

基于改进 YOLOv7 的遥感图像舰船目标检测算法

董紫微¹, 任鸿翔^{1,2}, 杨晓², 唐海娜³, 郑嘉琦²

1. 大连海事大学 人工智能学院, 辽宁 大连 116026;
2. 大连海事大学 航海学院, 辽宁 大连 116026;
3. 中国科学院大学 人工智能学院, 北京 100049)

摘要:针对遥感图像中舰船目标存在尺度变化大、大长宽比、排列密集、背景信息复杂等特点,提出了一种基于改进 YOLOv7 的遥感图像舰船目标检测算法。以 YOLOv7 作为基线网络,优化数据集先验锚框生成算法;使用长边表示法结合圆形平滑标签方法,解决边界回归的周期性带来的突变问题;并将全局注意力(GAM)与无参注意力(SimAM)嵌入到 YOLOv7 网络中,有效削弱遥感图像中复杂背景区域信息带来的干扰,提高目标检测准确率;优化坐标框损失函数,加快模型收敛速度。通过对舰船数据集(DOTA-ship),以及 HRSC2016 数据集执行单类检测与多类检测任务,mAP 结果分别为 86.1%、97.7%、87.1%,相比于原 YOLOv7 分别提升了 7.8%、4.6%、7.9%,验证了方法的有效性与优越性。

关键词:遥感图像;舰船目标;YOLOv7;注意力机制

中图分类号:TP391.4 文献标志码:A DOI:10.3969/j.issn.1000-4653.2025.03.017

Remote sensing image ship target detection algorithm based on improved YOLOv7

DONG Ziwei¹, REN Hongxiang^{1,2}, YANG Xiao², TANG Haina³, ZHENG Jiaqi²

1. College of Artificial Intelligence, Dalian Maritime University, Dalian 116026, China;
2. Navigation College, Dalian Maritime University, Dalian 116026, China;
3. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: To address challenges such as significant scale variations, high aspect ratios, dense arrangements, and complex backgrounds in ship target detection from remote sensing images, this paper proposes an improved YOLOv7-based algorithm. Using YOLOv7 as the baseline network, the prior anchor generation algorithm is optimized for the dataset. A long-edge representation method combined with circular smooth labeling is introduced to detect ship targets with uncertain rotation directions. The YOLOv7 network is enhanced by embedding both the GAM (Global Attention Mechanism) and SimAM (Simple Attention Mechanism) modules, which effectively suppress interference from complex background regions in remote sensing images and improve target detection accuracy. Additionally, the coordinate loss function is optimized to accelerate model convergence. Experimental results on the DOTA-ship and HRSC2016 datasets for both single-class and multi-class detection tasks show mAP values of 86.1%, 97.7%, and 87.1%, respectively-representing improvements of 7.8%, 4.6%, and 7.9% over the original YOLOv7 model. These results validate the effectiveness and superiority of the proposed method.

Key words: remote sensing images; ship target; YOLOv7; attention mechanisms

收稿日期:2024-07-29

基金项目:国家自然科学基金项目(52071312);广西重点研发计划项目(AB22080106);辽宁省应用基础研究计划项目(2023JH2/101300144);交通运输行业重点科技项目(2022-ZD3-035)

作者简介:董紫微(1999—),女,硕士研究生,研究方向为计算机视觉。E-mail:dzw201712@163.com

通信作者:任鸿翔(1974—),男,教授,研究方向为计算机影像学。E-mail:dmu_rhx@163.com

引用格式:董紫微,任鸿翔,杨晓,等.基于改进 YOLOv7 的遥感图像舰船目标检测算法[J].中国航海,2025,48(3):137-146.

DONG Z W, REN H X, YANG X, et al. Remote sensing image ship target detection algorithm based on improved YOLOv7[J]. Navigation of China, 2025, 48(3): 137-146. (in Chinese)

遥感图像对近岸和远海舰船目标检测在民用和军事领域都具有重要的应用价值。在民用领域上,该技术是监督渔业活动,为渔业管理提供实时信息的重要途径。在军事领域上,通过该技术控制舰船目标是确保海上交通和海岸安全的关键手段。实时监测舰船位置和行踪,军队能及时识别潜在威胁,维护海域秩序与安全。传统的舰船检测方法通常是利用图像中的纹理、形状、色彩等信息进行建立数学模型或者利用视觉显著性提取特征。然而这些方法大多依赖于舰船目标先验知识,无法应对背景复杂多变的遥感图像。同时,由于该技术采用人工提取特征的方式,往往会消耗大量人力和物力资源。随着遥感卫星和人工智能技术的发展,利用遥感图像快速且准确地检测舰船目标已成为热门的研究方向^[1-2]。遥感图像的舰船目标检测存在着许多挑战:1)舰船目标极易受到云雾、船舶尾迹、环绕海浪与海杂波的干扰,以及同一艘舰船在不同天气环境下外表特征变化大,准确识别较为困难;2)停靠在码头港口的舰船,往往排列密集,并且易受到岸边建筑物、码头堆积物等相似假目标的干扰;3)由于遥感图像为俯视视角,目标往往呈现不同的旋转方向,且因同一观测区域内舰船目标的尺度变化大,对于方向旋转极其敏感^[3]。

基于深度学习的目标检测算法流程主要分为三个阶段:信息区域选择、特征提取和分类。算法主要分为两类:一阶段目标检测算法与两阶段目标检测算法。一阶段目标检测的代表算法是 YOLO 系列算法^[4-5],单阶段多框检测器(Single Shot MultiBox Detector, SSD)算法^[6]等。两阶段目标检测算法代表算法为基于区域的卷积神经网络(Region-based Convolutional Neural Network, R-CNN)算法^[7]、FasterR-CNN^[8]算法等。在两阶段目标检测算法中先使用区域提议网络(Region Proposal Network, RPN)生成潜在的对象区域后,再对这些提议的区域进行对象分类和位置回归,速度较慢,不适合实时检测。相比与两阶段目标检测算法,一阶段检测算法直接进行目标对象检测,无需区域提议阶段,检测速度也更快。

近年来,有许多研究者在上述算法上进行改进,使其更适合遥感图像中的目标检测任务。一部分研究人员通过对网络结构进行改进,提高目标检测精度。例如:刘涛等^[9]通过将轻量级通道注意力机制引入到 YOLOv5 网络特征融合结构中,提高了网络对遥感图像中目标多尺度特征的提取能力;张徐等^[10]将 Swin Transformer 与 YOLOv7 网络结合提出了一种遥感小目标检测方法(cosSTR-YOLOv7);SUI 等^[11]将特征金字塔网络(Feature Pyramid Network,

FPN)与 SSD 算法结合,在不增加网络计算量的基础上,提高了 SSD 算法在遥感图像中目标检测的性能;还有一部分研究人员针对遥感图像中目标具有方向任意性的问题对算法进行改进。例如:XIE 等^[12]提出了一种定向区域卷积神经网络(Oriented Region-based Convolutional Neural Network, R-CNN),实现了遥感图像中任意方向的目标检测;刘鑫等^[13]提出了一种基于 CenterNet 网络的遥感旋转目标检测方法,并采用向量表示法进行旋转目标的检测与损失设计;周国庆等^[14]针对遥感舰船目标存在目标任意角度分布以及粗细粒度特征混杂等问题,提出了一种旋转目标候选框生成方法,并通过改进 Oriented R-CNN 网络,有效降低了背景冗余信息对检测效果的影响;张宇等^[15]通过融合注意力机制以及采用密集编码标签对角度信息进行离散化处理,解决了遥感图像目标方向性不定的问题;BO 等^[16]为解决遥感图像中细长旋转目标检测精度不高的问题,将可变形卷积网络(Deformable Convolutional Networks, DCN)引入到 YOLOv5 主干网络中,并边界框的长宽比信息添加到损失函数中,减少了背景信息对细长旋转目标的干扰。

尽管上述已有许多研究者针对遥感图像目标检测做了许多研究,但由于遥感图像中舰船目标存在背景信息复杂、目标尺度变化大、排列密集、小目标数目多以及具有方向不确定性等特点,容易出现错检、漏检。故本文提出了一种基于改进 YOLOv7 的遥感图像舰船目标检测算法,本文的主要工作如下:1)优化锚框生成算法;2)引入旋转目标检测方法解决舰船密集情况下的漏检问题以及舰船目标方向不确定性问题;3)在 YOLOv7 网络中嵌入无参注意力机制(Simple Attention Module, SimAM)与全局注意力机制(Global Attention Mechanism, GAM),提高网络的目标特征提取能力;4)采用 SiLU (SCYLLA-IoU Loss)函数计算坐标框损失,加快算法收敛速度。在 DOTA-ship 舰船数据集以及 HRSC2016 数据集上的试验表明,本文所提方法在遥感图像中舰船目标检测效果显著提升。

1 网络改进部分

YOLOv7 是由 WANG 等^[17]在 2022 年所提出的。YOLOv7 网络主要由 4 部分组成:输入端(Input)、主干网络(Backbone)、颈部(Neck)、预测网络(Head)。本文以 YOLOv7 网络为基线模型,在其颈部网络中嵌入 SimAM 注意力机制,用于调整 PAN 结构侧向连接处特征图的权重,能够在不增加模型参数量的前提下,降低网络上下文语义融合过程中

噪声信息的影响,提高网络提取特征的能力;将 GAM 全局注意力机制嵌入到 MP 模块中,以此抑制遥感图像中复杂的背景信息对舰船目标检测的影响;将长边表示法结合圆形平滑标签方法引入到 YOLOv7 网络中,解决旋转方向角度接近边界值时

发生突变问题;为了提高锚框的定位能力,进一步优化损失函数,将 YOLOv7 网络原有的边界框损失函数由 CIoU 替换为 SIoU;采用 K-means++ 聚类算法结合遗传算法生成数据集先验锚框,有效提高了锚框的定位精度。改进后的整体网络结构图如图 1 所示。

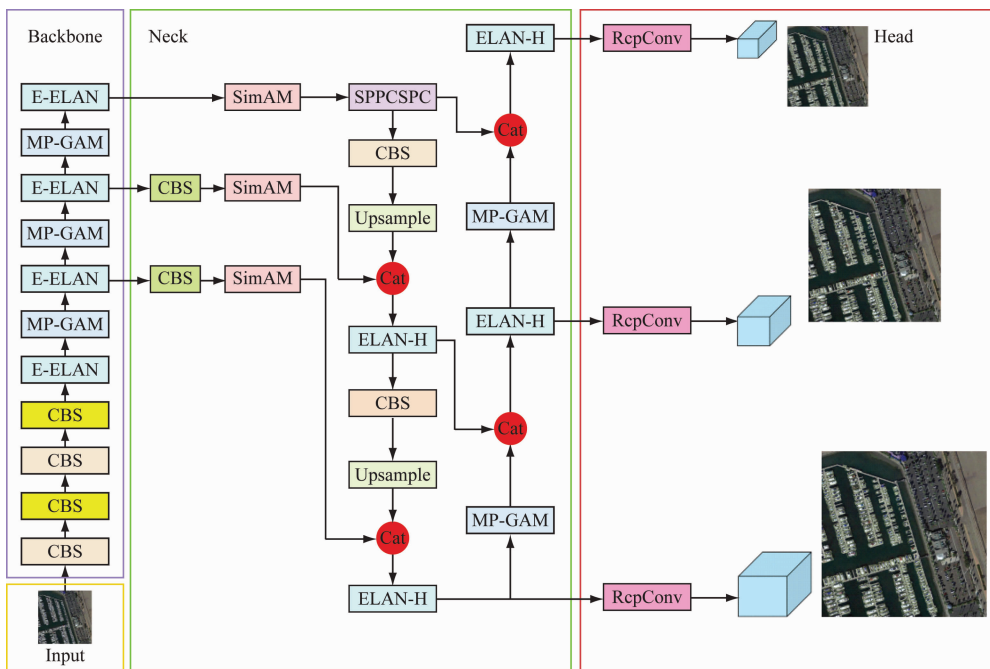


图 1 改进后的 YOLOv7 网络结构

Fig. 1 The improved network architecture of YOLOv7

1.1 优化先验锚框生成算法

在遥感图像的舰船目标检测中,准确生成先验锚框是提高检测性能的关键,为了提高检测的准确性和算法效率,本文采用 K-means++ 算法结合遗传算法生成数据集的先验锚框,而不是 YOLOv7 中传统的 K-means 算法。K-means++ 算法是由 ARTHUR 等在 K-means 算法的基础上提出的,K-means++ 聚类算法和 K-means 聚类算法的主要区别在于初始化聚类中心的方式。K-means 算法是通过从数据集中随机选择 K 个数据点作为初始的聚类中心,并不考虑数据点之间的距离或分布情况。K-means++ 则是通过最大化聚类中心之间的距离来选择初始聚类中心,这种初始化方式能够有助于避免算法陷入局部最小值,提高算法的收敛速度的目标检测先验锚框生成。K-means++ 算法初始化聚类中心的步骤为:

- 1) 随机选取一个数据点作为初始聚类中心;
- 2) 按照距离现有中心的远近,以概率分布方式选择下一个聚类中心;
- 3) 重复直到选出 K 个初始中心。

1.2 改进边界回归

与水平检测不同,在旋转目标检测的回归任务

中,主要面临两个关键问题:

1) 目标旋转方向角度接近边界值时发生角度跃变而引发的角度周期性 (Perpendicularity of Angle) 问题。

2) 目标旋转时需要同时缩放目标的宽度和高度,导致长边和短边交换造成的边的交换性 (Exchange of Edge) 问题。

在本文中,将圆形平滑标签 (Circular Smooth Label) 方法^[18]引入到 YOLOv7 网络中,将角度回归问题转变为分类问题,以此来避免角度周期性问题,如图 2 所示,标签值在角度边界处是连续的,不受周期性的影响,从而避免了损失突变误差。圆形平滑标签方法的表达式为式(1),参数 r 为窗口半径, x 为旋转角度, $g(x)$ 为窗函数,常用的窗函数有脉冲函数、矩形函数、三角函数、高斯函数等,在本文中使用高斯函数作为窗函数,如式(2)所示。为解决边的交换性问题,本文使用长边表示法区分目标对象的旋转度,如图 3 所示,该表示法格式为 (x, y, h, w, θ) ,其中 $\theta \in (-90, 90)$ 。

$$\text{CSL}(x) = \begin{cases} g(x), & \theta - r < x < \theta + r \\ 0, & \text{other} \end{cases} \quad (1)$$

$$g(x) = a \times \exp\left(-\frac{(x-b)^2}{2r^2}\right) \quad (2)$$

本文在 YOLOv7 网络的基础上添加角度参数,并将角度损失添加到 YOLOv7 网络损失函数中。损失函数公式为式(3)与式(4),其中 L_{box} 表示坐标损失,用于衡量预测框与真实框在位置和尺度上的差异; L_{obj} 表示置信度损失,用于区分前景目标与背景区域,提升目标存在性的预测准确性; L_{cls} 表示分类损失,用于约束网络对不同类别目标的识别能力; L_{theta} 表示角度损失,基于高斯角度标签建模目标的朝向信息,用于提升旋转目标的检测精度。图4是改进后的遥感图像中舰船目标检测的基本流程。

$$L = L_{\text{box}} + L_{\text{obj}} + L_{\text{cls}} \quad (3)$$

$$L_{\text{modified}} = L_{\text{box}} + L_{\text{obj}} + L_{\text{cls}} + L_{\text{theta}} \quad (4)$$

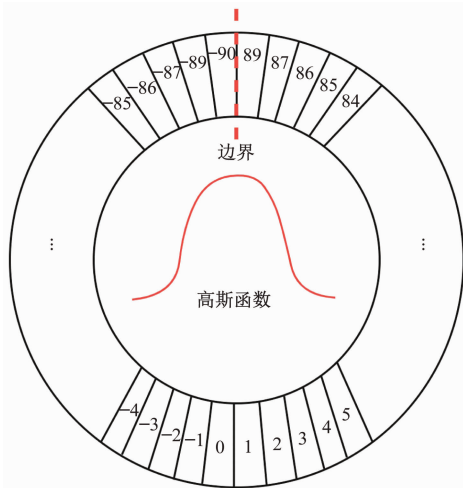


图2 圆形平滑标签方法示意

Fig. 2 Schematic diagram of Circular Smooth Labeling method

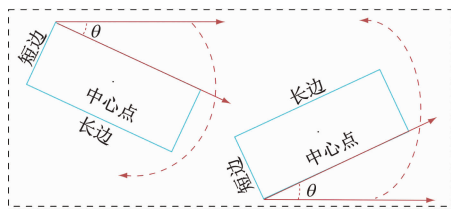


图3 长边表示法

Fig. 3 Long edge representation method

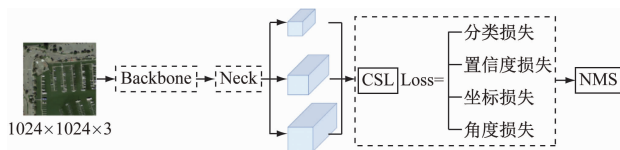


图4 改进后目标检测的基本流程

Fig. 4 The basic workflow of the improved YOLOv7 algorithm for rotated object detection is as follows

1.3 嵌入注意力机制

注意力机制通过赋予输入特征不同的权重,使模型能够更关注那些对最终任务至关重要的特征。这种方式有助于增强特征表示能力,尤其是在背景复杂、噪声较多的情况下。在神经网络中,注意力机制主要分为空间注意力和通道注意力,2017年,HU

等^[19]首次提出了通道注意力机制网络结构(Squeeze-and-Excitation Network, SE-Net),通过增强有用特征并建立通道间的依赖关系,从而提升网络性能。考虑遥感图像中舰船目标的特点与模型检测性能的要求,本文选择将 SimAM 注意力机制与 GAM 全局注意力机制嵌入到 YOLOv7 网络结构中。

1.3.1 改进 PAN 特征融合结构

由于遥感图像采样范围大且图像包含的信息丰富,因此有效聚焦舰船重要特征区域至关重要。注意力机制在深度学习中得到广泛应用,主要用于增强网络特征提取能力并聚焦于关键目标对象。YANG 等^[20]提出了一种无需额外参数的 SimAM 注意力机制,如图5所示,SimAM 注意力机制通过赋予具有关键特征信息的神经元更高的权重,并采用能量函数与注意力机制结合的方式,强化模型对目标关键特征的关注。

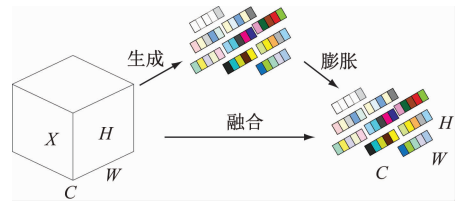


图5 SimAM 注意力机制

Fig. 5 SimAM Attention

$$e_i(w_i, b_i, y, x_i) = (y_i - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_o - \hat{x}_i)^2 \quad (5)$$

$$\bar{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (6)$$

式(5)与式(6)为 SimAM 根据神经元的重要性对其进行加权。其中式(5)定义了一个能量函数 e_i , 数值越小,表示该神经元既能被准确拟合,又能与周围神经元保持明显差异。 t 为目标神经元; x_i 为同一通道中其他神经元; y 为目标输出值,用来衡量线性可分性; w_i, b_i 为可学习的权重和偏置; X 为输入特征图; E 表示为对通道和空间两个维度上进行分组处理。

本文将 SimAM 注意力机制嵌入到 YOLOv7 网络的 PAN 结构中,如图6所示,通过调整 PAN 结构横向连接处特征图的权重大小,减少噪声信息及与背景无关信息的影响,保留舰船目标关键特征,提高模型在遥感图像中舰船目标检测任务中的性能。

1.3.2 GAM 全局注意力机制

遥感图像中的舰船目标检测常受到复杂背景的干扰,对舰船目标检测准确率产生显著影响。为了缓解背景信息引起的干扰,本文在 YOLOv7 网络中的 MP 模块中引入了 GAM 全局注意力机制^[21], GAM 注意力机制包括通道注意力子模块和空间注

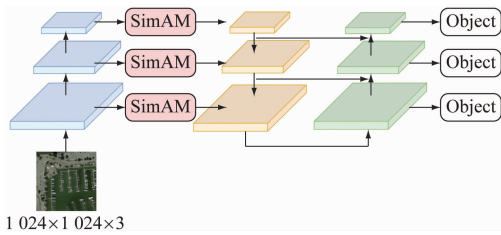


图6 引入 SimAM 注意力机制的特征融合结构

Fig. 6 Feature fusion structure with SimAM Attention mechanism

注意力子模块,通道注意力子模块通过三维转换保持输入特征的空间完整性,并通过两层多层感知器 (Multi-layer perceptron, MLP) 增强通道间的依赖关系,确保网络能够关注更有意义的前景区域。随后,空间注意力子模块进一步细化了这些特征,确保网络在复杂背景下依然能够有效检测目标,如图 7 所示,为该注意力的整体结构图。

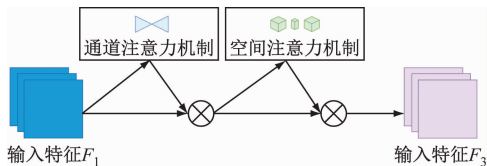


图7 GAM 注意力机制

Fig. 7 GAM Attention mechanism

在本文方法中将 MP 模块中的采用 1×1 卷积核的 CBS 模块替换为 GAM 注意力机制,如图 8 所示。该改进有助于减少下采样过程中特征信息的损失,增强全局维度特征的交互,有效地抑制了遥感图像中复杂背景信息引起的干扰,提高了舰船目标检测的准确性。

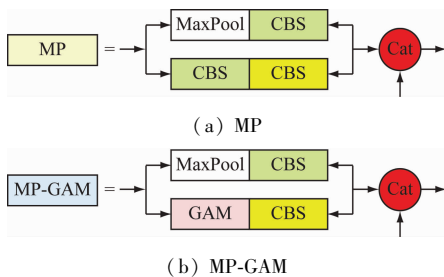


图8 MP 模块与 MP-GAM 模块的结构比较

Fig. 8 Comparison of the structure of the MP module and the MP-GAM module

1.4 损失函数改进

在 YOLOv7 网络中,边界框损失函数默认使用 CIoU 损失函数进行计算,公式为式(7)与式(8),其中 α 为权重系数; V 表示为两个预测框长宽比指标的相似性计算; GT (Ground Truth, 真实框) 表示数据集中标注的真实目标边界框, p 为预测框中心点坐标, p_{GT} 为真实框中心点坐标; c 为包围预测框和真实框的最小闭合框的对角线长度。

$$L_{CIoU} = 1 - I_{oU} + \frac{p^2(p, p_{gt})}{c^2} + \alpha V \quad (7)$$

$$\alpha = \begin{cases} 0, & I_{oU} < 0.5 \\ \frac{V}{(1 - I_{oU}) + V}, & I_{oU} \geq 0.5 \end{cases} \quad (8)$$

SIoU 是在 CIoU 的基础上进一步改进,重新定义了惩罚项并引入了角度损失这一概念,提高了训练速度与推理准确度。因此,本文选择 SIoU 进行计算边界框损失,其损失主要由角度成本、距离成本、形状成本和 IoU 成本 4 个部分组成。

1) 角度成本

角度成本计算公式为式(9), α 表示预测框和真实框之间中心点的距离, C_h 表示预测框的中心点坐标与真实框的高度差,则 $\frac{C_h}{\alpha}$ 等于 $\sin(\alpha)$ 。

$$\Lambda = 1 - 2 \times \sin^2\left(\arcsin(\sin(\alpha)) - \frac{\pi}{4}\right) \quad (9)$$

2) 距离成本

距离成本计算公式为式(10), ρ_x 代表预测框和真实框之间 x 轴坐标差异的平方比率, ρ_y 代表二者之间 y 轴坐标差异的平方比率。

$$\Delta = \sum_{t=x,y} (1 - e^{-\lambda \rho_t}) = 2 - e^{\lambda \rho_x - 2\rho_x} - e^{\lambda \rho_y - 2\rho_y} \quad (10)$$

3) 形状成本

形状成本可定义为式(11), θ 的值定义了形状成本的权重系数,控制网络对形状成本的关注程度。

$$\Omega = \sum_{t=x,h} (1 - e^{-\omega_t})^\theta \quad (11)$$

4) IoU 成本

IoU 成本的定义为预测框 B 与真实框 B_{GT} 交集与并集之比的绝对值,如式(12)所示。

$$I_{oU} = \left| \frac{B \cap B_{GT}}{B \cup B_{GT}} \right| \quad (12)$$

引入 SIoU 损失函数为坐标损失函数:

$$L_{SIoU} = 1 - I_{oU} + \frac{\Delta + \Omega}{2} \quad (13)$$

2 试验与分析

2.1 试验环境及参数设置

本文试验在装载 Ubuntu22.04 系统的 Linux 服务器下基于 Pytorch2.0.1 框架进行训练开发,试验所使用的显卡型号为的 NVIDIA RTX4070,Python 版本为 3.9,CUDA 版本为 11.8。

试验训练参数设置如下:输入图像尺寸为 1024×1024 ; 初始学习率为 0.01; 采用随机梯度下降 (stochastic gradient descent, SGD) 优化器优化学习率; 动量因子设置为 0.937; 权重衰减正则系数为

0.000 5,训练轮次为300个epoch;训练过程中使用随机翻转,Mosaic数据增强进行数据增强处理。

2.2 数据集简介与处理

本文所使用是DOTA数据集^[22]以及HRSC2016数据集。DOTA数据集是武汉大学发布的用于遥感目标检测的数据集。该数据集供有2806张航拍图像,其目标大小大多数为小目标,本试验从该数据集中提取出所有船舶的图像和位置信息,构建DOTA-shp数据集。由于该数据集的大多数图像尺寸过大,若直接输入到网络进行训练,会因为网络的统一缩放,导致图像中的小目标变形,使其更加难以检测。故需将图像进行裁剪处理,本文中将该图像统一裁剪,GAP设置为10%。

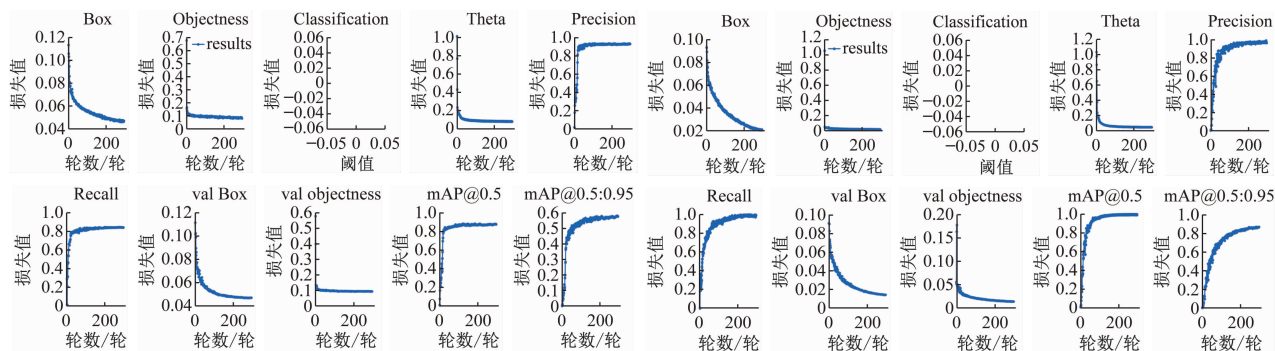
HRSC2016数据集是一个用于在遥感图像中进行舰船分类识别检测的数据集,数据集目标特点是呈大长宽比性,原数据集共分为30种不同类型的舰船,包括8种Aircraft-carrier(航母)类别(如尼米兹级航母、企业级航母、小鹰级航母等),12种Naval Ships(军舰)类别(如两栖舰、护卫舰、驱逐舰、巡洋舰等),8种Merchant-ship(商船)类别(如集装箱船、散装货船、油船等),Submarine(潜艇)类等,但由于该数据集每种类别的舰船数量不多,样本量少,会导致学习训练不够,模型鲁棒性较差。所以在本文中,将该数据集不再细分到每种类别,而是直接分为

Aircraft-carrier(航母)、Naval Ships(军舰)、Merchant-ship(商船)、Submarine(潜艇)、Other-ship(其他舰船)5种类别,重新构建HRSC2016-new数据集,弥补该数据集类间不平衡的问题。在本文中分别在原有数据集上进行单类目标检测和HRSC2016-new数据集进行多类目标检测。

2.3 试验结果及对比分析

2.3.1 试验结果

1)如图9所示,图9a为本文方法在DOTA-shp数据集上的训练过程图,图9b为HRSC2016数据集上的训练过程图;其中:Box(边界框损失)衡量预测框与真实框的差距;Objectness(置信度损失)衡量目标存在性的预测准确性;Classification(分类损失)计算类别预测误差;Theta(角度损失)表示预测与真实角度的差异。在验证阶段,val Box和val Objectness分别衡量验证集上的边界框误差与目标置信度。mAP@0.5(平均精度均值,IoU阈值为0.5)和mAP@0.5:0.95(IoU阈值0.5到0.95)用于评估检测精度。本文试验中以mAP(%)指代mAP@0.5。损失下降说明模型误差减小,Recall(召回率)和Precision(精确率)上升表明检测能力增强,mAP上升则表示检测精度提升。图10分别为本文方法与原YOLOv7在DOTA-shp数据集与HRSC2016数据集上mAP结果的对比。



(a) DOTA-shp 上训练过程

(b) HRSC2016 上训练过程

图9 训练过程

Fig.9 Training process diagram

2)如图11所示,图11a为本文方法在HRSC2016-new数据集上的训练过程图,图11b为本文方法与原YOLOv7在HRSC2016-new数据集上mAP结果的对比图,图12为本文方法在HRSC2016-new数据集各个类别的Precision曲线图。

2.3.2 试验对比分析

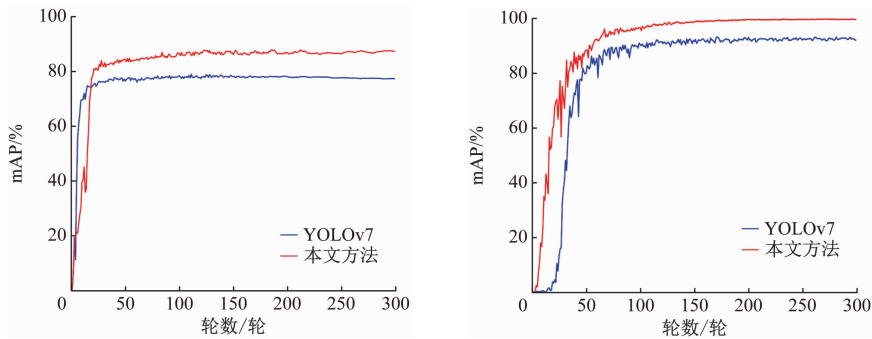
1)消融试验

为了验证每个模块算法改进的合理性和有效性,分别在三个数据集上进行消融试验,试验结果如表1所示。采用YOLOv7网络为基线网络,以本文方法具体改进部分进行4组试验,试验结果如表1

所示。本文方法与原YOLOv7相比,虽模型参数量与运算量比原YOLOv7模型增加了1.06%、1.09%,但在三个数据集上的mAP结果分别提升了7.8%、4.6%、7.9%。

2)对比试验

为了验证本文方法对于遥感场景下舰船目标检测的优越性,将本文方法与RoI-Transformer、Oriented RCNN旋转目标检测算法以及在VisDrone数据集竞赛获奖的TPH-YOLOv5算法,YOLOv5、YOLOv7、YOLOv8算法在相同试验环境下,采用相同的训练参数进行对比试验,试验结果如表2所示。

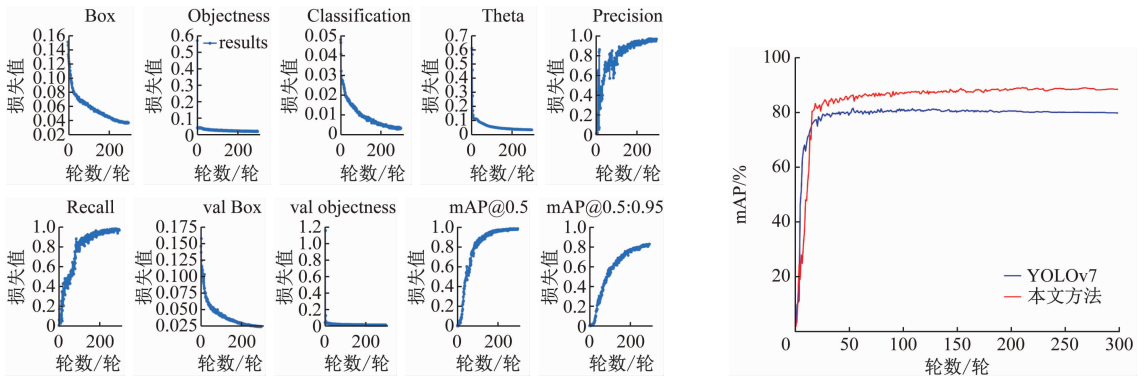


(a) DOTA-ship 上本文方法与 YOLOv7 对比

(b) HRSC2016 上本文方法与 YOLOv7 对比

图 10 本文方法与 YOLOv7 的 mAP 结果对比

Fig. 10 Comparison of mAP results between the algorithm in this paper and YOLOv7 algorithm



(a) HRSC2016-new 上训练过程

(b) HRSC2016-new 上本文方法与 YOLOv7 对比

图 11 HRSC2016-new 数据集训练结果

Fig. 11 Training results on the HRSC2016-new Dataset

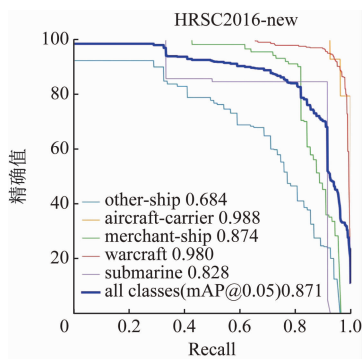


图 12 Precision 结果

Fig. 12 Precision results graph

本次研究对 YOLOv7 进行改进后,可以看出在三个数据集上检测性能超越 YOLOv8 算法,比 YOLOv8 算法的 mAP 结果分别高出 6.7%、5.3%、7.5%。只是在实时性检测上本文方法的 FPS 从表中可以看出比 YOLOv5 以及 YOLOv8 算法低,约为 35 帧/s,相比原 YOLOv7 改进后检测速度有所下降,但 FPS 仍大于 30,满足实时性检测要求。

2.3.3 试验结果分析

为了能够直观地展示本文方法的检测结果,分别对三个数据集进行检测结果可视化分析。如图 13 为分别为 YOLOv5、YOLOv7、以及本文方法在

表 1 消融试验结果对比

Tab. 1 Comparison of results for ablation experiments

组别	网络结构				mAP/%				
	CSL	SimAM	GAM	SIoU	DOTA-ship	HRSC2016	HRSC2016-new	Params/MB	GFLOPs/B
					78.3	93.1	79.2	9.32	26.1
改进 1	✓	-	-	-	83.1	95.6	83.3	9.8	28.2
改进 2	✓	✓	-	-	84.6	96.9	84.4	9.8	28.3
改进 3	✓	✓	✓	-	85.9	97.6	87.1	9.91	28.7
改进 4	✓	✓	✓	✓	86.1	97.7	87.1	9.91	28.7

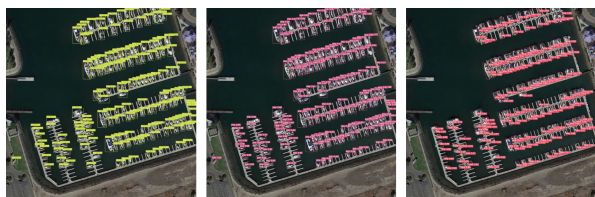
注:“✓”表示使用此方式,“-”表示不使用此方式。

表2 不同主流算法的 mAP 与 FPS 试验结果
Tab.2 mAP and FPS experimental results of different mainstream algorithms

单位: %

网络	DOTA-ship	HRSC2016	HRSC2016-new	FPS
RoI-Transformer	80.4	95.3	80.5	20.2
Oriented RCNN	83.2	96.5	83.7	18.5
TPH-YOLOv5 ^[23]	79.4	89.6	70.2	8.7
YOLOv5	78.8	91.9	74.6	51.2
YOLOv7	78.3	93.1	79.2	37.5
YOLOv8	79.4	92.4	79.6	72.3
本文方法	86.1	97.7	87.1	35.3

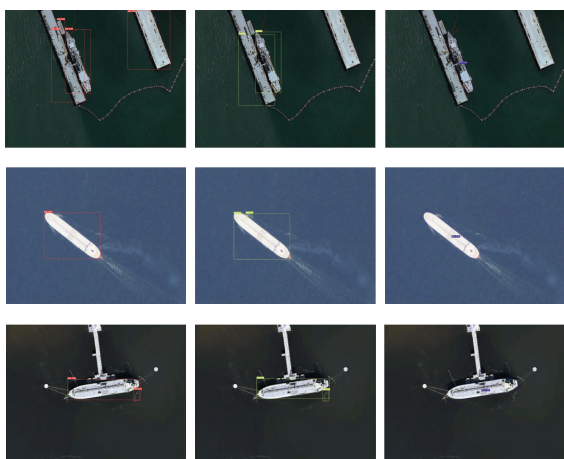
DOTA-ship 数据集上的检测结果,可以看出在舰船小目标聚集以及岸上假目标干扰的情况下, YOLOv5 与 YOLOv7 均出现了漏检和错检的情况, 本文方法则能够捕获更多的小目标特征信息, 小目标的检测准确率也更高, 根据 COCO 指标定义输出的该数据集的小目标检测准确率为 91.2%; 图 14 分别为 YOLOv5、YOLOv7 及本文方法在 HRSC2016 数据集上的检测结果, 可以看出在检测远洋航行舰船时, 检测效果极易受到尾杂波、海面反光的影响, YOLOv5 与 YOLOv7 均出现了错检的情况, 而本文方法在复杂背景环境中, 仍具有良好的检测效果, 能够准确识



(a) YOLOv5 (b) YOLOv7 (c) 本文方法

图 13 DOTA-ship 数据集上的检测结果可视化

Fig. 13 Visualization of detection results on the DOTA-ship dataset

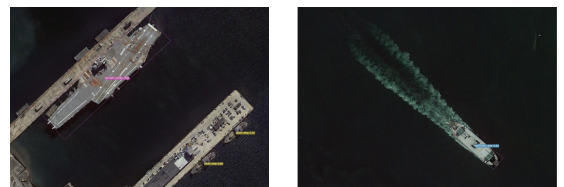


(a) YOLOv5 (b) YOLOv7 (c) 本文方法

图 14 HRSC2016 数据集上的检测结果可视化

Fig. 14 Visualization of detection results on HRSC2016 dataset

别出受强光照射以及码头相似假目标干扰下的舰船目标; 图 15 为展示本文方法在 HRSC2016-new 数据集上的多类检测任务上, 也依旧有着良好的检测效果, 能准确地识别出 Aircraft-carrier (航母)、Merchant-ship (商船)、Naval Ships (军舰)、Submarine (潜艇)、Other-ship (其他舰船) 多个目标。同时如图 16 所示在应对不同季节、光照条件变化引起的外观差异, 以



(a) 航母和其他舰船

(b) 商船



(c) 航母和军舰

(d) 潜艇

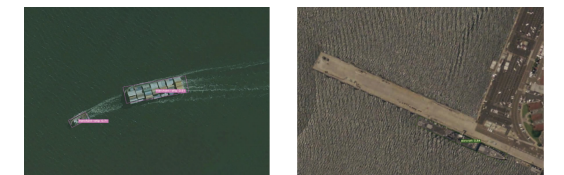
图 15 HRSC2016-new 数据集上的检测结果可视化

Fig. 15 Visualization of detection results on the HRSC2016-new dataset



(a) 不同季节、光照条件变化

(b) 云雾遮挡



(c) 航行尾迹

(d) 海面杂波

图 16 不同干扰环境下的检测效果

Fig. 16 The detection performance under challenging environmental conditions

及云雾遮挡、航行尾迹、海面杂波等外部环境干扰,本文算法模型仍能在一定程度上保持良好的检测效果,有着较强的鲁棒性。

3 结束语

本文针对遥感图像中舰船目标背景信息复杂、尺度变化大、方向具有不确定性、准确分类难等难点,提出了一种基于改进 YOLOv7 的遥感图像舰船目标检测算法。首先,根据舰船数据集特点通过使用 K-means++ 聚类算法替代 K-means 算法生成舰船数据集先验锚框,其次采用圆形平滑标签法结合长边表示法解决在旋转目标检测中因角度周期性和边的交换造成的损失值突变问题,以及通过在 YOLOv7 的 PAN 结构中嵌入 SimAM 无参注意力机制,在 MP 模块中嵌入 GAM 全局注意力机制,以此增强网络提取特征的能力,解决因遥感图像背景的复杂性而造成的小目标检测精度不高的问题,最后优化坐标框损失函数为 SIoU 函数,进一步提升网络性能。本文方法在构建的三个遥感图像舰船数据集上分别进行了单类和多类检测,试验表明,本文算法不仅在舰船密集情况下,舰船目标具有旋转不定性,还是在易受到不同干扰环境的舰船以及多类别舰船分类检测上均表现出优异的效果,有效地解决了遥感图像中舰船目标检测的难点。由于本文模型相比于原 YOLOv7 模型结构更加复杂,模型的参数量和计算量也有一定的提升,下一步将尝试在模型剪枝轻量化上做进一步研究。

参 考 文 献

- [1] 王慧赢,王春平,付强,等. 面向嵌入式平台的轻量级光学遥感图像舰船检测[J]. 光学学报, 2023, 43(12): 121-134.
WANG H Y, WANG C P, FU Q, et al. Ship detection based on lightweight optical remote sensing images for embedded platform[J]. Acta Optica Sinica, 2023, 43(12): 121-134. (in Chinese)
- [2] LI B, XIE X, WEI X, et al. Ship detection and classification from optical remote sensing images: a survey[J]. Chinese Journal of Aeronautics, 2021, 34(3): 145-163.
- [3] 徐芳,刘晶红,孙辉,等. 光学遥感图像海面船舶目标检测技术进展[J]. 光学精密工程, 2021, 29(4): 916-931.
XU F, LIU J H, SUN H, et al. Research progress on vessel detection using optical remote sensing image[J]. Optics and Precision Engineering, 2021, 29(4): 916-931. (in Chinese)
- [4] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [DB/OL]. (2020-04-23) [2020-04-23]. <https://arxiv.org/abs/2004.10934>.
- [5] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications [DB/OL]. (2022-09-07) [2022-09-07]. <https://arxiv.org/abs/2209.02976>.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [M]. Cham: Springer International Publishing, 2016: 21-37.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN [C] // Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [9] 刘涛,丁雪妍,张冰冰,等. 改进 YOLOv5 的遥感图像检测方法[J]. 计算机工程与应用, 2023, 59(10): 253-261.
LIU T, DING X Y, ZHANG B B, et al. Improved YOLOv5 for remote sensing image detection [J]. Computer Engineering and Applications, 2023, 59(10): 253-261. (in Chinese)
- [10] 张徐,朱正为,郭玉英,等. 基于 cosSTR-YOLOv7 的多尺度遥感小目标检测[J]. 电光与控制: 1-9.
ZHANG X, ZHU Z W, GUO Y Y, et al. Multi-scale remote sensing small target detection based on cosSTR-YOLOv7 [J]. Electronics Optics & Control: 1-9. (in Chinese)
- [11] SUI D L, LEI Z, WEN W W. Improving SSD for detecting small target in remote sensing image [C] // Proceedings of the 2020 Chinese Automation Congress (CAC). Shanghai, China: IEEE, 2020: 567-571.
- [12] XIE X, CHENG G, WANG J, et al. Oriented R-CNN for object detection [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2021: 3500-3509.
- [13] 刘鑫,黄进,杨瑛玮,等. 基于 CenterNet 的改进遥感旋转目标检测[J]. 遥感技术与应用, 2023, 38(5): 1081-1091.
LIU X, HUANG J, YANG Y W, et al. Improved remote sensing rotating object detection based on centerNet [J]. Remote Sensing Technology and Application, 2023, 38(5): 1081-1091. (in Chinese)
- [14] 周国庆,黄亮,孙乔. 改进 Oriented R-CNN 的遥感舰船目标细粒度检测方法[J]. 计算机工程与应用, 2024, 60(1-15): 307-317.
ZHOU G Q, HUANG L, SUN Q. Fine-grained detection method for remote sensing ship targets with improved oriented R-CNN [J]. Computer Engineering and Applications, 2024, 60(1-15): 307-317 (in Chinese)
- [15] 张宇,马杰,崔静雯,等. 融合注意力机制的遥感图像旋转目标检测算法[J]. 激光与光电子学进展, 2022, 59(24): 192-200.

- ZHANG Y, MA J, CUI J W, et al. Rotation target detection algorithm for remote sensing image using attention mechanism [J]. *Laser & Optoelectronics Progress*, 2022, 59(24): 192-200. (in Chinese)
- [16] BO Z, LU Y Y. Improved YOLOv5 in remote sensing slender and rotating target detection [C] // 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA). New York: IEEE, 2022: 918-923.
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New York: IEEE, 2023: 7464-7475.
- [18] YANG X, YAN J. On the arbitrary-oriented object detection: Classification based approaches revisited [J]. *International Journal of Computer Vision*, 2022, 130(5): 1340-1365.
- [19] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132-7141.
- [20] YANG L, ZHANG R Y, LI L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks [C] // Proceedings of the 38th International Conference on Machine Learning. New York: PMLR, 2021: 11863-11874.
- [21] LIU Y, SHAO Z, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions [DB/OL]. (2021-12-10) [2021-12-10]. <https://arxiv.org/abs/2112.05561>.
- [22] XIA G S, BAI X, DING J, et al. DOTA: a large-scale dataset for object detection in aerial images [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3974-3983.
- [23] ZHU X, LV S, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C] // Proceedings of the IEEE/CVF international conference on computer vision. New York: IEEE, 2021: 2778-2788.

~~~~~

(上接第136页)

- [14] LAI P, LIU Y, ZHANG W, et al. Intelligent controller for unmanned surface vehicles by deep reinforcement learning [J]. *Physics of Fluids*, 2023, 35(3): 1-11.
- [15] 张梦杰, 陈姚节, 邓江. 改进 TD3 算法在电机 PID 控制器中的应用 [J]. *计算机系统应用*, 2024, 33(5): 262-270.
- ZHANG M J, CHEN Y J, DENG J. Application of improved TD3 algorithm in motor PID controllers [J]. *Computer Systems & Applications*, 2024, 33(5): 262-270. (in Chinese)
- [16] 吕金旭, 葛万成. 智能制造领域中深度强化学习的应用综述 [J]. *信息与电脑 (理论版)*, 2023, 35(5): 186-193.
- LV J X, GE W C. A review of the application of deep reinforcement learning in the field of intelligent manufacturing [J]. *Information & Computer*, 2023, 35(5): 186-193. (in Chinese)
- [17] 曹宏业, 刘潇, 董绍康, 等. 面向强化学习的可解释性研究综述 [J]. *计算机学报*, 2024, 47(8): 1853-1882.
- CAO H Y, LIU X, DONG S K, et al. A survey of interpretability research methods for reinforcement learning [J]. *Chinese Journal of Computers*, 2024, 47(8): 1853-1882. (in Chinese)
- [18] 刘潇, 刘书洋, 庄韞恺, 等. 强化学习可解释性基础问题探索和方法综述 [J]. *软件学报*, 2023, 34(5): 2300-2316.
- LIU X, LIU S Y, ZHUANG Y K, et al. Explainable reinforcement learning: basic problems exploration and method survey [J]. *Journal of Software*, 2023, 34(5): 2300-2316. (in Chinese)
- [19] 陈奇, 赵炳春, 尚明健, 等. 变频阻垢除垢装置增量式数字 PID 恒流控制 [J]. *控制工程*, 2015, 22(增刊1): 89-93.
- CHEN Q, ZHAO B C, SHANG J M, et al. The constant current control of anti-scaling descaling device based on incremental digital PID controller [J]. *Control Engineering of China*, 2015, 22(Suppl. 1): 89-93. (in Chinese)
- [20] FOSSEN T. Handbook of marine craft hydrodynamics and motion control [M]. Sussex: John Wiley & Sons Ltd, 2011.
- [21] 姚培源, 魏潇龙, 俞利新, 等. 基于 Q-Learning 算法的无人机空战机动决策研究 [J]. *电光与控制*, 2023, 30(5): 16-22.
- YAO P Y, WEI X L, YU L X, et al. Research on UAV air combat maneuver decision based on Q-Learning algorithm [J]. *Electronics Optics & Control*, 2023, 30(5): 16-22. (in Chinese)
- [22] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor [C] // International Conference on Machine Learning. Stockholm, Swedn: PMLR, 2018: 2976-2989.
- [23] CUI Z, GUAN W, ZHANG X. Collision avoidance decision-making strategy for multiple USVs based on Deep Reinforcement Learning algorithm [J]. *Ocean Engineering*, 2024, 308: 118323.