

基于 YOLOv5 算法的多尺度小目标船舶识别方法

杨俊秀¹, 王荣杰^{1,2}, 林安辉^{1,2}, 王亦春^{1,2}, 曾广森¹, 蒋德松^{1,2}

(1. 集美大学轮机工程学院, 福建 厦门 361021; 2. 福建省船舶与海洋工程重点实验室, 福建 厦门 361021)

[摘要] 为提高海面多尺度小目标船舶的识别性能, 提出一种数据集划分方法, 并在 YOLOv5 算法中改进数据增强方法, 融合注意力机制, 改进损失函数。实验结果表明, 该方法能更好地识别海面上的多尺度小目标船舶, 平均精度 (mAP)、精确率 (P)、召回率 (R) 分别为 99.1%, 98.5%, 97.5%, 识别性能比经典深度学习算法和近几年的方法都高。

[关键词] 船舶; 多尺度小目标; 图像识别; 数据增强; YOLOv5 算法

[中图分类号] U 675.79; TP 391.4

Multi-Scale Small Target Ships Recognition Method Based on YOLOv5 Algorithm

YANG Junxiu¹, WANG Rongjie^{1,2}, LIN Anhui^{1,2}, WANG Yichun^{1,2}, ZENG Guangmiao¹, JIANG Desong^{1,2}

(1. School of Marine Engineering, Jimei University, Xiamen 361021, China;

2. Fujian Provincial Key Laboratory of Naval Architecture and Ocean Engineering, Xiamen 361021, China)

Abstract: A dataset division method was proposed to improve the recognition performance of multi-scale small target ships on the sea surface. The data augmentation method, which contains the YOLOv5 algorithm, was improved by integrating the attention mechanism and improving the loss function. Experimental results showed that this method could better identify multi-scale small target ships on the sea surface, with the average precision (mAP), precision (P) and recall (R) are 99.1%, 98.5% and 97.5%, respectively. It suggested that the detection performance was higher than that from classical deep learning algorithms and methods in the past several years.

Keywords: ship; multi-scale small targets; image recognition; data augmentation; YOLOv5 algorithm

0 引言

海面船舶识别研究在海防安全、海上搜救、海洋保护、港口船舶调度、海关运输管理、海面监控与缉私等方面发挥着巨大的作用。船舶雷达主要用于航行避让、船舶定位和狭水道引航等, 是船舶航行必要的导航设备, 但其也存在目标误识别、目标丢失和易受环境噪声影响等缺点。随着计算机视觉的发展, 图像分辨率有了很大的提高^[1], 获取图像的方式简单, 且图像中包含更多的目标特征信息,

[收稿日期] 2023-07-17

[基金项目] 机器视觉检测安徽省重点实验室开放基金项目“基于跨模态迁移的近海无人艇水面目标识别方法研究”(KLMVI-2023-HIT-14); 福建省自然科学基金项目“船用真空制淡/制冷一体机优化匹配及智能调控研究”(2020J01689); 福建省教育厅项目“基于深度强化学习的船舶自主靠泊运动控制方法研究”(JAT220173)

[作者简介] 通信作者: 王荣杰 (1981—), 博士, 教授, 从事智能信息处理、盲源分离和电力系统故障诊断研究, E-mail: roger811207@163.com

利用图像进行海面目标识别也得以迅速发展。而且, 计算机算力的不断提升和深度学习技术的快速发展也为目标识别奠定了基础^[2-4]。目前的目标检测主要集中在大中型目标上, MS COCO^[5]、PASCAL VOC^[6]等相关数据集主要包含大中型目标, 对小目标识别效果不好^[7]。直接用这些模型进行小目标的识别, 性能会很差, 且拍摄角度和像素分辨率也会导致检测性能不佳^[7]。这些问题使得小目标船舶的识别研究具有重要意义。

针对上述问题, 文献 [8] 提出了一种用于自主航运技术的小船检测的高效最优掩模区域卷积神经网络技术, 它采用数据增强技术解决小型船舶真实世界样本数量有限的问题, 使用轻量级卷积神经网络模型的实例分割算法来识别船舶, 使用优化器进行 SqueezeNet 模型的超参数调整, 采用加权正则化极限学习机技术的碰撞体优化算法对识别到的船舶进行分类。文献 [9] 提出了基于生物鹰眼视觉适应机制的海上小目标探测方法, 根据鹰眼的生理结构和特点, 基于鹰视觉的背景适应机制, 建立了眩光适应、暗淡适应、颜色适应和背景适应的数学模型, 利用重建图像的多尺度灰度差和高斯函数得到海上小目标的显著性图。文献 [10] 在网络各阶段的不同位置使用 3 个不同的 ResNet 块, 下采样阶段早期减少激活函数的使用, 增强网络结构中的信息流, 提出一种网络识别结构, 利用多个 ResNet 块的关键点估计和行人位置回归, 更好地平衡识别速度和精度, 引入抑制最大值的激活函数, 以防止数值爆炸和十进制敏化。

上述方法虽然能够识别到小目标, 但是检测能力弱, 仍存在漏检、误检、精确度低和召回率低等问题, 导致小目标检测困难。为提高海面小目标船舶的识别性能, 本文提出了一种数据集划分方法, 并在 YOLOv5 算法^[11]中改进数据增强方法, 加入双通道注意力机制, 增强网络对目标的特征提取能力, 改进损失函数, 以更好地提取到小目标船舶的特征信息。

1 YOLOv5 网络结构

YOLOv5^[12] 是一种单阶段目标识别算法, 有 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本, 其中, 最快和最小的版本是 YOLOv5s。YOLOv5 中 4 个版本的网络架构是相同的, 网络结构的大小由深度因子和宽度因子 2 个参数控制。例如, YOLOv5s 中的 C3 操作只执行一次, 而 YOLOv5l 的深度是 YOLOv5s 的 3 倍, 因此执行 3 个 C3 操作^[12]。YOLOv5s 在保证精确度的同时, 也具有体积较小、算力要求低、运行速度快的优点。本文使用的是 6.0 版本的 YOLOv5s, 其结构如图 1 所示, 可见, 该算法由输入端 (Input)、主干网络端 (Backbone)、颈部 (Neck) 端和预测端 (Prediction) 4 部分组成。

1.1 YOLOv5 输入端

YOLOv5 沿用了 YOLOv4 的马赛克数据增强方法, 通过随机缩放、随机裁剪、随机排列等方式进行拼接^[13], 核心思想是将图像局部区域替换为来自另一个训练图像的补丁^[14], 丰富了输入数据的多样性, 增加了单批小目标的数量, 提高了网络识别小目标的能力。

1.2 YOLOv5 主干网络

YOLOv5 主干网络包括 CSPDarknet53^[15]、卷积模块 (Conv) 和空间金字塔池化 (spatial pyramid pooling-fast, SPPF)^[16] 等模块。C3 模块是基于 CSPDarknet53 模块来学习残差特征的, 主干网末端的 SPPF 模块是基于空间金字塔池化的改进^[16]。

1.3 YOLOv5 颈部

YOLOv5 颈部采用“特征金字塔路径聚合网络”^[17-18] 的结构, FPN 结构将顶部的强语义特征传到底部特征图中^[19]。此外, PAN 结构将强大的定位特征从底部特征图传达到顶部特征图中, FPN 和 PAN 共同增强了颈部网络的特征融合能力。

1.4 YOLOv5 预测端

预测端包括预测目标框损失函数和非极大值抑制。YOLOv5 的损失函数由目标框损失函数、置信度损失函数和分类损失函数 3 部分组成^[20]。其中, 目标框损失函数使用的是 GIOU, 计算公式为

$$\text{GIOW} = \text{IOU} - |A_c - U|/|A_c|。$$

(1)

其中： A_c 表示任意两个凸形状的最小外凸形状的面积^[21]； U 为 A_c 中除两个凸形状之外的区域面积^[22]； IOU 为两个凸形状的交并比。在目标预测阶段使用加权非极大值抑制。目标置信度损失函数（ Loss_{obj} ）通过动态缩放的交叉熵损失^[22]计算得出，具体计算公式为

$$\text{Loss}_{\text{obj}} = \begin{cases} -\alpha(1 - C_i^j)^{\gamma} \lg C_i^j, \text{Obj} \in \text{pos}, \\ -\alpha(1 - C_i^j)^{\gamma} \lg (1 - C_i^j), \text{Obj} \in \text{neg}. \end{cases}$$

(2)

其中： Obj 为预测框内的目标； α 是平衡因子，用来平衡正负样本的重要性； γ 可以减少易分类样本的权重； pos 和 neg 分别为目标和背景集合； C_i^j 为第 i 个网格第 j 个预测框中包含目标的置信度，计算公式为

$$C_i^j = P_r \times \text{IOU}。$$

(3)

其中： P_r 是目标的先验分类概率。

在目标预测阶段，对于每个目标模型都会预测出多个候选框，当目标之间比较紧密时，非极大值抑制能够降低冗余边界框对网络参数更新的影响。

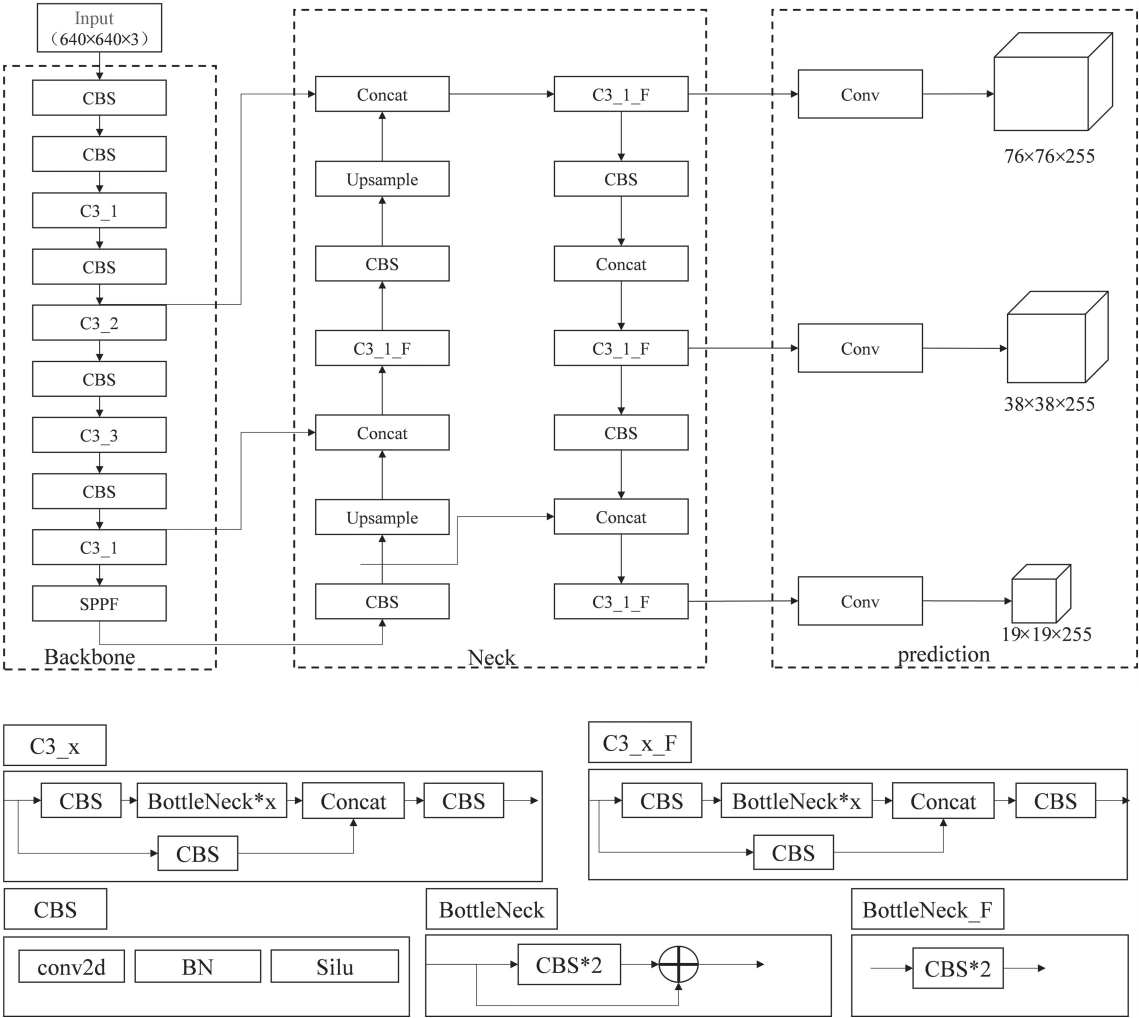


图 1 YOLOv5 网络模型
Fig.1 YOLOv5 network model

2 数据集及筛选方法

2.1 数据集

SeaShips^[23] 是一个公开船舶数据集, 该数据集共有 7 000 张图片, 包括矿石运输船、散货船、普通货船、集装箱船、渔船、客船。所有图像都来自大约 10 080 个真实世界的视频片段, 主要涵盖所有可能的成像变化, 如不同的比例、船体部分、照明、视点、背景和遮挡。数据集中天气情况是多样的, 船舶也有被遮挡的情况, 视野遮挡或者船与船之间的遮挡。此外, 同一张图中可能存在多个不同尺度的船舶。部分数据集如图 2 所示。抽取的数据集中矿石运输船、散货船、普通货船、集装箱船、渔船、客船的数量分别为 1 715, 493, 1 185, 633, 1 496, 360 艘。



图 2 部分数据集
Fig.2 Partial dataset

图 3 为数据集中标签框形状的分布。由图 3 可见, 由于数据集中船舶形状的特点, 所以标注的标签长方形居多。图 4 为标签中心位置分布图。从图 4 中可以看出, 船舶的中心位置几乎都在整张图像的下半部分, 由于航海船舶所处环境的特点, 船舶在整张图中的位置比较偏下。

图 5 为标签尺寸分布图。从图 5 中可以看出，数据集中船舶尺寸的多样性，其中有大大尺寸目标船舶，也有小尺寸的目标船舶标。从图 2 和图 5 中可知，样本尺寸差异大，且每类船的数量分布不均衡，样本尺寸和类别分布的特点更增加了小目标船舶识别的难度。

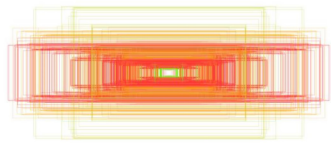


图 3 标签框形状分布
Fig.3 Distribution of label box shape

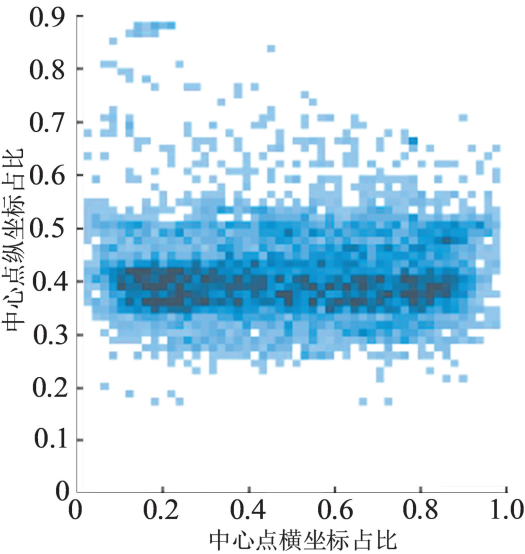


图 4 标签中心位置分布

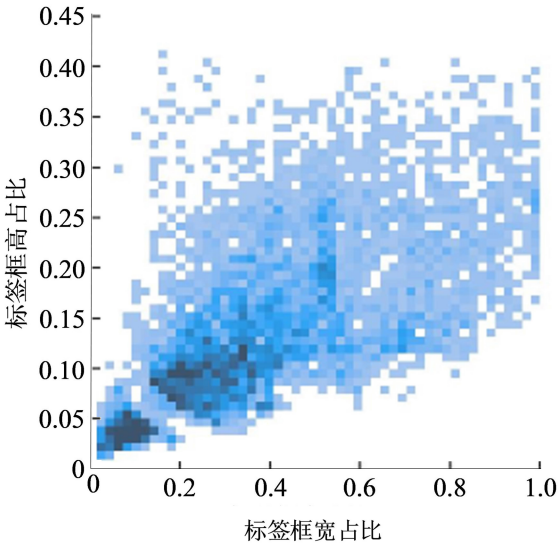


图 5 标签尺寸分布

Fig.4 Distribution of the center position of the label

Fig.5 Distribution of label box size

2.2 数据集筛选方法

由于现有大型相机公开可见光数据集中小目标占比非常少，尤其是船舶相机可见光图像数据集，缺乏小目标船舶数据集，没有足够的小目标船舶数据集用于训练，所以本文提出一种数据集划分方法，将大目标船舶图像作为训练集，小目标船舶图像作为验证集。这种数据集划分方式使得训练数据（源域）和测试数据（目标域）所属的分布不同，在训练数据上，训练模型使得模型具有很好的泛化能力，通过基于领域自适应的数据增强方法，将源域数据的分布转换成目标域数据分布，使得训练后的模型能够适用于与训练数据分布不同的目标数据。

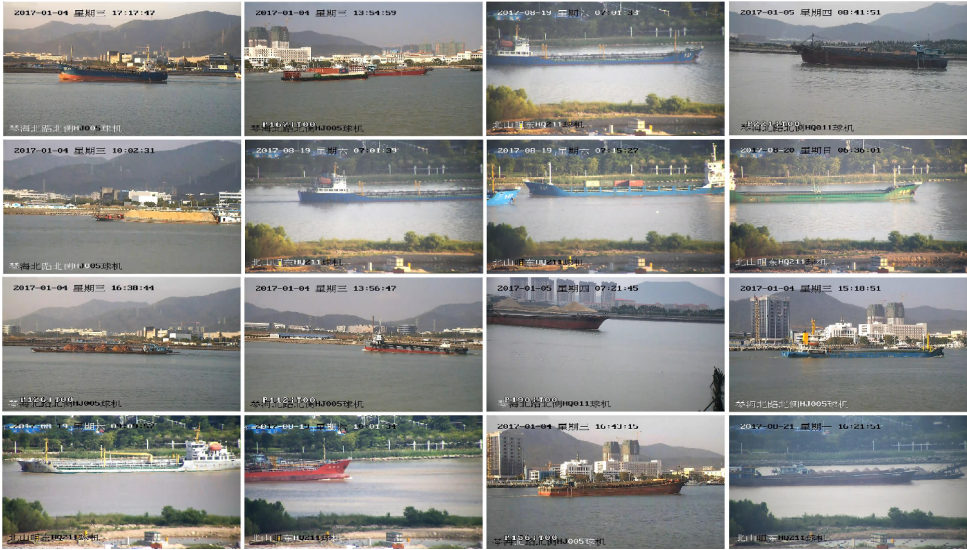
本实验从 SeaShips 数据集中抽取 6 105 张图片进行实验，从抽取的图像中筛选出小目标船舶图像作为验证集，筛选小目标数据集的伪代码为：

```
for 目标框 boxi in 图像 imageId,  
    xmax: 目标框的最大横坐标,  
    xmin: 目标框的最小横坐标,  
    ymax: 目标框的最大纵坐标,  
    ymin: 目标框的最小纵坐标,  
    w: 图像的宽,  
    h: 图像的高,  
    x = xmax - xmin ,  
    y = ymax - ymin ,  
    s = x * y ,  
    s0 = w * h ,  
    if s < s0 * 0.12 ,
```

then imgId。

该伪代码的思想是, 分别计算出目标框和整张图的面积, 若图片中存在某一船舶目标框的面积小于整张图面积的 0.12, 就把这些图片作为验证集, 实验筛选出 750 张小目标船舶图像。验证集中同一张图片中可能不仅包含小目标船舶, 也可能小目标船舶和大目标船舶同时存在, 即存在多个尺度的船舶, 更增加了数据的尺度多样性。

选取的数据集中, 750 张小目标图片作为验证数据集, 剩余 5 355 张图片作为训练数据集。其中, 训练集和验证集都涵盖了 6 种类型的船舶, 经过算法筛选后, 训练集和验证集的比例约为 7:1。图 6 展示了数据集中部分船舶图像, 其中: 图 6a 为部分训练集图片; 图 6b 为部分验证集图片。为了方便观察, 验证集中的船舶用红色矩形框标出。



a)部分训练集



b)部分验证集

图 6 训练集和验证集图片对比

Fig.6 Comparison of images from training set and validation set

在实际应用中, 还没有一个公开的、专门用于训练小目标的数据集^[24], 而现实情况是, 小目标可能单独出现在探测器视野中, 也可能和大目标同时出现在视野中。在小目标数据集缺乏的情况下,

本文将大目标图像作为训练数据集，将多尺度小目标图像作为验证数据集，并对网络进行优化。

3 YOLOv5 网络的改进

3.1 数据增强方法

为了增加数据的多样性并防止过度拟合，在目标识别中经常使用多种数据增强方法。大量实验表明，数据增强方法对识别精度的影响较大^[25]。YOLOv5 中使用的数据增强方法是对 4 张图片进行预处理。由于本文研究的是小目标识别，需要增加一个批次中小目标图像的数量，所以在数据预处理部分，本文随机选取 6 张图片，采用和马赛克一样的方法进行数据预处理，并将此方法称为 Mosaic6 数据增强，其效果如图 7 所示。

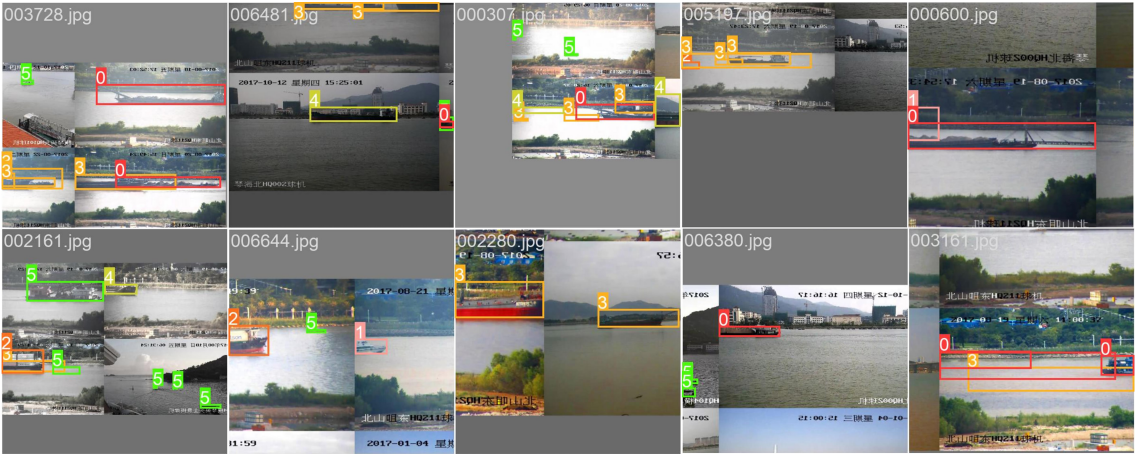


图 7 Mosaic6 数据增强效果图
Fig.7 Enhancement effect from Mosaic6 data

3.2 卷积注意力模块

计算机视觉中的注意力机制的目的是模仿人类视觉系统，能够有效地在复杂场景中找到显著区域，通过注意力机制，可以对输入图像的特征进行动态权重调整。注意力机制已经在许多视觉任务中取得了成效^[26]。

卷积注意力模块^[27]从通道和空间 2 个维度推断注意力，是一个即插即用模块，它可以放在网络模型的任何位置，其结构图如图 8 所示。在图像分类和目标识别任务中，研究者们通过实验证明，嵌入卷积注意力机制的网络模型性能得到了提升，只是不同算法在应用到不同场景时，嵌入的注意力模块的位置不一样。

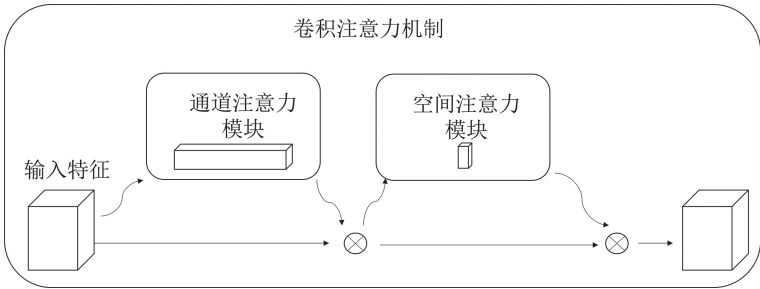
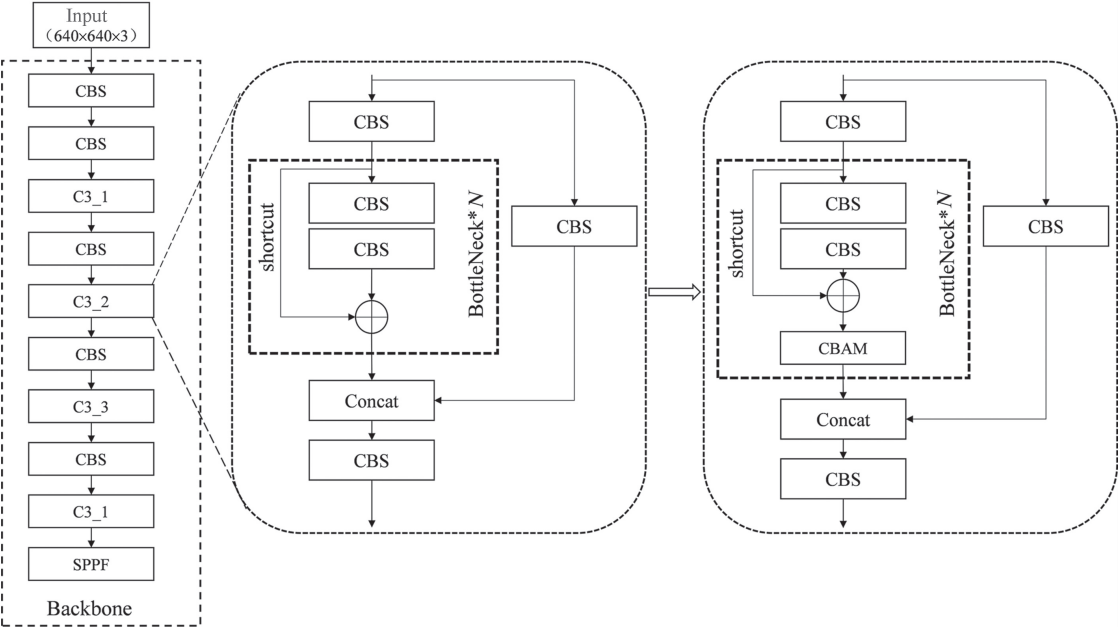
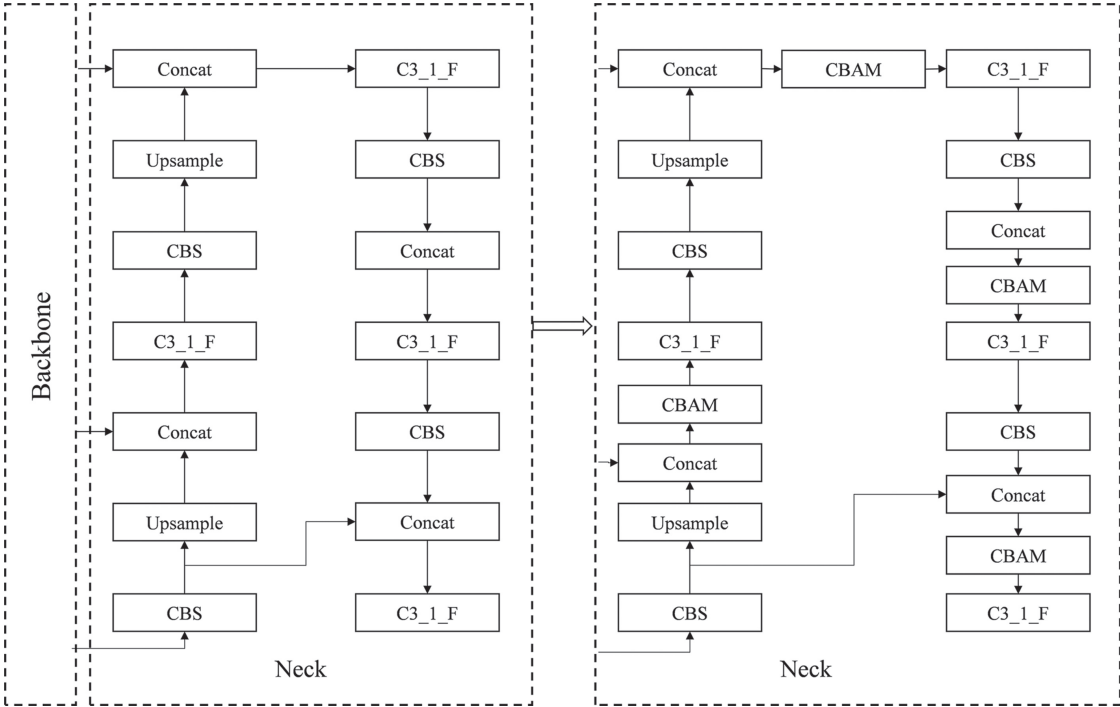


图 8 卷积注意力模块
Fig.8 Convolutional attention module

本文将卷积注意力模块分别嵌入到 YOLOv5 的主干、颈部、检测头, 观察不同位置的卷积注意力机制对模型性能的影响。分别在主干 C3 中 BottleNeck 内部 add 后, 颈部 Concat 后及检测头前面添加卷积注意力模块, 从而产生 3 种新的网络模型 YOLOv5-A、YOLOv5-B、YOLOv5-C, 卷积注意力模块具体所加位置如图 9 所示。经过多次实验对比发现, 将 CBAM 注意力模块加在主干 C3 中 BottleNeck 内部 add 后能有效提升模型的识别性能。



a) 主干加入卷积注意力



b) 颈部加入卷积注意力

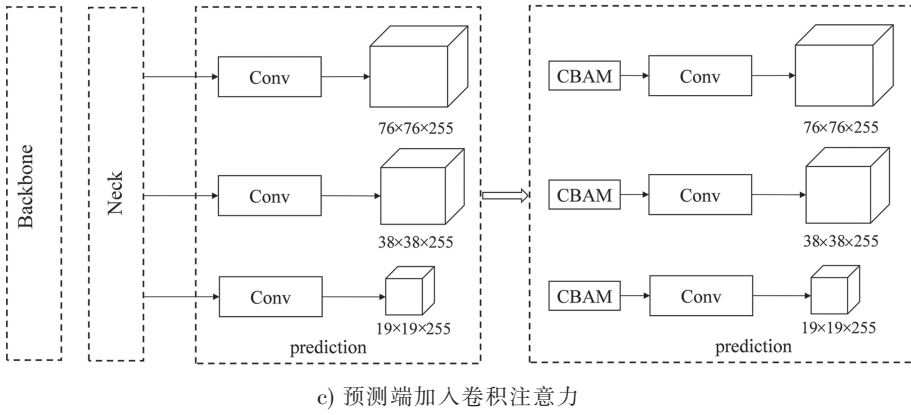


图 9 3 种嵌入卷积注意力模块的 YOLOv5 模型

Fig.9 Three YOLOv5 models embedded in CBAM modules

3.3 融合注意力模块

YOLOv5 的位置损失函数计算采用的是 GIOU 损失^[28]。与普通 IOU 计算方式相比, GIOU 虽然改善了重叠面积的计算方式, 但并没有考虑到两框中心点间的距离和宽高比不同的情况, 也没有考虑到两框之间方向不匹配的问题。因此, 本文采用了一种新的损失函数 SIOU, 考虑到距离、形状、IOU 和所需回归之间的向量角度, 重新定义了惩罚指标^[29]。向量角度的回归可以让预测框很快地移动到最近的轴, 仅需一个横坐标或者纵坐标的回归。该损失函数加快了训练收敛速度^[30], 优化了训练效果。

该损失函数的惩罚项包括角度、距离、形状和交并比 4 个成本函数^[31], 其计算公式分别为

$$\Lambda = 1 - 2\sin^2(\arcsin((C_h/\sigma) - \pi/4)), \quad (4)$$

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma\rho_t}), \quad (5)$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta, \quad (6)$$

$$\text{IOU} = \frac{|B \cap B^{\text{GT}}|}{|B \cup B^{\text{GT}}|}. \quad (7)$$

距离成本函数中 γ 与角度成本函数 Λ 的关系为

$$\gamma = 2 - \Lambda. \quad (8)$$

总损失函数为

$$I_{\text{box}} = 1 - \text{IOU} + (\Delta + \Omega)/2. \quad (9)$$

式 (4)~式 (9) 中: σ 为目标框和预测框中心点之间的距离; C_h 为目标框和预测框最小外接矩形的高; ρ_x 为目标框和预测框中心点之间的水平距离与最小外接矩形宽比值的平方; ρ_y 为目标框和预测框中心点之间的垂直距离与最小外接矩形高比值的平方; γ 为赋予时间优先的距离值; ω_w 为目标框和预测框最小外接矩形的宽和预测框宽的差值与外接矩形宽的比值; ω_h 为目标框和预测框最小外接矩形的宽和预测框高的差值与外接矩形高的比值; B 为目标的预测框; B^{GT} 为目标的真实框。

4 实验结果及分析

4.1 实验环境

本文实验的硬件环境为: 操作系统为 Windows11 (64 位), 处理器为 Intel (R) Core (TM) i7-10700F CPU@2.90GHz 2.90 GHz, 运行内存为 32 GB, 显卡为 NVIDIA GeForce RTX 3060 Ti, Python 版本为 3.9.7, CUDA 版本为 11.7。

4.2 评估指标

为了评估模型的识别精度,采用平均精度 (mAP)、精确率 (P)、召回率 (R)、 F_1 值作为衡量指标^[12]。 P 和 R 的计算公式分别为

$$P = TP / (TP + TN), \tag{10}$$

$$R = TP / (TP + FN). \tag{11}$$

其中: TP 代表真阳性; FP 代表假阳性; TN 代表真阴性; FN 代表假阴性。

AP 代表单个类别的平均精确率,是在多标签分类中衡量分类结果的一种评价指标。对于具有 n 个样本的数据集,对每个样本的真实标签进行预测,并计算概率,最后取平均值。AP 的计算公式为

$$AP = \int_0^1 P(R) dR. \tag{12}$$

mAP 是对所有类别的 AP 求取均值后所得数据^[14],其计算公式为

$$mAP = \sum_{i=1}^c (AP_i / c). \tag{13}$$

F_1 值的计算公式为

$$F_1 = 2PR / (P + R). \tag{14}$$

mAP 比精确率、召回率更能反映全局性能。如果一个系统的性能越好,其曲线向上凸出的部分越大,从几何角度讲,曲线与坐标轴之间的面积也越大。最理想的系统包含的面积应是 1,而所有系统的包含的面积都应大于 0。

4.3 网络参数与训练

实验涉及的超参数如表 1 所示,通过多次实验手动调优获得。网络训练过程中各收敛曲线如图 10 所示。

表 1 训练超参数

Tab. 1 Training hyperparameters	
实验超参	数值
初始学习率	0.010 0
最终单次学习率	0.010 0
动量	0.937 0
初始权重	yolov5s. yaml
权重衰变	0.000 5
识别框损失权重	0.050 0
分类损失权重	0.500 0
置信度损失权重	1.000 0
输入尺寸	640.000 0
训练代数	180.000 0

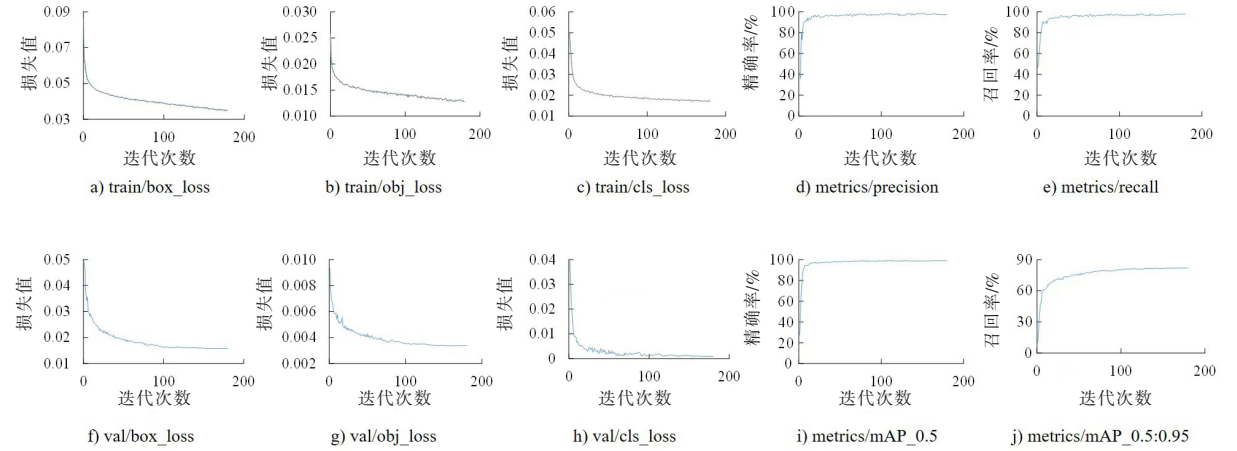


图 10 收敛曲线
Fig.10 Convergence curve

4.4 结果分析

为了验证本文方法的有效性,使用节 4.2 中的评估方法来评估小目标船舶识别的性能。本文方法训练改进后的 YOLOv5 网络,生成的混淆矩阵如图 11 所示,AP、 P 、 R 和 F_1 值分别如图 12 ~ 图 15 所示,部分船舶识别可视化效果如图 16 所示。

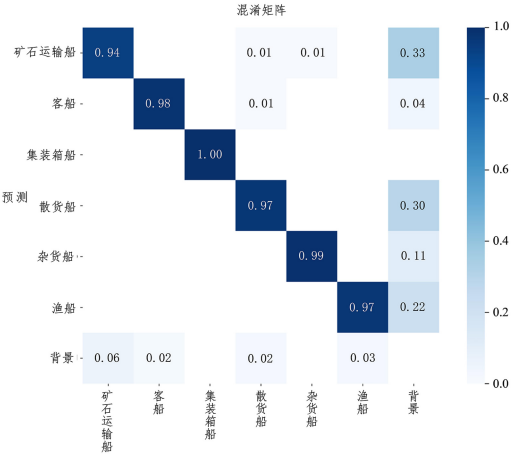


图 11 混淆矩阵
Fig.11 Confusion matrix

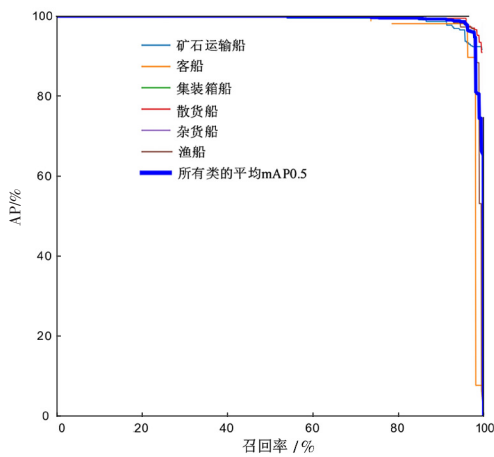


图 12 6 类船 AP 值与召回率关系曲线
Fig.12 Relationship curve between AP value and recall rate for 6 types of ships

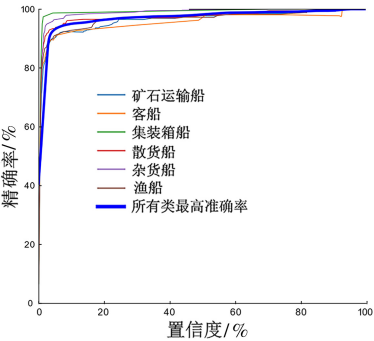


图 13 6 类船精确率与置信度关系曲线
Fig.13 Relationship curve between accuracy and confidence for 6 types of ships

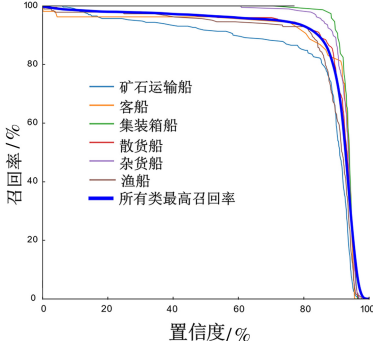


图 14 6 类船召回率与置信度关系曲线
Fig.14 Relationship curve between recall rate and confidence for 6 types of ships

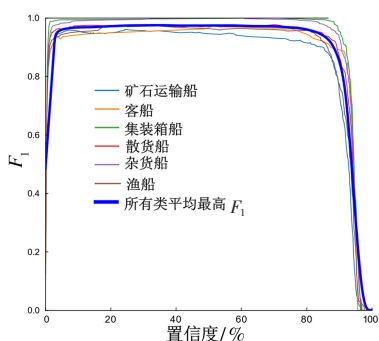


图 15 6 类船 F_1 值与置信度关系曲线
Fig.15 Relationship curve between F_1 value and confidence for 6 types of ships



图 16 部分船舶识别结果
Fig.16 Results from identification of partial ships

图 12 展示了 6 类船舶对应的 AP 值, 其 mAP 值为 99.0%。可以看出, 各类船舶的 AP 值都非常接近于 1。由图 13 可以看出, 集装箱船识别的精确度最高, 为 99.5%, 客船识别的精确度最低, 为 97.9%, 其他 4 类船舶识别的精确率均在 98% 以上。由图 14 可以看出, 集装箱船的召回率最高, 非常接近于 100%, 矿石运输船的召回率相对较低, 所有类别的平均召回率在 95% 以上。由图 15 可以看出, 几乎所有船的 F_1 值都在 95% 以上, 集装箱船的 F_1 值最高, 非常接近于 100%, 矿石运输船的 F_1 值相对较低, 所有类别平均 F_1 值约为 97%。由图 16 可以看出, 改进后的模型识别船舶的精确率几乎都非常接近 100%, 有近 1/3 的船舶识别准确率为 100%, 还有一些视野遮挡小目标船舶识别准确率接近 100%。由此可见, 改进后的模型能够较准确地识别多尺度小目标船舶。

4.5 结果对比

为了验证本文对 YOLOv5 算法改进的有效性, 与经典深度学习方法和近几年的改进方法进行对比。为了控制实验变量, 所有实验都按本文提出的数据集划分方式来划分数据集。实验结果表明, 本文提出的小目标识别方法在 P 、 R 、mAP@0.5 和 mAP@0.5:0.95 上都有显著提升, 结果如表 2 所示。分别选择一个二阶段和一阶段的深度学习算法 FasterRCNN 和 SSD 作为对比实验, 与近几年的改进算法对比时, 算法都是基于 YOLOv5 改进的算法。

表 2 实验结果对比
Tab. 2 Comparison of experimental results

单位: %					
数据集	识别方法	P	R	mAP @ 0.5	mAP @ 0.5:0.95
Seaships	Faster-RCNN	74.2	97.7	96.6	—
	SSD	96.4	87.0	96.6	—
	YOLOv5	97.2	95.8	98.4	—
	文献[34]	93.1	94.1	96.3	—
	文献[30]	—	91.4	97.3	—
	本文方法	98.5	97.5	99.1	81.3
Caltech pedestrian	文献[12]	42.4	27.5	—	—
	本文方法	54.2	46.5	—	—

由表 2 可以看出, 本文方法比 FasterRCNN 的 P 值高 24.3%, R 值仅差 0.02%, mAP 值高 2.5%, 比其他对比实验的 P 值、 R 值和 mAP 值都高。本文方法比文献 [32] 的 P 值、 R 值和 mAP@0.5 分别提升了 5.4%、3.4% 和 2.8%, R 值和 mAP@0.5 比文献 [28] 分别提升了 6.1% 和 1.8%。查阅大量文献发现, 很多文献都没有给出 mAP@0.5:0.95 的值, 仅有部分给出了 mAP@0.5:0.95 的值, 并且 mAP@0.5:0.95 值不高, 而本文中 mAP@0.5:0.95 值达到了 81.3%。

文献 [32] 和文献 [28] 也是基于改进 YOLOv5 深度学习的海上船舶识别算法, 这两种方法的识别效果不错, 但是识别平均精度不高, 对小目标船舶识别效果不好。

本文还对比了文献 [10]。由于文献 [10] 没有公开代码, 所以本文使用文献 [10] 中使用的 Caltech 行人检测数据集来训练模型。结果显示, P 值和 R 值比文献 [10] 分别提升了 11.8% 和 19%, 表明本文所提方法的有效性, 也表明所提方法有较强的泛化性, 可以应用到其他领域。

综上, 本文的改进 YOLOv5 算法, 提高了多尺度小目标船舶识别的综合性能, 并且具有较强的泛化性。

5 结论

在现实应用中,由于小目标船舶像素少,特征不明显,这导致小目标的识别性能较差。本文提出一种数据集划分方法,并在 YOLOv5 算法中改进数据增强方法,加入注意力机制,改进损失函数,整体提升了多尺度小目标识别的性能。但是,当目标出现严重相互遮挡或视觉遮挡时,识别性能还有待提升,所以,小目标遮挡问题是下一步值得研究的方向。

[参 考 文 献]

- [1] DU J M, LU H Z, ZHANG L P, et al. A spatial-temporal feature-based detection framework for infrared dim small target[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60(10): 1-12.
- [2] CHEN X L, SU N Y, HUANG Y, et al. False-alarm-controllable radar detection for marine target based on multi-features fusion via cnns[J]. IEEE Sensors Journal, 2021, 21(7): 9099-9111.
- [3] MCINTOSH B, VENKATARAMANAN S, MAHALANOBIS A, et al. Infrared target detection in cluttered environments by maximization of a target to clutter ratio (TCR) metric using a convolutional neural network[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(1): 485-496.
- [4] ZHOU X Y, YANG K D, RUI D, et al. Deep learning based on striation images for underwater and surface target classification [J]. IEEE Signal Processing Letters, 2019, 26(9): 1378-1382.
- [5] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]//Computer Vision-ECCV. Cham, Switzerland: Springer, 2014: 740-755.
- [6] EVERINGHAM M, WAN G L, WILLIAMS C K I, et al. The pascal visual object classes (voc) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [7] WEN Z H, SU J, ZHANG Y X. SIE-YOLOv5: improved YOLOv5 for small object detection in drone-captured-scenarios[C]//Knowledge Science, Engineering and Management. Cham, Switzerland: Springer Nature, 2023: 39-46.
- [8] ESCORCIA G J, GAMARRAM, BELENO K, et al. Intelligent deep learning-enabled autonomous small ship detection and classification mode[J]. Computers and Electrical Engineering, 2022, 100(3): 107871.
- [9] DUAN H B, XU X B, DENG Y M, et al. Unmanned aerial vehicle recognition of maritime small-target based on biological eagle-eye vision adaptation mechanism[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(5): 3368-3382.
- [10] WANG M Y, MA H, LIU S C, et al. A novel small-scale pedestrian detection method base on residual block group of center-net[J]. Computer Standards & Interfaces, 2023, 84(3): 103702.
- [11] 成倩, 李佳, 杜娟. 基于 YOLOv5 的光学遥感图像舰船目标检测算法[J]. 系统工程与电子技术, 2023, 45(5): 1270-1276.
- [12] DAI Y, LIU W M, WANG H, et al. YOLO-Former: marrying yolo and transformer for foreign object detection[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-14.
- [13] WANG H R. HOS-YOLOv5: an improved high-precision remote sensing image target detection algorithm based on YOLOv5 [C]//International Conference on Data Mining and Big Data. Singapore: Springer, 2022, 1745: 43-56.
- [14] YUN S D, HAN D Y, CHUN S Y, et al. Cutmix: regularization strategy to train strong classifiers with localizable features [C]//IEEE/CVF International Conference on Computer Vision. Seoul: IEEE/CVF, 2019: 6022-6031.
- [15] WANG C Y, LIAO H Y, WU Y H, et al. Cspnet: a new backbone that can enhance learning capability of CNN[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE, 2020: 1571-1580.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [17] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759-8768.
- [18] WANG H, ZHANG S, ZHAO S L, et al. Real-time detection and tracking of fish abnormal behavior based on improved yolov5 and siamrpn + + [J]. Computers and Electronics in Agriculture, 2022, 192(1): 106512.
- [19] DONG X D, YAN S, DUAN C Q, et al. A lightweight vehicles detection network model based on yolov5[J]. Engineering Applications of Artificial Intelligence, 2022, 113(1): 104914.

- [20] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu:IEEE,2017:936-944.
- [21] HAMID R, NATHAN T, JUNYOUNG G, et al. Generalized intersection over union;a metric and a loss for bounding box regression[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach,CA,USA:IEEE,2019:658-666.
- [22] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(2):318-327.
- [23] SHAO Z F, WU W J, WANG Z Y, et al. Seaships;a large-scale precisely-annotated dataset for ship detection[J]. IEEE Transactions on Multimedia,2018,20(10):10-12.
- [24] TONG K, WU Y Q. Deep learning-based detection from the perspective of small or tiny objects;a survey[J]. Image and Vision Computing,2022,123(7):104471.
- [25] ZHAO Y, SHI Y, WANG Z L. The improved yolov5 algorithm and its application in small target detection[C]//15th International Conference on Intelligent Robotics and Applications(ICIRA)-Smart Robotics for Society. Harbin:Springer,2022:679-688.
- [26] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision;a survey[J]. Computational Visual Media, 2022,8(3):1-38.
- [27] MA K, ZHAN C A, YANG F. Multi-classification of arrhythmias using resnet with cham on cwgan-gp augmented ecg gramian angular summation field[J]. Biomedical Signal Processing and Control,2022,77(3):103684.
- [28] 韦一,张妙藏,武智强. 基于改进 YoloV5 的海面小目标检测方法[C]//2022 年无人系统高峰论坛(USS2022)论文集. 西安:USS,2022:80-85.
- [29] HAN P, LIAO D Y, HAN B B, et al. Sean;a simple and efficient attention network for aircraft detection in sar images[J]. Remote Sensing,2022,14(18):4669.
- [30] ZHENG J F, WU H, ZHANG H, et al. Insulator-defect detection algorithm based on improved yolov7[J]. Sensors,2022,22(22):8801.
- [31] AN Q L, WANG K, LI Z Y, et al. Real-time monitoring method of strawberry fruit growth state based on yolo improved model[J]. IEEE Access,2022,10:124363-124372.
- [32] ZHANG H T, TIAN M, SHAO G P, et al. Target detection of forward-looking sonar image based on improved yolov5[J]. IEEE Access,2022,10(1):18023-18034.

(责任编辑 马建华 英文审校 郑青榕)