA-YOLO: 无人机视角下的小目标检测模型[J/OL]. 计算机科学与探索, 1-16. https://kns.cnki.net/kcms2/article/abstract?v=VUvWpoE9A3ITi70SHDHk8XxLy6S4ylaGPE8nLnY-XZX064l4mP2hQs-ao09cZfB4PtFrEZOF0Ym4MMcsNO4QeA-trHO5OA0jIXtzP0sLLJJ4tqzjM9evB-praBqEnY0HyWDCmqxhSwNHNjv0NMYUhQ-fyO-ccax_E0PQmZO-OsN6FXtVIJk2T0q1TtpxKasFidKs9U5zxHw=&uniplatform=NZKPT, 2025-10-19.
- [8] 何懿璇, 叶兆元, 郑凯扬, 等. 基于 ViT 的农作物检测方法与应用研究[J]. 南方农机, 2025, 56(15): 55-58.
- [9] 祝欣宇, 窦迅, 牛鹏艺, 等. 基于改进 ViT 模型的电网关键线路智能预测方法[J/OL]. 电力自动化设备, 1-15. https://kns.cnki.net/kcms2/article/abstract?v=VUvWpoE9A3L2kQe65OSkbPvCEOCb0r5-WTWy6gyLRd5DzLwCZ2p_pF_p3B6LUaJX78F3pc1cJa-cQE_BIVm80FETjc54v6itVR3x62jCVySIF3jKLminXKDduJ4km4hXfaWvisg1rl5yKnRAjRp6pA6q0de4-feP2oeE-HwCJEQKAoSvSoPUM_9RPQH3dQ3xPoB5HLoggB_2o=&uniplatform=NZKPT, 2025-10-19.
- [10] 孙杰. 基于时频域解耦与双向注意力机制的显著性检测算法研究与应用[D]: [硕士学位论文]. 杭州: 浙江大学, 2023.