

## 基于改进 YOLOv5 的潜艇目标检测算法

梅礼坤, 陈智利

(西安工业大学光电工程学院, 陕西 西安 710021)

**摘要:** 为了解决传统潜艇目标检测缺乏对复杂背景和噪声的鲁棒性、对光照变化和视角变化敏感、难以处理大规模数据集等问题, 提出了一种基于改进 YOLOv5 潜艇目标检测器。通过 C3\_Transformer 结构, 有效提升了特征的全局上下文建模能力和长距离依赖性捕捉能力; 通过 simOTA 解决 anchorbased 算法中正负样本不平衡问题, 增强模型对小目标和困难样本的学习能力; 利用 decoupledhead 的思想解决分类和位置预测任务的互斥性问题, 提高检测精度和鲁棒性。实验结果表明: 相较于原始 YOLOv5, 改进后的模型 Precision、Recall、mAP0.5、mAP 0.5 : 0.95 分别提高了 2.8%、10.9%、3.8%、14.7%, 这表明改进后的模型在潜艇目标检测的准确性、召回率以及在不同置信度阈值下的平均准确率等方面取得了明显的进步, 同时在实际检测任务中改进后的模型有效解决了“漏检”、“误检”的问题。

**关键词:** 上下文; 样本不平衡; 互斥性; 最优传输问题

**中图分类号:** TP391; TN29 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2024.06.020

## Submarine target detection algorithm based on improved YOLOv5

MEI Li-kun, CHEN Zhi-li

(School of Optoelectronic Engineering, Xi'an Technological University, Xi'an 710021, China)

**Abstract:** In this paper, an improved YOLOv5-based submarine target detector is proposed to solve the problems that traditional submarine target detection lacks robustness to complex backgrounds and noises, is sensitive to changes in illumination and viewing angle, and is difficult to deal with large-scale datasets. With the C3\_Transformer structure, the global context modeling ability of the features and the long-range dependency capturing ability are effectively improved. And the simOTA algorithm is employed to address the issue of imbalanced positive and negative samples in anchor-based algorithms, thereby enhancing the model's learning capabilities for small targets and challenging samples. Additionally, the decoupledhead approach is utilized to overcome the mutual exclusivity problem between classification and position prediction tasks, resulting in improved detection accuracy and robustness. The experimental results show that compared to the original YOLOv5, the improved model shows significant advancements in terms of Precision, Recall, mAP@0.5, and mAP@0.5 : 0.95, with improvements of 2.8%, 10.9%, 3.8%, and 14.7% respectively, which indicates that the improved model achieves notable progress in terms of accuracy, recall rate, and average precision at different confidence thresholds in submarine target detection. Furthermore, the improved model effectively addresses the issues of "missed detection" and "false detection" in the actual detection task.

**Keywords:** context; sample imbalance; mutual exclusion; optimal transmission problem

基金项目: 陕西省科技厅基金项目 (No. 2023-YBGY-369) 资助。

作者简介: 梅礼坤 (1996 -), 男, 硕士研究生, 主要研究方向为数字图像处理、深度学习技术。E-mail: 2112868362@qq.com

通讯作者: 陈智利 (1973 -), 男, 博士, 教授, 硕士生导师, 主要研究方向为光学设计、目标检测技术。E-mail: medichen@163.com

收稿日期: 2023-06-28; 修订日期: 2023-08-13



### 3 改进 YOLOV5

#### 3.1 C3\_Transformer

在目标检测领域,传统的检测算法如 YOLO (You Only Look Once) 系列使用了 CNN 来提取图像特征和回归边界框。然而,这些方法无法充分捕捉到不同位置的全局上下文信息,导致检测性能的限制。Transformer<sup>[15]</sup> 的自注意力机制能够全局地关注图像中的不同位置,有助于捕捉目标的上下文信息,提高目标检测的准确性。其次,Transformer 模型能够更好地处理图像中的长距离依赖关系,这对于目标检测任务中的大尺度目标或复杂场景具有重要意义。最后,Transformer 融合的主干网络能够学习更丰富、更有表达力的特征表示,有助于提升目标检测算法的性能。图 2 为 C3\_Transformer 框架图。

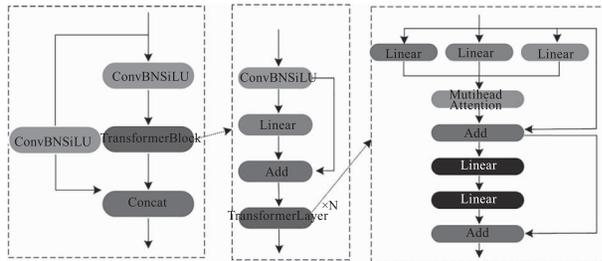


图 2 C3\_Transformer

Fig. 2 C3\_Transformer

在 Transformer block 中,首先使用一个 ConvBN-SiLU 层进行特征提取,然后是一个 Linear 层进行特征变换,Linear 层的作用是对输入特征进行线性变换。它通过将输入特征映射到不同的表示空间来改变特征的维度和表达能力,从而提供给后续的操作更丰富的信息,如公式(1)所示。接着,通过将输入与 Linear 层的输出进行 Add 操作,将特征传递到下一个阶段的 Transformer layer 中,Add 操作将输入特征与经过线性变换后的特征相加,形成了残差连接。多个 Transformer layer 被堆叠在一起以增加模型的深度,通过堆叠多个 Transformer layer 可以增加模型的深度,提升模型的表达能力、非线性能力和长距离依赖建模能力,如式:

$$\text{Linear}_i(X) = W_i X + b_i, i = 1, 2, 3 \quad (1)$$

其中,  $X$  是输入特征;  $W_i$  和  $b_i$  分别是第  $i$  个 Linear 操作的权重和偏置。

多头注意力机制具有捕捉输入序列中不同位置之间的依赖关系并加强全局信息的交互的作用。多头注意力机制包括三个关键步骤:查询(Query)、键

(Key) 和值(Value) 的线性变换、计算注意力权重以及权重和值的加权求和,如式:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

其中,  $Q$  表示查询的线性变换结果;  $K$  表示键的线性变换结果;  $V$  表示值的线性变换结果;  $d_k$  表示键的维度。通过将查询与键进行相似度计算,然后通过 softmax 函数得到注意力权重,最后将注意力权重与值相乘,得到多头注意力机制的输出。最终,多头注意力机制的输出表示为多个注意力头的加权结果的拼接,如式:

$$\text{MultiHeadAttention}(Q, K, V) = \text{concat}(head_1, head_2, \dots, head_h) W_o \quad (3)$$

其中,  $h$  表示注意力头的数量;  $head_i$  表示第  $i$  个注意力头的输出结果;  $W_o$  是输出的线性变换权重。

通过 Transformer 与主干部分的有效融合, YOLOv5 能够更好地利用全局上下文信息,提高目标检测算法在复杂场景下的性能。Transformer 的自注意力机制和线性变换层使模型能够捕捉输入特征之间的关系,并提供更丰富的表示能力,从而改善目标的定位和识别准确性。

#### 3.2 解耦头

IoU-Net<sup>[16]</sup> 和 Double-Head R-CNN<sup>[17]</sup> 通过引入额外的头部或解耦头部的方法来解决目标检测中分类和定位任务之间的冲突。然而,这些方法尽管采取了一些措施来分离任务,但仍然依赖于从相同的候选边界框中提取特征。为了消除任务之间的冲突,任务感知空间解缠(Task-Aware Spatial Disentanglement, TSD) 采用了解耦的策略,为分类和定位任务提供了任务特定的特征表示。TSD 的基本思想是在神经网络模型中引入空间解耦模块,通过学习将图像分解为任务相关的子图像和任务无关的子图像,通过将特征提取过程分离为任务特定的子特征提取器,能够消除分类和定位任务之间的权衡和冲突。这种简单而有效的设计使得 TSD 能够显著提高性能,因为它允许网络专注于每个任务的特定需求,从而提供更准确的分类和定位结果,如图 3 所示。此外,两个头部结构(即全连接头和卷积头)在不同的任务上具有特异性,全连接头(Fc-head)更适合分类任务,而卷积头(Conv-head)更适合定位任务。

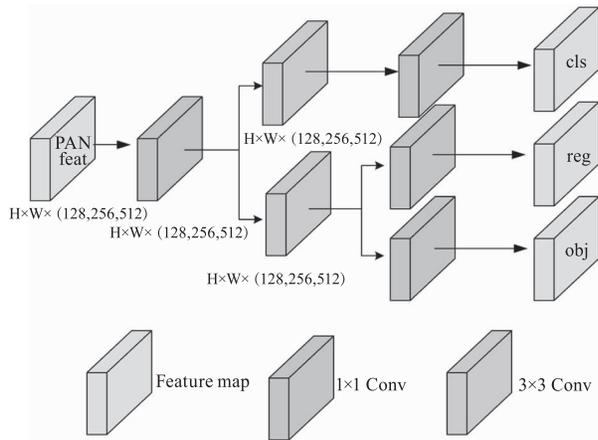


图 3 Decoupled head

Fig. 3 Decoupled head

在 YOLOv5 模型中,检测头部具有耦合设计,分类和定位分支之间共享参数。这意味着两个分支使用相同的卷积层集,这限制了每个分支的灵活性。相比之下,像 FCOS<sup>[18]</sup> 和 YOLOX<sup>[19]</sup> 这样的模型解耦了检测头部的分类和定位分支,使得每个分支都有自己的参数集,增加了模型的灵活性。此外,FCOS 和 YOLOX 在每个分支中引入了两个额外的  $3 \times 3$  卷积层,以提升性能。YOLOv6<sup>[20]</sup> 采用了一种混合通道策略来构建更高效的解耦头部,它将中间的  $3 \times 3$  卷积层数量减少到仅有一个,简化了架构。

传统的 YOLOv5 在分类和定位任务之间共享特征提取器,导致任务之间的权衡和冲突。解耦头通过为分类和定位任务引入独立的特征提取器,使每个任务能够更好地关注自己所需的信息。通过独立的特征提取器,可以捕捉到每个任务所需的特定信息,从而提高分类和定位的准确性。

### 3.3 simOTA

在目标检测中,存在一些边界区域的 anchor box,它们位于正样本和负样本之间。如果一个边界区域的 anchor box 同时包含多个目标的一部分,那么判断它是否为正样本以及属于哪个目标的正样本,就需要考虑全局信息来进行划分。最优传输分配(Optimal Transport Assignment, OTA)<sup>[21]</sup> 将标签分配问题形式化为最优传输问题。在 OTA 中,将每个真实对象和背景视为标签的供应者,将每个 anchor box 视为标签的需求者,目标是找到最优传输方案,以最小化供应者和需求者之间的运输成本,同时考虑正负标签,运输成本定义为对分类和回归损失的成对加权求和,如图 4 所示。

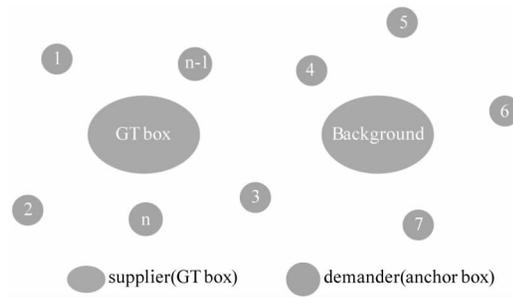


图 4 最优传输问题

Fig. 4 Optimal transport assignment

当将 OTA 算法应用于目标检测中的样本分配时,其主要目标是将预测的边界框与真实目标框(Ground Truth boxes, GT)进行最优匹配。这种匹配过程是基于两个关键方面的考虑:物体之间的相似性和位置之间的关系。通过使用 OTA 算法进行目标检测中的样本分配,可以充分考虑物体之间的相似性和位置关系,从而提高目标检测模型的准确性和稳定性。这种方法能够根据不同目标的特征和位置信息进行动态样本分配,有助于有效地处理目标检测中的复杂场景和多目标情况。假设有两个概率分布  $P$  和  $Q$ ,其中  $P$  表示源分布(预测的边界框分布), $Q$  表示目标分布(真实的边界框分布)。目标是找到一个最优的转移方案,以最小化总的转移成本。Sinkhorn 算法的核心思想是通过迭代调整两个分布的权重向量,使得它们的转移方案逐步收敛到最优解。详细步骤如下:

1) 定义一个成本矩阵  $C$ ,其中  $C[i, j]$  表示将  $P$  中的第  $i$  个元素转移到  $Q$  中的第  $j$  个元素的成本。

2) 初始化两个正数向量  $u$  和  $v$ ,它们的长度与  $P$  和  $Q$  的元素个数相等。这些向量用于调整成本矩阵。

3) 设置一个收敛条件,如迭代次数达到一定阈值或者权重向量的变化小于某个阈值。

4) 迭代更新  $u$  和  $v$ ,直到满足收敛条件为止。对于每个元素  $i$ :

$$u[i] = \log(\text{sum}(\exp((-C[i, j] + v[j])/\text{epsilon}))) \quad (4)$$

其中,epsilon 是一个正数,用于控制更新的平滑度。对于每个元素  $j$ :

$$v[j] = \log(\text{sum}(\exp((-C[i, j] + u[i])/\text{epsilon}))) \quad (5)$$

其中,exp 是指数函数;log 是自然对数函数。

5) 根据更新后的  $u$  和  $v$  计算最终的转移方案。转移方案可以通过以下公式计算:

$$T[i,j] = \exp((u[i] + v[j] - C[i,j]) / \epsilon) \quad (6)$$

6) 根据转移方案  $T$ , 可以计算出  $P$  中的每个元素与  $Q$  中的元素之间的匹配关系。

Sinkhorn 算法的核心思想是通过多次迭代, 通过调整  $u$  和  $v$  来逐步优化转移方案, 使得总的转移成本最小化。然而, Sinkhorn 算法求解 OTA 问题会增加 25% 的训练时间, 在进行多个 epoch 的训练任务中显得额外缓慢。为了简化了这个过程, 使用 simOTA 的动态 top-k 策略, 以获得近似解。simOTA 首先计算每个预测和真实对象之间的配对匹配程度, 可以使用成本来表示, 成本由分类损失和回归损失组成, 通过一个平衡系数进行加权, 如式:

$$c_{ij} = L_{ij}^{cls} + \lambda L_{ij}^{reg} \quad (7)$$

其中,  $\lambda$  为平衡系数;  $L_{ij}^{cls}$  表示预测框与真实框之间的类别损失;  $L_{ij}^{reg}$  表示当前预测框与真实框之间的回归损失 (IoU)。然后, 针对每个真实对象, 我们选择在固定中心区域内成本最低的前  $k$  个预测作为其正样本。最后, 将这些正样本所对应的区域标记为正样本, 其他区域则标记为负样本。simOTA 不仅减少了训练时间, 还避免了 Sinkhorn 算法中的额外求解器超参数。

## 4 实验准备与结果分析

### 4.1 数据集与预处理

潜艇数据集是一个包含不同区域的不同港口的 1000 张图像的潜艇目标检测数据集, 每张图像的分辨率为  $1024 \times 1024$ , 位深度为 24。首先, 使用数据增强技术扩充数据集, 通过旋转、缩放、平移、镜像翻转等操作生成更多的训练样本, 以增加数据的多样性和泛化能力。其次, 需要对图像进行标注和标签, 将潜艇目标的位置和边界框进行标注, 并为每个目标赋予相应的类别标签。然后, 可以将数据集划分为训练集、验证集和测试集, 用于模型的训练、调参和最终性能评估。在进行目标检测之前, 需要对图像进行预处理, 包括调整图像尺寸、归一化、亮度和对比度调整等操作。训练过程中的部分 batch 展示如图 5 所示。

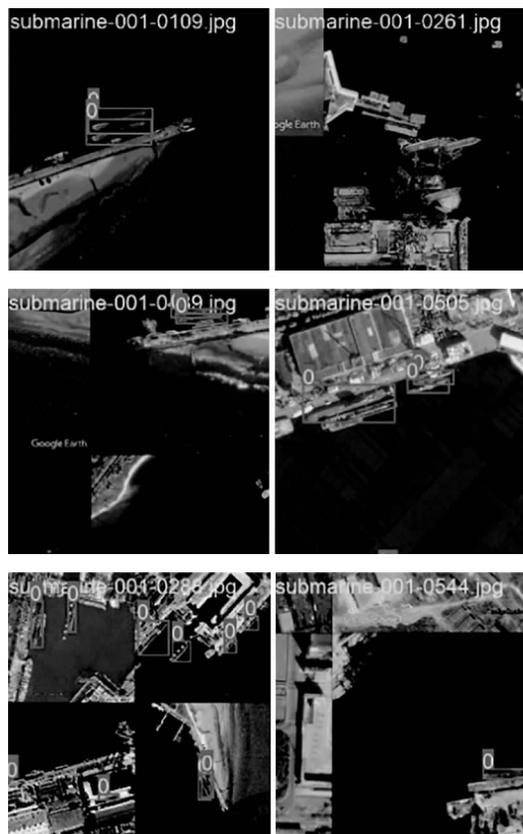
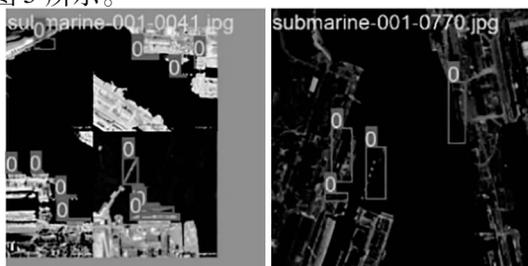


图 5 数据预处理效果展示

Fig. 5 Display of data preprocessing effect

### 4.2 训练参数设置

为了满足任务需求和计算资源的限制, 使用 YOLOv5s 作为基础模型架构。模型训练的总轮数 (epochs) 设置为 80, 每次迭代中用于训练的样本数量 (batch-size) 为 32。为了平滑调整学习率并提高模型性能, 采用余弦退火策略, 权重衰减 (weight decay) 为 0.00025, 以控制模型的正则化程度, 避免过拟合。初始学习率为 0.0031, 并使用 SGD 优化器进行参数优化, 动量因子为 0.74747。

### 4.3 训练结果

#### 4.3.1 训练轮次对 LOSS 的影响分析

在 YOLOv5 中, 通过损失函数衡量模型预测和真实目标之间的差异, 指导模型的参数优化和训练过程。该损失函数主要由两个部分组成: 目标定位损失 (Localization Loss) 和目标置信度损失 (Confidence Loss)。通过优化损失函数, YOLOv5 能够更好地理解图像中的目标并进行准确的定位和分类, 从而实现高效的目标检测。

如图 6 所示, 改进算法在相同的训练轮次内展现出更低的训练损失曲线。这意味着改进算法在训练过程中能够更好地适应训练数据, 并更有效地优化模

型参数。具体而言,观察到改进算法的训练损失曲线在训练过程中下降的速度更快,并且最终达到了更低的损失值。这表明改进算法能够更准确地学习目标特征,并更好地理解目标检测任务的规律。通过在相同的训练时间内实现更低的损失,改进算法在相对较短的时间内提供了更强的学习能力和收敛性。

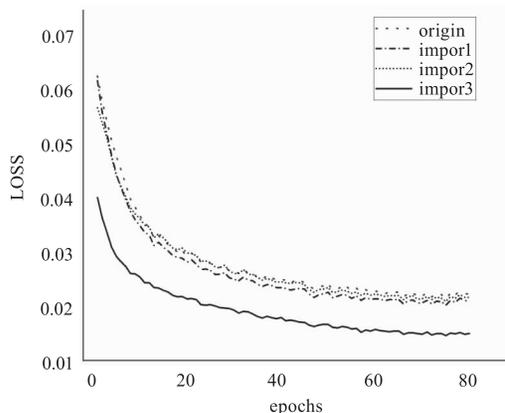
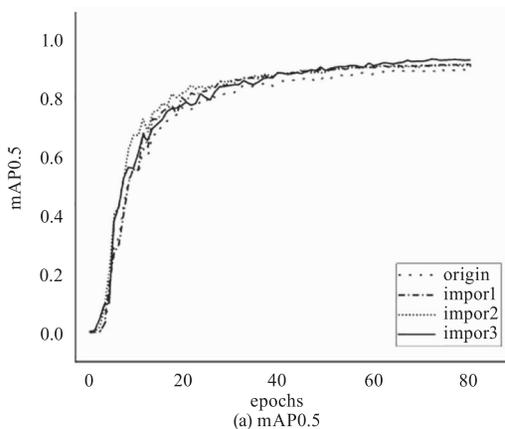


图 6 改进前后 YOLOv5 训练损失曲线对比图  
Fig. 6 Comparison of YOLOv5 training loss curve before and after improvement

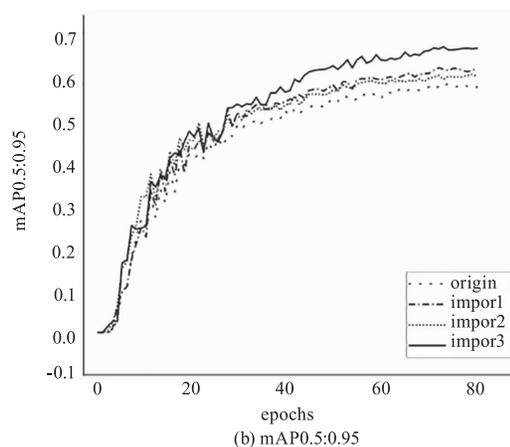
#### 4.3.2 训练轮次对 mAP 的影响及分析

在本实验中,我们使用 YOLOv5 作为目标检测算法的基线。针对潜艇目标进行检测,  $mAP@0.5$  衡量了模型在预测框与真实框之间 IoU 大于 0.5 时的平均准确率,而  $mAP@0.5:0.95$  则考虑了 IoU 在 0.5 到 0.95 范围内的所有值,我们通过记录和比较  $mAP@0.5$  和  $mAP@0.5:0.95$  指标的变化,来评估改进前后算法的优越性,如图 7 所示。

改进算法在这两个指标上都取得了提升,意味着它相对于原始算法在目标检测任务中更为有效。 $mAP@0.5$  的提升表示改进算法在精确定位和识别目标方面的性能得到了改善。而  $mAP@0.5:0.95$  的提升则说明改进算法能够在不同大小和形状的目标上取得更好的检测结果,具有更强的泛化能力。



(a)  $mAP@0.5$



(b)  $mAP@0.5:0.95$

图 7 改进前后模型的  $mAP@0.5$  与  $mAP@0.5:0.95$  随 epochs 曲线图  
Fig. 7  $mAP@0.5$  and  $mAP@0.5:0.95$  with epochs before and after the improvement

#### 4.3.3 评价指标对比

Precision 衡量了模型在预测框中的正例中真实目标的比例, Recall 衡量了模型能够正确检测出真实目标的能力。我们对不同改进程度的 YOLOv5 模型进行了性能评估,并得到了相应的 Precision、Recall、 $mAP@0.5$  和  $mAP@0.5:0.95$  的结果:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

其中,真正例(TruePositives, TP),指模型正确预测为正例的样本数;假负例(FalseNegatives, FN),指模型未能正确预测为正例的样本数;假正例(FalsePositives, FP),指模型错误地预测为正例的样本数。

$$AP = \sum_{n=1}^N (R_n - R_{n-1}) \cdot P_n \quad (10)$$

其中,  $R_n$  是第  $n$  个点的召回率(Recall, R);  $R_{n-1}$  是第  $n-1$  个点的 Recall;  $P_n$  是第  $n$  个点的精确率(Precision, P)。

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP_i \quad (11)$$

$$mAP@0.5:0.95 = \frac{1}{N} \sum_{i=1}^N AP_i^{interp} \quad (12)$$

其中,  $N$  表示类别的总数;  $AP_i$  表示第  $i$  个类别的 Average Precision;  $AP_i^{interp}$  表示对第  $i$  个类别的 Precision-Recall 曲线进行插值后的 Average Precision。

如表 1 所示,首先,我们注意到改进后的算法在 Precision 方面表现出更好的性能。Precision 的提升意味着改进算法能够更准确地识别出真正的正样本,减少了误报的情况。这证明改进算法在目标检

测任务中具有更好的准确性和可靠性。其次, Recall 指标也显示了改进后的算法的提升。Recall 衡量了算法能够检测到的真实正样本的比例。改进算法在召回率方面的改善意味着它能够更好地捕捉目标,减少了漏检的情况。此外,mAP@0.5 和 mAP@0.5:0.95 这两个指标都显示出了改进算法相对于未改进算法的提升。mAP@0.5:0.95 是对不同 IoU 阈值下平均精度的度量,而改进算法在这两个指标上的提升意味着在不同 IoU 阈值范围内都能够更好地预测和定位目标。

表 1 消融实验

Tab.1 Ablation experiment

Model	P	R	mAP0.5	mAP0.5:0.95
origin	0.909	0.773	0.905	0.589
+ C3TR	0.894	0.826	0.918	0.625
+ Decoupledhead	0.905	0.829	0.921	0.611
+ simOTA	0.934 ↑	0.857 ↑	0.939 ↑	0.676 ↑

综上所述,通过对改进前后的算法进行实验和指标对比分析,我们验证了改进算法在 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 等指标上的优越性。这些结果进一步验证了改进算法在目标检测任务中的性能提升和有效性。

#### 4.3.4 更加聚焦于前景

我们使用了特定的颜色编码来表示目标检测模型在不同位置对于潜艇目标的关注程度或重要性,如图 8 所示。

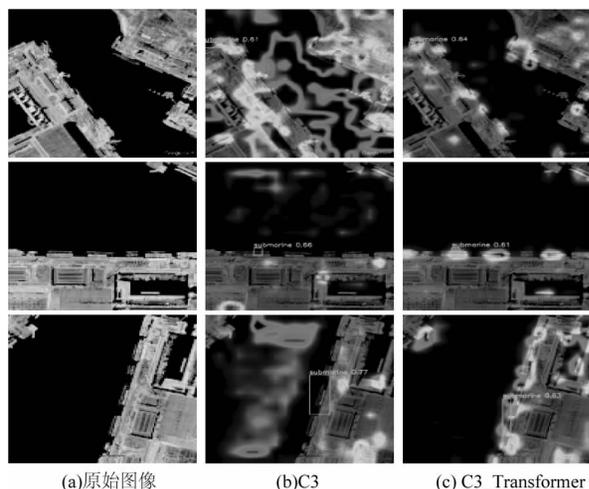


图 8 改进前后显著性区域

Fig.8 Significance area before and after improvement

在改进前的 YOLOv5 中,热力图显示出较为平

均的颜色分布,没有明显的高亮区域,注意力更加分散,意味着算法在整个图像上均匀关注。然而,在改进后的 YOLOv5 中热力图显示出更明显的高亮区域,集中在目标的位置上,更加集中地关注目标的重要区域。这意味着模型更有在目标的关键部分进行定位和分类,从而提高了准确性和鲁棒性。

通过对比改进前后的热力图,我们可以观察到关注点分布的差异。改进后的热力图显示出更集中、更明亮的区域,与目标位置更吻合,表明改进后的 YOLOv5 在目标检测中更加准确地关注目标区域。这种注意力的改变有助于提高模型的性能,因为它能够更有针对性地处理目标的特征,提高目标的定位和分类准确性。通过更好地关注目标的关键部分,改进后的 YOLOv5 可以更好地识别目标,减少误报和漏报的情况,提高整体的检测性能。

#### 4.3.5 改进模型效果对比

在实验部分,我们对潜艇目标检测任务使用了改进前后的 YOLOv5 模型,并通过图像展示了其对潜艇图像的识别效果。通过对测试数据集使用改进前后的模型进行测试,并将同一幅图像的检测结果显示出来,可以更好地说明改进后的模型在目标检测任务中的性能提升,如图 9 所示。

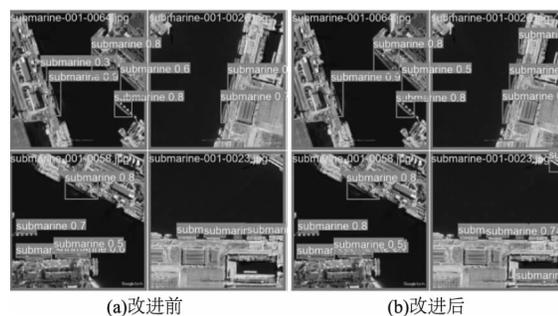


图 9 改进前后 YOLO 预测结果

Fig.9 YOLO prediction results before and after improvement

对于改进前的模型,检测结果会显示出一些误报和漏报的情况。例如,会将一些背景区域错误地识别为目标,或者未能正确地检测到一些真实的目标。这些错误与模型的关注点分布、定位准确性或分类能力等因素有关。然而,通过使用改进后的模型进行测试,我们可以观察到检测结果的改善。在对图 9(a)、(b)右下角以及右下角对比中可以发现原始 YOLOv5 漏检的目标在改进之后得到明显的改善,解决了模型的漏检问题。改进后的模型能更准确地定位和分类目标,减少了误报和漏报的情况。

#### 4.3.6 典型算法类比

在实验中,首先计算了每个模型的 P、R 和 mAP@0.5 值。然后,我们将这些结果进行对比,以展示改进后的 YOLOv5 相对于其他算法的优异性,如表 2 所示。

表 2 不同算法指标评价

Tab.2 Evaluation of different algorithm indexes

Model	P	R	mAP0.5
Faster-RCNN	0.826	0.659	0.698
SSD	0.746	0.719	0.725
YOLOv3	0.728	0.625	0.814
YOLOv5-6.0	0.883	0.702	0.876
YOLOv5	0.909	0.773	0.905
Ours	0.934 ↑	0.857 ↑	0.939 ↑

相对于 Faster-RCNN、SSD、YOLOv3 和原始 YOLOv5,改进后的 YOLOv5 在 P、R 和 mAP@0.5 上取得了更高的数值。较高的精确率意味着更少的误报,即更少地将背景误判为目标,从而提高了目标检测系统的准确性。而较高的召回率意味着更少的漏报,即更少地将实际目标错判为背景,增加了目标检测系统的完整性。通过提高精确率、召回率和平均精度的数值,改进后的 YOLOv5 显示出了在目标检测任务中更好的性能。这些结果进一步证明了我们所提出的改进方法对于潜艇目标检测具有优异性和有效性。

## 5 结论

本文主要研究了基于改进 YOLOv5 在潜艇目标识别领域的应用,以提升海军潜艇发现与跟踪的准确性与效率。通过改进 YOLOv5 潜艇目标检测器,有效解决了传统潜艇目标检测算法的局限性。实验结果表明,相较于原始 YOLOv5,改进后的模型在 Precision、Recall、mAP0.5 和 mAP0.5 : 0.95 指标上分别提高了 2.8%、10.9%、3.8% 和 14.7%。同样,相较于原始 YOLOv5 以及其他常用算法如 Faster-RCNN、SSD 和 YOLOv3,在 Precision、Recall、mAP0.5 和 mAP0.5 : 0.95 指标上,改进后的模型分别取得了显著的提升。这说明改进后的模型在目标检测的准确性、召回率以及在不同置信度阈值下的平均准确率等方面取得了明显的进步,并且改善了原始模型在漏检问题上的表现。与其他算法相比,改进后的模型展现出更好的性能表现,有

效地解决了传统算法对复杂背景和噪声的鲁棒性差、对光照和视角变化敏感以及难以处理大规模数据集等问题。

综上所述,改进后的模型相较于原始 YOLOv5 以及其他常用算法,在潜艇目标检测任务中取得了显著的提升,不仅在与原始模型的对比中显示出更好的性能,而且在与其它算法的对比中也展现出更高的准确性和召回率。这些结果为潜艇目标识别技术的发展提供了重要的参考和推动,具有实际应用的潜力和推广价值。

## 参考文献:

- [1] Yu K, Park D, Choi J, et al. Effect of skew on the tonal noise characteristics of a full-scale submarine propeller [J]. *Ocean Engineering*, 2023, 276: 114218.
- [2] Pulido Azpíroz A. Maritime front on the basque coast, a battlefield in the rear of the great war: U-boats and espionage [J]. *The International History Review*, 2023; 45: 478-495.
- [3] Wang R, Yao Z, Qi X B, et al. Applications of supercavitation drag reduction technology in naval gun underwater attack and defense warfare [J]. *Journal of Physics: Conference Series*, 2023, 2460(1): 012030.
- [4] Yamada K. Mobilizing citizens' ears: aural training as civil defense, 1941-1945 [J]. *Technology and Culture*, 2023, 64(2): 359-378.
- [5] XU Yinghao. Simulation research on sea surface reflection and target characteristics of full appendage submarine [D]. Harbin: Harbin Institute of Technology, 2021. (in Chinese)  
许英豪. 全附体潜艇海面映波及目标特性仿真研究 [D]. 哈尔滨: 哈尔滨工业大学, 2021.
- [6] Zheng H, Li D, Huang X, et al. Underwater submarine image recognition based on improved faster R-CNN algorithm [C] // 2021 IEEE 3rd International Conference on Computer Communication and the Internet (ICCCI), 2021.
- [7] Ghosh D, Paul A, Ghosh S. A deep learning based approach for underwater target detection and classification [J]. *IEEE Access*, 2021, 9: 73923-73934.
- [8] Chen Y, Fu K, Lin Y, et al. Real-time defect detection for large-scale industrial products using YOLO-based convolutional neural network [J]. *Journal of Manufacturing Systems*, 2020, 58, 247-257.

- [9] Li Y, Xie L, Cai C, et al. Research on aircraft detection based on YOLO-v3. In 2020 3rd International Conference on Image [ C ] // Vision and Computing (ICIVC), IEEE, 2020:1 – 5.
- [10] Raza A, Li J, Zhang M. A novel approach to detect defects in welding joints using YOLO-V3 and canny edge detection [ J ]. IEEE Access, 2021, 9:59823 – 59832.
- [11] Ultralytics. YOLOv5: a unified framework for object detection [ J/OL ]. arXiv preprint. 2020, arXiv:2004.10934.
- [12] Wang A, Zhang X. CSPDarknet53: a lightweight convolutional neural network for object detection [ J/OL ]. arXiv preprint. 2020, arXiv:1911.11929.
- [13] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [ C ] // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017:2117 – 2125.
- [14] Liu S, Qi L, Qin H. Path aggregation network for instance segmentation [ C ] // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018:8759 – 8768.
- [15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [ J ]. In Advances in Neural Information Processing Systems. 2017:5998 – 6008.
- [16] Borui Jiang, Ruixuan Luo, Jiayuan Mao, et al. Acquisition of localization confidence for accurate object detection [ C ] // Computer Vision and Pattern Recognition (CVPR). 2018.
- [17] Yue Wu, Yinpeng Chen, Lu Yuan, et al. Rethinking classification and localization for object detection [ C ] // Computer Vision and Pattern Recognition (CVPR). 2020.
- [18] Tian Z, Shen C, Chen H, et al. FCOS: fully convolutional one-stage object detection [ C ] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2020.
- [19] Zheng Ge, Songtao Liu, Feng Wang, et al. YOLOX: exceeding YOLO series in 2021 [ C ] // Computer Vision and Pattern Recognition (CVPR). 2021.
- [20] Chuyi Li, Lulu Li, Hongliang Jiang. YOLOv6: a single-stage object detection framework for industrial applications [ C ] // Computer Vision and Pattern Recognition (CVPR). 2022.
- [21] Zheng Ge, Songtao Liu, Zeming Li, et al. OTA: optimal transport assignment for object detection [ C ] // Computer Vision and Pattern Recognition (CVPR). 2021.