

基于改进 YOLOv7 的水下小目标检测算法研究

杜 锋

(江苏电子信息职业学院, 江苏 淮安 223003)

摘要: 目标检测研究一直是水下小目标检测的难题; 针对水下小目标检测任务漏检率高、水下场景识别效果差的问题, 提出一种利用 YOLOv7 改进的水下小目标检测技术; 为了达到准确率的同时兼顾高检测速度, 采用 YOLOv7 网络作为基础网络; 该网络通过融合 SENet 注意力机制、增强 FPN 网络拓扑、结合 EIoU 损失函数, 集中小目标更关键的特征信息, 提高检测精度, 同时降低模型复杂度; 通过模拟测试, 在测试集上确认了 mAP 、 P 和 R 指标, 并与其他传统目标检测技术进行了对比; 结果表明, 增强的算法优于竞争网络, 并成功提高了测试集的检测精度。

关键词: YOLOv7; 水下小目标检测; 注意力机制; FPN; EIoU

Research on Underwater Small Target Detection Algorithm Based on Improved YOLOv7

DU Feng

(Jiangsu Vocational College of Electronics and Information, Huai'an 223003, China)

Abstract: Research on target detection has always been a challenge for underwater small target detection. To address the issues of high miss detection rate and poor underwater scene recognition in underwater small target detection tasks, an improved underwater small target detection technique based on YOLOv7 is proposed. In order to achieve the accuracy and balance the detection speed, the YOLOv7 network is adopted as the basic network. By fusing the SENet attention mechanism, enhancing the FPN network topology, and incorporating the EIoU loss function, the crucial feature information of small targets is concentrated in the network to increase the detection accuracy while reducing the complexity of the model. Through simulation tests, the indexes of mAP and P as well as R are confirmed on the test set, and compared with other conventional target detection techniques, the results show that the enhanced algorithm is superior to the competing networks and successfully improves the detection accuracy on the test set.

Keywords: YOLOv7; underwater small target detection; attention mechanism; FPN; EIoU

0 引言

目标检测是计算机视觉领域的一个重要任务, 旨在识别和定位图像或视频中的特定物体或目标。与图像分类任务不同, 目标检测不仅要确定图像中是否存在目标, 还要准确地确定目标的位置, 通常是通过边界框来描述目标的位置和大小。

小目标检测是目标检测领域的一个重要分支, 专注于在图像或视频中检测和定位尺寸较小的目标, 这些目标通常具有低分辨率、低对比度或受到遮挡等挑战。小目标检测的关键挑战包括有效地提取小目标的特征、准确地定位小目标的位置以及抑制背景干扰等。针对这些挑战, 研究者们提出了一系列针对小目标的特定算法和技术, 包括使用更加敏感的特征提取方法、设计更加精细的目标定位策略以及优化模型架构以适应小目标的特性。小目标检测在

许多领域具有广泛的应用, 如无人机监测、微小物体识别等, 对于提高目标检测系统的全面性和实用性具有重要意义。

水下小目标检测是指在水下环境中, 针对尺寸较小的目标进行检测和定位的技术, 如水下生物、潜水器材等。该技术在海洋科学研究、水下探测、海洋资源开发等领域具有重要应用, 可用于水下生物监测、水下考古勘探、海底管道检测等任务, 为海洋工程和科学研究提供了关键支持, 有助于保护海洋生态环境和推动海洋资源的可持续利用。

与标准陆地区域不同, 复杂的水下环境、光线条件和较小的目标给水下人工作业带来了巨大的挑战。水下目标检测^[1]旨在定位和识别水下场景中的物体, 这项研究可以使用远程操作潜水器 (ROV)^[2] 和自主水下航行器 (AUV)^[3] 等水下设备进行, 这使得定位和识别特定的小型

收稿日期: 2024-03-19; 修回日期: 2024-04-30。

基金项目: 2023 年江苏省产学研合作项目 (BY20231025)。

作者简介: 杜 锋 (1977-), 男, 硕士, 副教授。

引用格式: 杜 锋. 基于改进 YOLOv7 的水下小目标检测算法研究[J]. 计算机测量与控制, 2024, 32(9): 108-117.

目标养殖生物成为可能。水下目标检测对于开展水下生物监测、代替人类捕鱼等活动提供了技术支持，解决了水下人工作业的难题。

小目标检测精度差、误检漏检现象普遍，一直是水下目标检测的问题。本研究提出了一种基于 YOLOv7 (You Only Look Once v7) 的增强型水下小目标生物识别算法^[4] Under Water Small Critter-YOLOv7 (UWSC-YOLOv7) 检测框架克服了这些问题。为了减少错误和漏检，同时提高检测精度，对网络进行了 3 种修改：通道特征信息增强、多尺度检查检测^[5]和损失函数。使用 mAP (平均精度) 值评估网络的性能，随后在国家水下机器人竞赛数据集上对改进的模型进行训练和测试，以衡量算法的有效性。它比其他网络实现了更高的性能，P 指标得分为 84.91%，R 指标得分为 80.73%，mAP 指标得分为 83.76%。

UWSC-YOLOv7 算法作为一项研究，通过增强 YOLOv7 网络来提高水下小目标识别的效率，以提高检测精度并在现有研究成果的基础上获得更好的结果。基于此，它可以被潜艇用于真实的水下领域探测，支持人类水下作业，全面定位和识别小型目标生物^[6]。该算法的功能还将实现各种专业用途，包括深海研究和浅海养殖^[7]。

1 水下小目标检测

基于深度学习的目标检测方法主要分为两阶段模型和单阶段模型^[8]。R-CNN (具有卷积神经网络特征的区域^[9]、Fast R-CNN^[10]和 Faster R-CNN^[11]等算法就是前者的例子。两种最流行的单级变体是 YOLO^[12]和 SSD (单次多盒检测器)^[13]。在传统的路面目标检测领域，上述方法均显示出良好的效果。近年来，这些相同的算法越来越多地用于水下情况。此外，水下物体识别数据集的缺乏限制了水下物体检测方法的进步。

与传统的路面检测相比，水下目标识别的历史很晚。在 URPC2017 (2017 年水下机器人专业大赛) 数据集上，针对 3 种类型的海参、海胆和海星目标的 mAP 为 46.3%，Chen 等人^[14]提出了一种用于水下目标小样本检测的 SWIPENet。然后，该网络使用 URPC2018 数据集，在 2018 年添加了 mAP 为 64.5% 的海星目标。林等人^[15]使用 RoIMix 的增强方法在 2019 年改进了 Faster R-CNN，在 URPC2018 数据集上达到了 74.92% 的 mAP。2020 年，刘等人^[16]增强了 YOLOv3 网络，并创建了与 DG-YOLO 配对的 Water Quality Transfer，在 URPC2019 数据集上获得了 63.83% 的 mAP。徐等人^[17]认为 SA-FPN (Scale-Sware-Feature Pyramid Network) 在 PASCAL VOC 数据集中达到 76.27% mAP。通过增强 SSD 网络，Jiang 等人^[18]能够在 URPC2018 数据集上达到 66.9% 的 mAP。Han 等人使用 Deep CNN 技术完成了海参、海胆和扇贝 3 个目标^[19]，实现了 91.2% 的 mAP。Wang 等人^[20]对 YOLOv5 模型进行了升级。2021 年，他们使用 YOLOv5x 在 URPC 数据集上获得了 69.3% 的 mAP。在 URPC2020 数据集上，Chen 等

人^[21]升级了 YOLOv4 并创建了 YOLOv4-uw (水下) 网络模型达到 75.34% mAP。在水下测试数据集上，Zeng 等人^[22]使用更快、更好的 RCNN 网络进行验证，并实现了 72.1% 的 mAP。为了检测独特的鱼类，Li 等人^[23]在 2022 年提出了 CME-YOLOv5 网络模型，并达到了 92.3% 的 mAP。Yu 等人^[24]提出了 mAP 为 79.6% 的多注意路径聚合网络 (APAN)。2023 年在 URPC2019 数据集上，Wen 等人^[25]提出了 YOLOv5s-CA (协调注意力) 网络，mAP 为 80.9%。Liu 等人提出了 YOLO 的 TC 网络模型^[26]，它在 RUIE2020 (真实世界水下图像增强) 数据集上实现了 83.1% 的 mAP。近年来提出的算法效果呈现出不断改进的格局，CNN、SSD、YOLO 等改进算法更广泛地应用于水下动物的目标检测算法。此外，URPC 数据集有 4 种不同类别的目标，即使数据集中的目标数量随着年数的增加而增加，它们从一开始就保持不变。

现在最常见的研究问题是小目标问题。2014 年，在通用数据集 MSCOCO^[27]中，分辨率小于 32×32 像素的目标被归类为小目标。继 Lim 等人之后，更多的人研究了小目标问题。^[28]在 2021 年通过上下文注意机制改进了小目标识别。根据 Cheng 等人的说法^[29]在 2022 年采用了绝对面积要求，面积小于 1 024 像素 (即 $S \leq 1 024$) 的实例被认为是小物体。光照对水下场景的影响显著降低了水下图像的质量^[30]；另一方面，水下生物目标由于拍摄距离长和遮挡问题，有效像素面积较小，因此本文将这些目标定义为小目标物体。在特征提取阶段仅从小目标中提取少量特征，这使得网络很难支持后续的检测过程^[31]，后续将详细介绍针对小目标进行的网络增强。

2 方法设计

图 1 描述了我们在本文中使用的 YOLOv7 网络设计。输入部分经过部分数据增强和尺寸缩放预处理后，将图像发送到 Backbone 对预处理后的图像进行特征提取；将提取的特征送入 Neck 增强特征提取网络进行特征融合，得到 3 种尺寸的特征：大 (20×20)、中 (40×40)、小 (80×80)；接下来是检测步骤。融合后的特征被传送给头部进行检查，并输出结果。

1) SENet 注意力机制：

注意力机制^[32]是一种常见的数据处理方式。通过引入注意力机制可以解决小目标特征不明显的问题。

因为通道注意力机制^[33]会自动学习每个通道的重要性，增强有用的特征并抑制对当前任务没什么用处的特征。通过引入 SENet 注意力机制可以解决小目标特征模糊或丢失的问题，可以更好地保持特征提取过程中容易丢失的小目标的特征。

Jie H 等人在牛津大学开发了挤压和激励网络 (SENet)^[34]，旨在通过模拟卷积特征通道之间的相互依赖性来增强网络的表示能力。SENet 设计如图 2 所示。

SENet 使用刺激和挤压来校准特征。Squeeze 是全局信

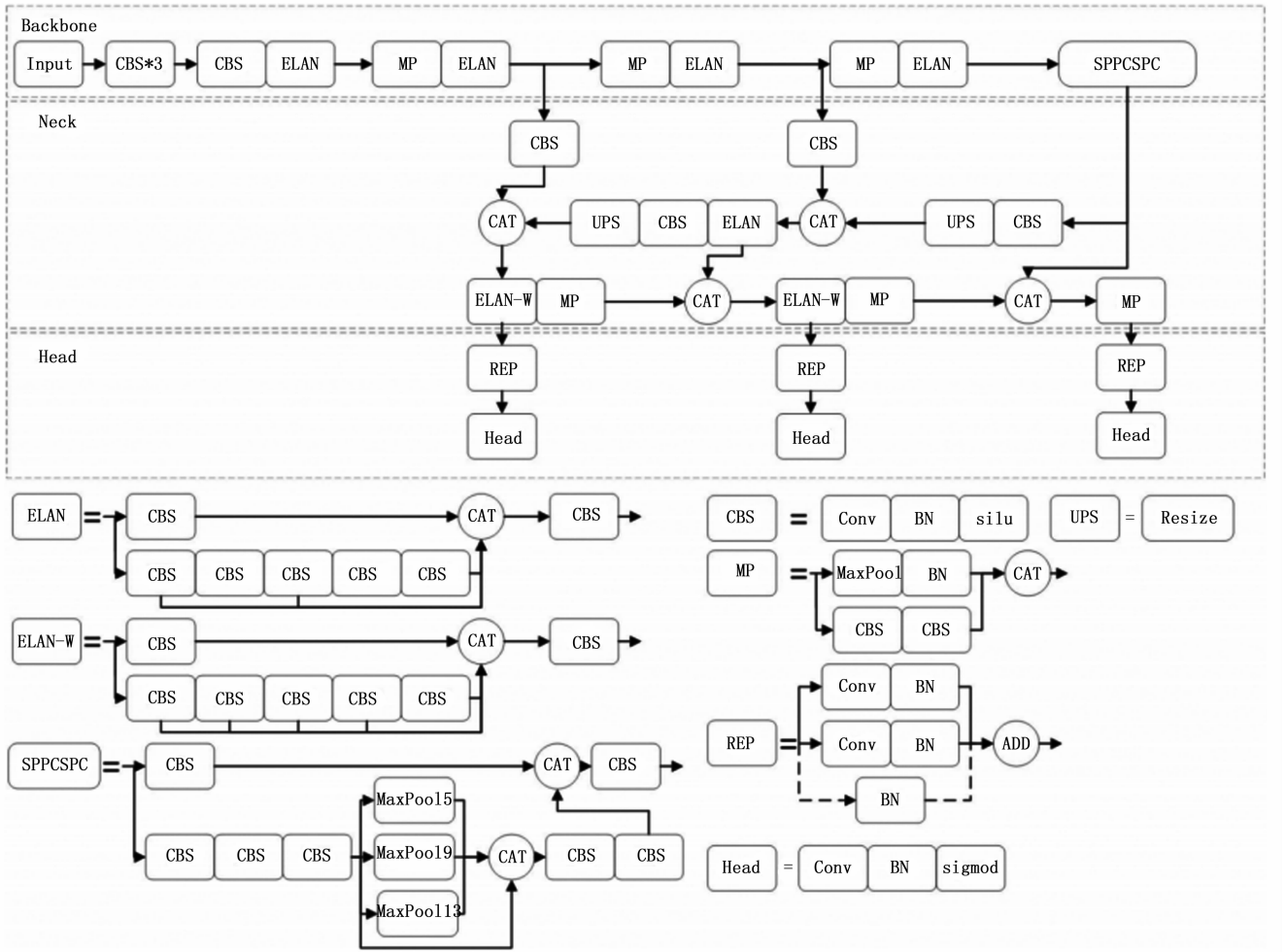


图 1 YOLOv7 模块的架构

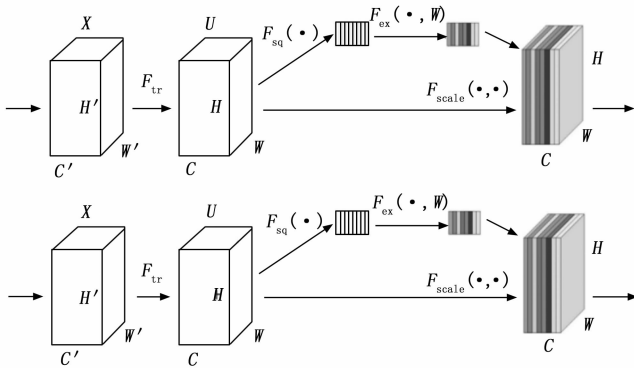


图 2 SENet 模块的架构

息的嵌入，它聚合跨空间维度 $H \times W$ 的特征以生成将 $H \times W \times C$ 维度映射到 $1 \times 1 \times C$ 维度的通道描述符。上述通道描述符已将全局空间信息压缩到其中，以便其输入层可以如公式 (1) 中所述使用它。

$$Z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

刺激是一个自适应调谐过程，其中每个通道通过通道相关性来学习样本特定的激活。它旨在学习如何使用全局信息有选择地强调信息方面并最小化不相关的特征。根据公式 (2)，具有两个完全链接层的瓶颈（指定为 W_1 用于降维， W_2 用于降维）用作门控机制的参数化。

$$s = F_{ex}(z, W) = \sigma[W_2 \delta(W_1 z)] \quad (2)$$

过多的模块添加会导致整个网络的计算效率下降。为了更好地从小目标中提取关键特征信息，本文仅在每个 ELAN 之后必须获得有效特征层的位置添加 SENet^[35]。

2) 强化特征提取网络 FPN:

FPN (Feature Pyramid Network)^[36] 是一种用于目标检测和语义分割等任务的特征提取网络，旨在解决不同尺度目标的定位和识别问题。FPN 通过自顶向下的递归方式构建特征金字塔，将来自不同层次的语义信息融合到同一级别的特征图中，从而使模型能够同时处理不同尺度的目标。FPN 有效地利用了深度卷积网络的特征层次结构，提高了目标检测和语义分割任务的性能和鲁棒性。

随着卷积层数的增加，小目标特征将变得越来越少；然而，处理小目标可以通过合并较小的有效特征层来完成。

FPN 的目的是通过上采样和下采样将小目标特征与大目标特征叠加来保证小目标的特征和信息。

对于特殊融合，YOLOv7 使用了 20×20 、 40×40 和 80×80 三个不同大小的特征有效层。本文^[37]针对小目标添加了 160×160 的第四个有效特征层，并且改进后的 FPN 如图 3 所示。

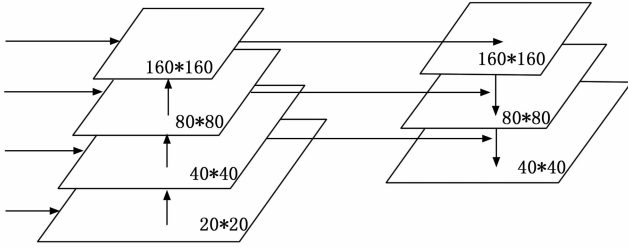


图 3 改进的 FPN 架构

引入 160×160 特征层，它从 Backbone 携带较小的目标信息，并在 160×160 层上分别添加一层上采样和下采样，以在下采样之前融合较小特征层的信息。为了提高模型速度，减少 20×20 层下采样并丢弃具有巨大目标特征的底层。

卷积神经网络对于尺度变化对表现良好，但对于位置变化对表现不佳；为了解决尺度变化问题，实现了多尺度检测。

多尺度检测是在不同尺度的特征图上预测多个帧。升级后的网络消除了 20×20 层特征图检测头以提高网络速度，并使用 YOLOv7 来检测特征图大小为 40×40 、 80×80 和 160×160 的微小目标。图 4 描绘了增强后的网络检测头。

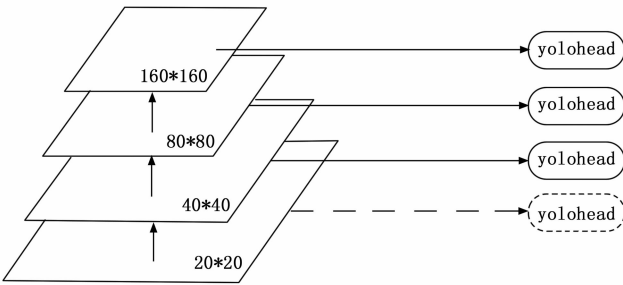


图 4 改进的检测头

3) EIOU 损失函数：

Intersection over Union (IoU) 损失函数是目标检测任务中常用的损失函数之一，用于衡量预测边界框与真实边界框之间的匹配程度。它通过计算两个边界框的交集面积与它们的并集面积之间的比率来评估预测的准确性，即 IoU 值。IoU 损失函数通常用于训练目标检测模型，帮助模型学习正确地定位和识别目标。

根据公式 (3)，IoU，也称为交叉合并比，是目标检测中应用最广泛的统计量。

$$IoU = (Box_t \cap Box_r) / (Box_t \cup Box_r) \quad (3)$$

具体算法的目的是确定反映预测检测框和真实检测框检测效果的正负样本，如图 5 所示。

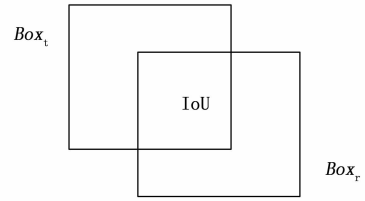


图 5 IoU 图

真实框 (Box_t) 和预测框 (Box_r) 的交集而不是真实框 (Box_t) 和预测框 (Box_r) 的串联是真实框与算法的交集。损失函数的目标是通过使预测帧更接近真实帧来提高预测精度。YOLOv7 使用 CIoU 损失函数。DIoU 和 CIoU 都是由 Zheng 等人提出的^[38]，这有助于使目标框回归更加稳定。通过增加长宽的损失来增加检测框尺度的损失，CIoU 是在 DIoU 的基础上实现的。在公式 (4) 中，显示了 CIoULoss。

$$1 - CIoU = 1 - \left(IoU - \frac{d_0^2}{d_c^2} - \frac{v^2}{1 - IoU + v} \right) \quad (4)$$

式中， d_0 为目标帧与真实帧之间的质心欧式距离， d_c 为目标帧对角线距离， v 为表示长宽比一致性的参数，如式 (5) 所示。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gr}}{h^{gr}} - \arctan \frac{w^p}{h^p} \right)^2 \quad (5)$$

真实目标帧的宽度和高度为 w^{gr} 、 h^{gr} ，预期帧的宽度和高度为 w^p 、 h^p 。在此基础上，结合了重叠面积、质心距离和纵横比这 3 个关键几何元素的损耗。然而，CIoU 也存在两个问题：第一，长宽比表示的相对值存在一定的不确定性；其次，没有考虑挑战性样品的平衡问题。UWSC-YOLOv7 模型中 EIoU 损失取代了 CIoU 损失来解决这两个问题。

EIOU (Extended Intersection over Union) 损失函数是一种用于目标检测任务的损失函数，旨在解决传统 IoU 损失函数对于小目标定位不准确的问题。EIOU 损失函数将目标边界框之间的差异转化为面积差异，通过扩展 IoU 计算方式来准确度量预测边界框和真实边界框之间的匹配程度，从而提高了小目标检测的准确性和鲁棒性。

张等人^[39]提出 ELoU 函数分离基于 CIoU 的纵横比，并添加具有公式 (6) 中描述的 EIoULoss 的焦点质量锚框。

$$1 - \left(IoU - \frac{d^2(b^p, b^{gr})}{(w^r)^2 + (h^r)^2} - \frac{d^2(w^p, w^{gr})}{(w^r)^2} - \frac{d^2(h^p, h^{gr})}{(h^r)^2} \right) \quad (6)$$

预测帧和实际帧的最小外框 w 被值覆盖。重叠损失、中心距离损失和宽度高度损失是损失函数的 3 个组成部分。该方法的前两部分是 CIoU 的延续，但宽高损失直接降低了真实帧和预测帧之间的宽高差，加速收敛。

4) UWSC-YOLOv7:

YOLO (You Only Look Once) 是一种实时目标检测算法, 将目标检测任务视为单个神经网络的回归问题, 通过将图像分成网格并在每个网格单元内预测边界框及其类别概率来实现目标检测。其步骤包括将图像分割为固定大小的网格单元、在每个网格单元内预测边界框和类别概率, 以及通过非极大值抑制策略筛选最终的目标检测结果。

YOLO 算法作为 one-stage 目标检测算法最典型的代表, 其基于神经网络进行对象的识别和定位, 运行速度很快, 可以用于实时系统。YOLOv7 是目前 YOLO 系列最先进的算法, 在准确率和速度上超越了以往的 YOLO 系列。

图 6 描述了经过上述改进后的网络。为了改变 FPN 结构并使用 3 个较小的检测头进行微小目标生物检测, 在 Back-bone 网的有效特征层中添加了 4 个 SENet。

卷积核的层数和通道数是决定 YOLOv7 模型复杂度和计算时间的重要因素, 而复杂度和计算时间又受到模型复杂度的影响。例如, 增加网络中卷积核的数量以执行从 256 通道到 512 通道再回到 256 通道的维度转换, 会导致模型更复杂, 计算时间更长。SENet 注意力方法以较低的计算时间成本实现了小目标特征的改进表示, 因为输出等于输入通道的数量。这可以防止增加模型复杂性。为了改进 FPN 而添加 160 × 160 层和删除 20 × 20 层并不会影响模型的复杂度。然而, 随着网络变深, 通道数增加, 160 × 160 层网络的通道数明显少于 20 × 20 网络, 这在一定程度上加快了计算速度。为了计算损失的长宽比, EIoU 损失分割了长宽比, 这将需要更长的时间来计算。尽管有所改进, 整个模型的复杂度并没有改变, 计算时间与之前的模

型相当。

3 实验结果与分析

1) 数据集描述:

2020 年全国水下机器人目标检测算法竞赛提供了 3 个可供实验使用的替代数据集。数据集中包含 5 个小型生物目标, 包括 20 088 个海胆目标、6 184 个海星目标、5 954 个扇贝目标、4 935 个海参目标和 70 个水藻目标, 总共 5 543 张 jpg 格式的水下光学图像。样本照片质量较差, 目标数量分布范围广, 且小目标较多, 数据量过少容易导致过拟合问题。UPRC2021 收藏包括 4 个微小生物目标: 分别是 21, 879 个海胆目标、7 854 个海星目标、6 484 个扇贝目标和 5 471 个海参目标。它还包括 6 575 张 jpg 格式的水下光学照片以及匹配的注释结果。该示例具有普遍存在的问题, 与上述数据集类似。咸水数据集中共包含 1 969 张 png 格式的水下光学图像, 其中还包括 6 个小型生物目标, 其中分别包括 1 809 个螃蟹目标、1 479 个小鱼目标、1 145 个海星目标、460 个鱼类目标、92 个虾目标、87 个水母目标。

结合问题, 本文选择第一个数据集作为改进实验的数据集, 另外两个数据集作为验证使用, 将数据集按以下两种比例划分: (训练集+验证集): 测试集=9: 1; 训练集: 验证集=9: 1。由于 2020 数据集和 UPRC2021 数据集比较相似, 这里仅提供 2020 数据集和 Brackish 数据集样本。样本包括海星、海胆、虾、蟹、小鱼和其他目标。部分数据集样本如图 7 所示。

2) 评估标准:

实验上, 使用 Precision (P)、Recall (R) 和 Average Precision (AP) 3 个性能指标来评估网络的性能^[40], P、R

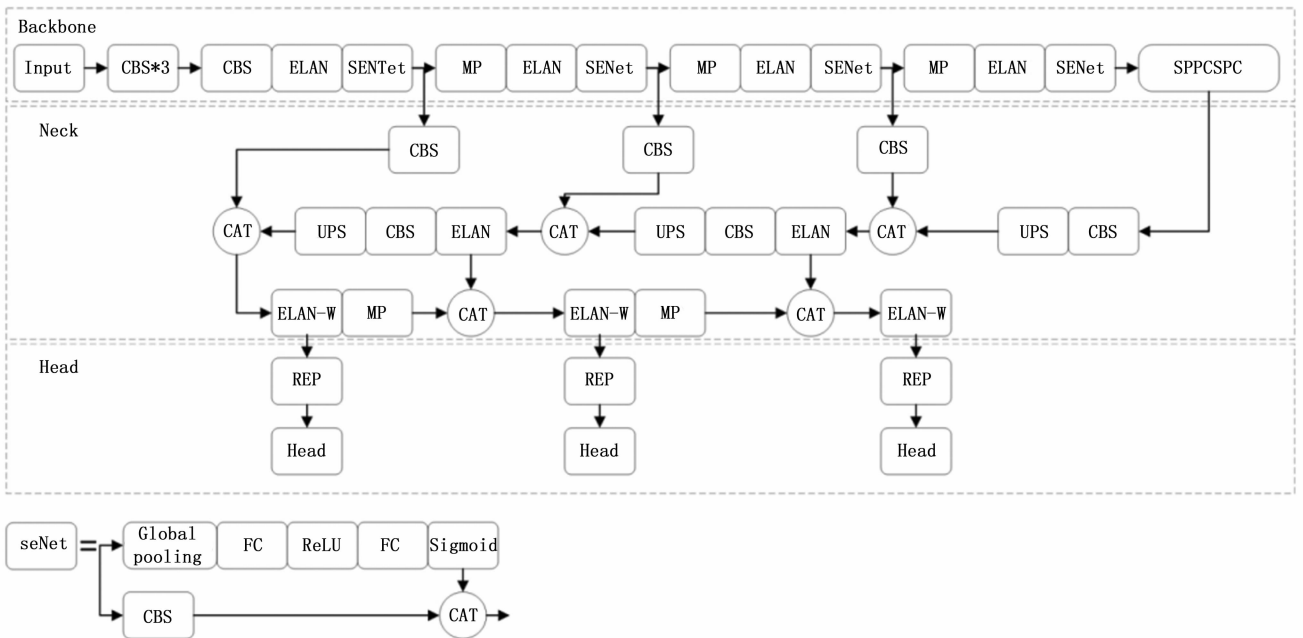


图 6 UWSC-YOLOv7 模块的架构

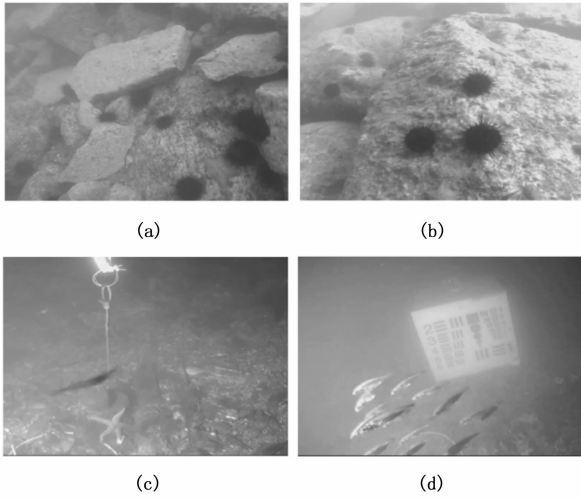


图 7 部分数据集示例

和 AP 的计算结果如表 1 所示，其中公式如下：

表 1 样本分类

	真正例	真负例
预测正例	TP	FP
预测负例	FN	TN

$$R = \frac{T_p}{T_p + F_n} \times 100\%$$

$$P = \frac{T_p}{T_p + F_p} \times 100\%$$

$$AP = \int_0^1 P(R) dR \quad (7)$$

Precision (P): Precision 是指在所有被分类为正例（即目标存在）的样本中，有多少比例是真正的正例。换句话说，Precision 衡量了模型在预测正例时的准确性，即模型的预测中有多少是正确的。Precision 侧重于评估模型预测的准确性，尤其是关注减少误报的情况，因此在一些对误报敏感的任务中很重要，比如医学诊断或者安全检测。

Recall (R): Recall 是指在所有实际正例样本中，有多少比例被正确地检测出来了。Recall 衡量了模型在所有正例中识别出的比例，即模型的预测中有多少是正确的。Recall 侧重于评估模型对于正例的覆盖率，尤其是关注避免漏报的情况，因此在一些对漏报敏感的任务中很重要。

Average Precision (AP): Average Precision 是 Precision-Recall 曲线下面积的平均值。Precision-Recall 曲线是在不同阈值下计算的 Precision 和 Recall 值的曲线。AP 衡量了模型在不同 Recall 水平下的平均 Precision，因此可以综合评价模型在不同阈值下的性能。AP 综合考虑了 Precision 和 Recall 两个方面的性能指标，更全面地评估了模型的性能。

3) 消融实验：

消融实验可以帮助研究人员深入理解模型的工作原理，识别模型中关键的组成部分，并验证这些组成部分的重

要性。

每个改进的实验数据如表 2 所示，所有实验的 epoch 设置为 200。从表中可以看出 YOLOv7 的 3 个指标都是低于 UWSC-YOLOv7 的 3 个指标。添加更多 FPN 层并切换为多尺寸检测头后，Neck 网络的 AP 提高了 1.43%。此外，通过将 SENet 纳入骨干网络内的几个位置，AP 增强了 0.32%。引入 EIou 损失函数替代 CIou 损失函数后 AP 提高了 0.78%。P 改善 4.33%、R 改进 7.21%、AP 改进 2.53%是所有改进的结果。

表 2 改进和推广效果

改进	P/%	R/%	mAP/%
YOLOv7	80.58	73.52	81.23
v7+newFPN	88.04	69.72	82.66
v7+newFPN+SENet	87.79	76.91	82.98
v7+newFPN+SENet+EIou	84.91	80.73	83.76

图 8 描绘了 YOLOv7 和 UWSC-YOLOv7 的 mAP 修改过程。对于相同的训练集，两个网络的 mAP 值在前 10 个 epoch 中增长得更快，并且更好的网络实现了更高的 60% AP 值。虽然增强网络已经达到融合 30 个 epoch 后，YOLOv7 需要 80 个 epoch 才能完成此操作。这意味着网络收敛得更快。

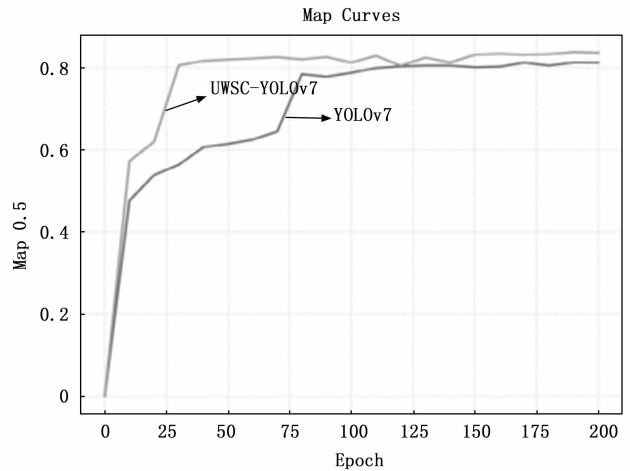


图 8 YOLOv7 和 UWSC-YOLOv7 的 mAP 变化

两种 Losses 的变化过程如图 9 所示。训练损失和验证损失在最初的 10 个 Epoch 中下降得更快，分别达到 0.11 和 0.07，并且训练损失在 140 个 Epoch 左右迅速下降，因为前 4 个目标具有更多数量的目标已达到最佳拟合。在 200 个 Epochs 时，YOLOv7 和 UWSC-YOLOv7 的训练损失达到最低点 0.05 和 0.04，UWSC-YOLOv7 的损失更低。

表 3 显示了 YOLOv7 和 UWSC-YOLOv7 之间所有类型目标的 P、R 和 AP 指标的比较结果。对于前 4 个目标比例较高的小生物目标，YOLOv7 的 AP 值较大，升级后的网络已达到峰值并出现过拟合问题。然而，增强网络在最少水厂目标的 R 和 AP 指标上实现了惊人的 100%，并且所有

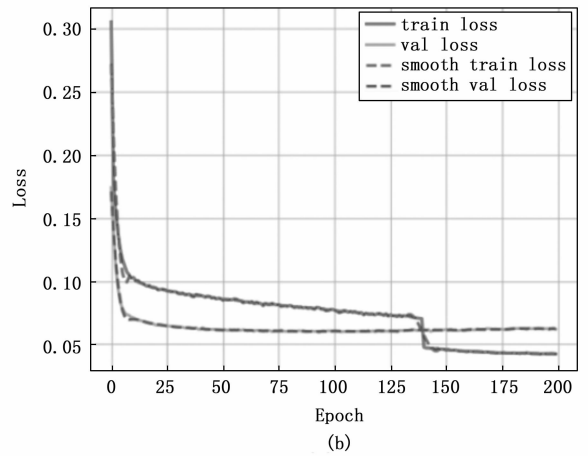
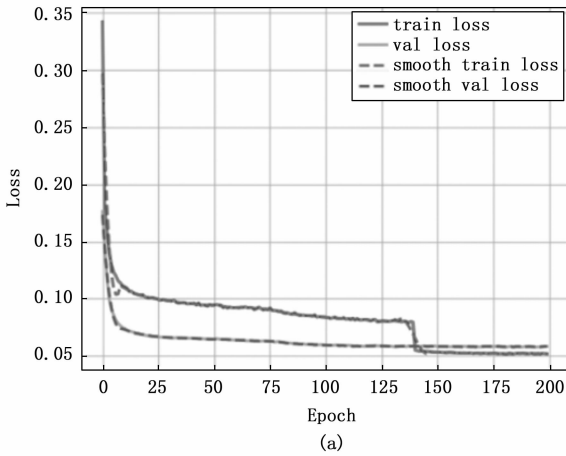


图 9 YOLOv7 (a)、UWSC-YOLOv7 (b) 的损耗变化

3 个指标都远远高于 YOLOv7。增强的网络对于样本较少的目标物种表现更好，并在数量非常有限的数据集上产生了出色的训练结果。

表 3 不同类别的比较

分类	YOLOv7			UWSC-YOLOv7		
	P/%	R/%	mAP/%	P/%	R/%	mAP/%
海胆	86.38	84.66	88.23	86.31	84.76	86.85
海星	84.84	79.30	86.59	85.34	76.91	82.93
扇贝	84.04	67.27	79.43	86.11	72.73	78.60
海参	80.98	69.69	73.36	81.09	69.25	70.42
水草	66.67	66.67	78.57	85.71	100	100
平均	80.58	73.52	81.23	84.91	80.73	83.76

4) 对比实验:

在实验中，UWSC-YOLOv7 与 YOLOv7、YOLOX^[41]、YOLOv5、YOLOv4^[42]、YOLOv3^[43]、SSD 和 Faster RCNN 模型进行了对比。结果如表 4 所示。获得了最好的结果，并且所有 3 个 UWSC-YOLOv7 的指标比大多数网络更高且更平均。

表 4 2020 年数据集不同版本的比较

模型	P/%	R/%	mAP/%
Faster RCNN	38.27	80.41	71.18
SSD	91.46	38.83	62.61
YOLOv3	65.91	46.94	59.61
YOLOv4	86.19	40.84	50.47
YOLOv5	68.28	53.88	61.93
YOLOX	82.67	70.95	80.50
YOLOv7	80.58	73.52	81.23
UWSC-YOLOv7	84.91	80.73	83.76

在本书的第二部分中，UWSC-YOLOv7 与模型、数据集和 mAP 值的一些算法进行了比较。产生较差结果的算法未包含在比较中，结果显示在表 5 中。UWSC-YOLOv7 的 mAP 高于大多数其他网络，显示出显著的进步。尽管 Deep

CNN 网络实现了最大 mAP 值 91.2%，但其数据集只有 3 个类别，相比之下本研究使用的 2020dataset 数据集有 5 个类别。与更具创新性的方法相比，UWSC-YOLOv7 也可以产生积极的结果。

表 5 其他比较

模型	数据集	mAP/%
SA-FPN	PASCAL VOC	76.27
Deep CNN	URPC(3 分类)	91.20
YOLOv4-uw	UPRC2020	75.34
APAN	Underwater Image	79.60
YOLOv5s-CA	URPC2019	80.90
TC-YOLO	RUIE2020	83.10
UWSC-YOLOv7	URPC2021	83.11

5) 效果图:

具体结果如图 10 所示，训练后可以成功预测水下小型生物目标。对于 (a)、(b) 图中较大的海胆、海星、海参，检测精度可达 0.85 左右。鉴于距离很远，部分覆盖的海胆和海星的准确度可以合理地达到 0.7 以内。(c) 和 (d) 中虾的准确度为 0.79，螃蟹的准确度为 0.88，海星的准确度

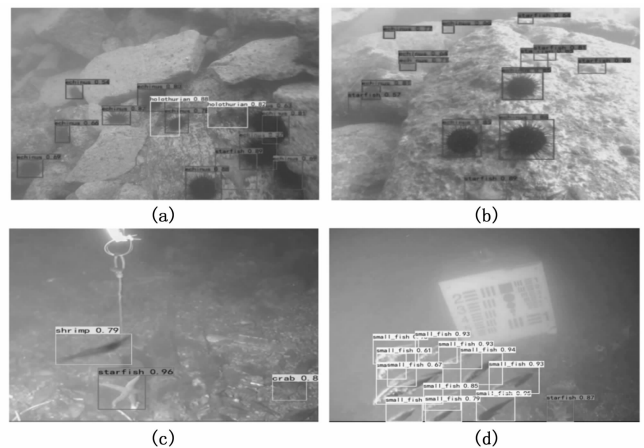


图 10 测试结果图

约为 0.9, 大多数小鱼目标都在 0.9 以上, 少数为 0.6 或 0.7。效果图展示了算法如何有效地发现和分类水下小目标生物。结果表明, UWSC-YOLOv7 算法在水下小目标生物的分类和检测方面表现良好, 显著降低了漏检和误检的问题, 同时提高了网络的检测精度。

本研究在 3 个不同的数据集上进行了对比测试, 以证实增强网络模型的通用性, 结果如表 6 所示。在 3 个数据集上, 修正后的模型分别显示出 2.53%、1.12% 和 3.19% 的改进。所有 3 个数据集的显著改进可能表明网络模型效果具有更好的普遍性。值得注意的是, 该模型可以在某些方面产生出色的结果差异较小的类别, 海星在 Brackish 数据集上的 *mAP* 值也超过 100%。

表 6 3 个数据集的比较

分类	YOLOv7 <i>mAP</i> /%	UWSC-YOLOv7 <i>mAP</i> /%
2020dataset	81.23	83.76
UPRC2021	82.19	83.31
Brackish	85.10	88.29

在这 3 个数据集中, 根据实验结果, UWSC-YOLOv7 的表现令人满意, 尽管升级后的方法也存在一定的缺陷。

1) 当两个目标彼此靠近时, 每个网格单元只有两个边界框, 并且它们是同一类的成员, 检测效果很差。当存在一群较小的目标时, 它也不是很好。

2) 对于具有新颖或非典型纵横比或组合的物体, 泛化能力较弱。

3) 需要更平衡的样本量, 因为当样本量高度不均匀时, 某些目标类型可能会过度拟合, 但其他目标类型也可能无法检测到, 这会阻止模型产生最佳结果。

4) 样本质量差, 如光线昏暗或样本图片严重遮挡等, 会降低检测效果, 增加漏检或误检的可能性。

未来针对该问题的研究可以通过平衡数据样本、改进损失函数以及在训练和检测前对图像进行预处理来增强样本质量, 进一步增强网络效果。

4 结束语

本文针对水下小目标检测漏检、检测精度低的问题, 提出了一种增强型 UWSC-YOLOv7 水下微小生物识别方法。通过在 Backbone 特征提取网络中加入 SENet, 修改增强型特征提取网络部分中的 FPN 结构, 并用 EIoU 损失函数代替原来的 CIoU 损失函数, 对模型进行了改进。根据消融实验和对比实验的结果, UWSC-YOLOv7 网络模型达到了 83.76% 的 *mAP*, 检测效果优于原始模型和其他常规经典目标检测网络模型。这有效提高了模型的检测精度, 同时实现了对水下小型生物的更准确检测。

本文列举了早期关于小目标的研究成果, 将小目标定义为属于特定数据集或具有某些特征的样本, 并阐述了研究的主要问题。使用 YOLOv7 算法探索水下目标检测, 并在同一数据集上与其他算法进行比较, 以证明其优越性。

通过增强 YOLOv7 算法对小物体的通道特征信息改进、多尺度检测和损失函数, 减少误检和漏检, 同时提高检测精度和网络效应性能。将改进网络应用于不同的数据集, 验证改进算法的可行性和普遍适用性, 进而达到目标检测效果提升的目的。

然而, 该实验仍然存在缺陷, 因为数据集的目标非常分散, 使得模型不可能强制所有检测目标同时达到峰值。该模型还可以通过未来数据集进一步改进。

水下小目标检测的未来的研究方向一是改进特征提取技术, 针对水下环境中目标的低对比度、复杂背景等问题设计更加有效的特征表示方法; 二是优化目标检测模型, 结合深度学习、强化学习等技术, 提高水下小目标检测的准确性和鲁棒性; 三是研究水下成像技术, 改进水下图像采集装置, 提高水下图像的质量和清晰度; 四是开发更多样化的数据集, 包括各种水下场景和目标类型, 促进水下小目标检测算法的训练和评估。综合考虑水下环境的特殊性和未来技术发展趋势, 这些方向有望推动水下小目标检测技术迈向更加准确、可靠和实用的阶段。

参考文献:

- [1] GUO Y, LI H, ZHUANG P. Underwater image enhancement using a multiscale dense generative adversarial network [J]. IEEE Journal of Oceanic Engineering, 2019, 45 (3): 862-870.
- [2] LEONARD J J, BAHAR A. Autonomous underwater vehicle navigation [J]. Springer Handbook of Ocean Engineering, 2016: 341-358.
- [3] KINSEY J C, EUSTICE R M, WHITCOMB L L. A survey of underwater vehicle navigation: Recent advances and new challenges [C] //IFAC Conference of Manoeuvring and Control of Marine Craft, 2006, 88: 1-12.
- [4] 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, 2023: 7464-7475.
- [5] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2016, 29.
- [6] SHOBANA M, MADHAVAN M, MEHA V S, et al. Ai-underwater drone in protection of waterways by relating design thinking framework [C] //2023 International Conference on Computer Communication and Informatics (ICCCI), IEEE, 2023: 1-4.
- [7] ER M J, CHEN J, ZHANG Y, et al. Research challenges, recent advances, and popular datasets in deep learning-based underwater marine object detection: A review [J]. Sensors, 2023, 23 (4): 1990.
- [8] ZOU Z, CHEN K, SHI Z, et al. Object detection in 20 years: A survey [J]. Proceedings of the IEEE, 2023, 111 (3): 257-276.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich fea-

- ture hierarchies for accurate object detection and semantic segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014; 580–587.
- [10] HE K, ZHANG X, REN S, 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.
- [11] GIRSHICK R. “Fast R-CNN,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, 1440–1448.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once; Unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016; 779–788.
- [13] LIU W, ANGUILOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C] //Computer Vision-ECCV 2016; 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016; 21–37.
- [14] CHEN L, LIU Z, TONG L, et al. Underwater object detection using Invert Multi-Class Adaboost with deep learning [C] //2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020; 1–8.
- [15] LIN W H, ZHONG J X, LIU S, et al. Roimix: proposal-fusion among multiple images for underwater object detection [C] //ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020; 2588–2592.
- [16] LIU H, SONG P, DING R. WQT and DG-YOLO: Towards domain generalization in underwater object detection [J]. ArXiv Preprint ArXiv; 2004.06333, 2020.
- [17] XU F, WANG H, PENG J, et al. Scale-aware feature pyramid architecture for marine object detection [J]. Neural Computing and Applications, 2021, 33; 3637–3653.
- [18] JIANG Z, WANG R. Underwater object detection based on improved single shot multibox detector [C] //Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, 2020; 1–7.
- [19] HAN F, YAO J, ZHU H, et al. Marine organism detection and classification from underwater vision based on the deep CNN method [J]. Mathematical Problems in Engineering, 2020, 2020; 1–11.
- [20] WANG H, SUN S, WU X, et al. A YOLOv5 baseline for underwater object detection [C] //OCEANS 2021; San Diego-Porto, IEEE, 2021; 1–4.
- [21] CHEN L, ZHENG M, DUAN S, et al. Underwater target recognition based on improved YOLOv4 neural network [J]. Electronics, 2021, 10 (14): 1634.
- [22] ZENG L, SUN B, ZHU D. Underwater target detection based on Faster R-CNN and adversarial occlusion network [J]. Engineering Applications of Artificial Intelligence, 2021, 100; 104190.
- [23] LI J, LIU C, LU X, et al. CME-YOLOv5: An efficient object detection network for densely spaced fish and small targets [J]. Water, 2022, 14 (15): 2412.
- [24] YU H, LI X, FENG Y, et al. Multiple attentional path aggregation network for marine object detection [J]. Applied Intelligence, 2023, 53 (2): 2434–2451.
- [25] WEN G, LI S, LIU F, et al. YOLOv5s-CA: A modified YOLOv5s network with coordinate attention for underwater target detection [J]. Sensors, 2023, 23 (7): 3367.
- [26] LIU K, PENG L, TANG S. Underwater object detection using TC-YOLO with attention mechanisms [J]. Sensors, 2023, 23 (5): 2567.
- [27] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C] //Computer Vision-ECCV 2014; 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13, Springer International Publishing, 2014; 740–755.
- [28] LIM J S, ASTRID M, YOON H J, et al. Small object detection using context and attention [C] //2021 International Conference on Artificial Intelligence in Information and Communication (ICAIC), IEEE, 2021; 181–186.
- [29] CHENG G, YUAN X, YAO X, et al. Towards large-scale small object detection: Survey and benchmarks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.
- [30] LI C, GUO C, REN W, et al. An underwater image enhancement benchmark dataset and beyond [J]. IEEE Transactions on Image Processing, 2019, 29; 4376–4389.
- [31] LIU W, ANGUILOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C] //Computer Vision-ECCV 2016; 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016; 21–37.
- [32] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing systems, 2017, 30.
- [33] 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): 331–368.
- [34] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018; 7132–7141.
- [35] ZHAO Z, YANG X, ZHOU Y, et al. Real-time detection of particleboard surface defects based on improved YOLOV5 target detection [J]. Scientific Reports, 2021, 11 (1): 21777.
- [36] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017; 2117–2125.

[37] BENJUMEA A, TEETI I, CUZZOLIN F, et al. YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles [J]. ArXiv Preprint ArXiv: 2112.11798, 2021.

[38] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (7): 12993 - 13000.

[39] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146 - 157.

[40] DAVIS J, GOADRICH M. The relationship between Precision-Recall and ROC curves [C] // Proceedings of the 23rd International Conference on Machine Learning, 2006: 233 - 240.

[41] GE Z, LIU S, WANG F, et al. Yolox: Exceeding yolo series in 2021 [J]. ArXiv Preprint ArXiv: 2107.08430, 2021.

[42] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint ArXiv: 2004.10934, 2020.

[43] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. ArXiv Preprint ArXiv: 1804.02767, 2018.

(上接第 26 页)

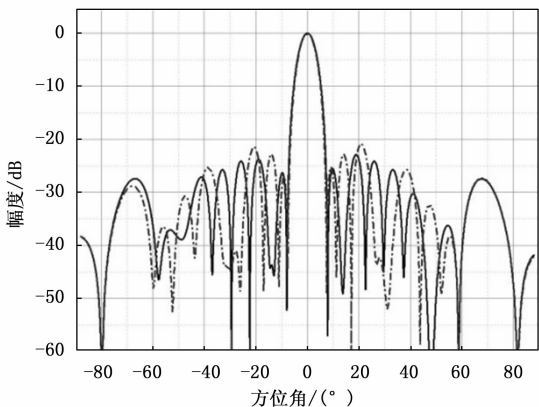


图 11 波瓣图数据比对

表 1 测试数据比对

	增益/dB	波束宽度/(°)	副瓣/dB
机械臂天线测试系统	16.25	14.5	11.77
传统平面天线测试系统	16.37	14.3	11.61

从测试数据可以得出，基于机械臂毫米波天线测试系统与传统平面天线测试系统，测试结果增益偏差 0.73%，波束宽度偏差 1.4%，副瓣偏差 1.37%，两种系统测试结果一致性较好，故基于机械臂毫米波天线测试系统能够满足毫米波天线测试要求。

4 结束语

本文提出了基于机械臂的毫米波天线测试系统，与传统的平面天线测试系统相比，该系统具有多方面优势。基于机械臂天线测试系统部署灵活，能够满足复杂多变的测试环境；成本低，无需占用较大的测试场地，减少了暗室建造成本；测试能力强，不仅能够执行平面近场测试，还能够进行柱面近场、球面近场以及远场测试任务。实际应用表明，基于机械臂毫米波天线测试系统能够精确测量毫米波天线波瓣图，在毫米波天线测试领域具有广阔应用前景。

参考文献:

[1] 洪伟, 余超, 陈继新, 等. 5G 及其演进中的毫米波技术 [J]. 微波学报, 2020, 36 (1): 12 - 16.

[2] 邵 晖. 汽车雷达生产测试 [J]. 国外电子测量技术, 2017, 36 (5): 4 - 6.

[3] 据泽东, 黄晓红. 基于毫米波雷达的动态目标生命体征检测研究 [J]. 电子测量技术, 2022, 45 (16): 27 - 33.

[4] 吕思涵, 何宇奇, 赵鲁豫, 等. 高性能的 5G 毫米波天线 (阵列) 实现与性能评估 [J]. 移动通信, 2022, 46 (3): 13 - 20.

[5] John D. Kraus. 天线 (第三版) [M]. 章文勋译. 北京: 电子工业出版社, 2012.

[6] 张士选, 李 勇, 张福顺, 等. 超低副瓣天线近场测试关键技术探讨 [J]. 西安电子科技大学学报, 2000, 27 (3): 368 - 373.

[7] 侯 飞. 相控阵雷达天线近场测试系统发展与研制 [D]. 南京: 南京理工大学, 2009.

[8] 侯 飞, 柏 利, 乔淑君. 数字化相控阵天线测试方法及测试系统设计 [J]. 计算机测量与控制, 2017, 25 (1): 47 - 49.

[9] 胡鸿飞, 高 雪, 傅德明. 基于 FFT 的近场一口径场变换方法 [J]. 微波学报, 2001, 17 (2): 1 - 6.

[10] 李建新, 徐 慧, 胡明春, 邵江达. 基于 FFT 的阵列方向图快速计算 [J]. 现代雷达, 2007, 23 (1): 10 - 14.

[11] 陈晓昕, 石 磊, 肖 鸿. 一种基于平面近场测量的线阵天线快速检测方法研究 [J]. 雷达与对抗, 2015, 35 (4): 50 - 53.

[12] 宋 诚. 平面近场测量的近-远场及近一口径场变换研究 [D]. 西安: 西安电子科技大学, 2014.

[13] 乔淑君, 李晓峰. 有源相控阵天线测试通用平台设计与实现 [J]. 计算机测量与控制, 2022, 30 (7): 20 - 25.

[14] 吴天一. 毫米波信道测量与特征参数提取 [D]. 南京: 东南大学, 2017.

[15] 张福顺, 张进民. 天线测量 [M]. 西安: 西安电子科技大学出版社, 1995.

[16] FRANCIS M H, WITTMANN R C. Modern antenna handbook [M]. Hoboken: Wiley, 2008.

[17] 机械臂系统设计及其自适应神经网络控制研究 [D]. 成都: 电子科技大学, 2018.

[18] 余 乐. 六轴机械臂对人手臂动作模仿的研究 [D]. 厦门: 华侨大学, 2018.

[19] 王雅君. 基于 Qt 开发框架的多波束测深声呐显控技术研究 [J]. 电声技术, 2022, 46 (3): 101 - 104.

[20] 李晓峰, 侯 飞, 白 雪. 天线近场测试软件通用化和多任务设计 [J]. 计算机测量与控制, 2016, 24 (11): 48 - 51.