智能系统学报
CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

面向失效增强和改进YOLOv8的目标检测

储文娟, 李震, 黄炜嘉, 王宇轩

引用本文:
储文娟, 李震, 黄炜嘉, 等. 面向失效增强和改进YOLOv8的目标检测[J]. 智能系统学报, 2026, 21(2): 353-364.
CHU Wenjuan, LI Zhen, HUANG Weijia, et al. A failure enhancement and improvement of YOLOv8 for target detection[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(2): 353-364.

在线阅读 View online: https://dx.doi.org/10.11992/tis.202503010

DOI: 10.11992/tis.202503010
网络出版地址: https://link.cnki.net/urlid/23.1538.tp.20260209.1447.002

面向失效增强和改进 YOLOv8 的目标检测

储文娟, 李震, 黄炜嘉, 王宇轩
(江苏科技大学 海洋学院, 江苏 镇江 212003)

摘　要: 针对当前在光照、天气、遮挡等复杂背景条件下进行目标检测技术的检测性能较低、泛化能力弱等问题, 文章提出一种基于失效增强和改进 YOLOv8 的目标检测算法(asymptotic structure of YOLO, AS_YOLO)。1)基于复杂场景构建了多种目标单元数据集, 并设计面向应用环境的图像失效增强技术; 2)引入通道–空间并行注意力机制同时关注复杂环境下目标的特征信息与位置信息; 3)采用 AFPN 结构强化非相邻层级的特征融合效果; 4)采用了 Inner_IoU(inner intersection over union) 损失函数改善现有 IoU(intersection over union) 损失函数, 在不同检测任务中的泛化能力不足的问题, 并在 WSODD 多目标数据集下进行迁移实验。实验结果表明, 改进后的算法与基线模型 YOLOv8n 相比, $mAP_{0.5}$ 达到了 94.0%, 提升 12.5 百分点, $mAP_{0.95}$ 达到了 72.5%, 提升 15.7 百分点, 具有更好的检测性能。

关键词: 计算机视觉; 复杂环境; 目标检测; YOLO; 图像增强; 注意力机制; 特征融合; 损失函数
中图分类号: TP391.41　　**文献标志码**: A　　**文章编号**: 1673−4785(2026)02−0353−12

中文引用格式: 储文娟, 李震, 黄炜嘉, 等. 面向失效增强和改进 YOLOv8 的目标检测 [J]. 智能系统学报, 2026, 21(2): 353–364.
英文引用格式: CHU Wenjuan, LI Zhen, HUANG Weijia, et al. A failure enhancement and improvement of YOLOv8 for target detection[J]. CAAI transactions on intelligent systems, 2026, 21(2): 353–364.

A failure enhancement and improvement of YOLOv8 for target detection

CHU Wenjuan, LI Zhen, HUANG Weijia, WANG Yuxuan
(Ocean College, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: To address the issues of low detection performance and weak generalization ability in target detection under complex background conditions such as illumination, weather, and occlusion, this paper proposes an improved object detection algorithm based on failure augmentation and enhanced YOLOv8 (AS_YOLO). First, a variety of target unit datasets were constructed based on complex military scenarios, and an image failure augmentation technique tailored to the application environment was developed. Second, a channel-spatial parallel attention mechanism was introduced to simultaneously focus on feature and position information of targets in complex environments. Then, the AFPN structure was used to enhance feature fusion of non-adjacent hierarchical layers. Finally, the Inner_IoU loss function was adopted to address the generalization limitations of existing IoU loss functions in different detection tasks. Transfer experiments were conducted on the WSODD multi-target dataset. The experimental results show that the improved algorithm achieves an $mAP_{0.5}$ of 94.0%, a 12.5 percentage point improvement over the baseline YOLOv8n model, and an $mAP_{0.95}$ of 72.5%, a 15.7 percentage point improvement, indicating superior detection performance.

Keywords: computer vision; complex environment; object detection; YOLO; image enhancement; attention mechanism; feature fusion; loss function

近年来, 目标检测技术成为计算机视觉领域的核心技术之一, 对整个社会产生了深远影响。目标检测算法以结合卷积神经网络 (convolutional neural networks, CNN)[1] 和多尺度特征提取技术为基础, 以提高检测精度和速度。经典方法包括基于候选区域的两阶段检测算法如 Faster R-CNN (faster region-based convolutional neural networks)[2]

收稿日期: 2025−03−06.　**网络出版日期**: 2026−02−10.
基金项目: 国家自然科学基金项目(62276285); 教育部学位与研究生教育发展中心主题案例库项目(ZT-231028914); 江苏省研究生科研与实践创新计划项目(KYCX24-4178); 中国科学院软件研究所合作项目(220507-2325).
通信作者: 李震. E-mail: justlz@just.edu.cn.

和直接预测目标位置与类别的单阶段算法如 YOLO(you only look once)[3] 系列。单阶段检测器在一个阶段中同时进行目标分类和边界框回归，模型直接从输入图像生成预测。双阶段检测的第一阶段是对输入图像抽取兴趣区域 (region of interest, ROI)，第二阶段对每个 ROI 进行对象分类和边界框回归。

复杂环境下的目标成像面临诸多挑战，例如多变的天气、光照条件和烟雾遮挡等，这些因素导致图像质量下降，目标尺寸变小，成像模糊等问题，严重影响了检测器提取有效特征的能力。现有的基于深度学习的复杂环境下的目标检测方法大多应用于退化去除的清晰图像，但这一策略在复杂环境下的目标检测中存在诸多局限性，例如单一的退化还原方法难以全面提升复杂环境下的目标检测性能。同时研究人员对模型架构设计关注较多[4]，很少考虑到真实场景下光学图像通常受到雾霾、光照、雨雪、遮挡的影响。此外，由于成像距离远，某些图像的分辨率低，缺乏细节和纹理信息，且受到图像噪声的影响，这些因素共同增加了模型过拟合的风险，从而影响其泛化能力。因此，利用图像增强技术和先进的网络模型，为目标检测任务提供丰富的样本和更加优异的算法，对于提升目标检测任务在复杂环境下的准确性、可靠性和泛化能力，具有重要的研究意义。

Wang 等[5] 提出一种真实场景遥感变化检测的数据增强方法，深度学习模型通过从雾霾或薄云覆盖区域或光谱差异较大的缝线区域中学习到足够的语义信息来提高深度学习模型在雾霾/薄云或大片地区的检测精度；Wu 等[6] 提出了一种免训练的图像增强方法，利用预训练分段任意模型 SAM(segment anything model) 模型作为数据增强工具 PTSAM-DA(pre-trained segment anything model for data augmentation) 来生成图像的增强注释；肖晶晶等[7] 提出一种雾环境下船舶检测方法，通过构建数据集、引入 GSConv(ghost shuffle convolution)、Slim-Nec(slim neck network) 及优化损失函数改进 YOLOv5 结构以提升检测精度与模型轻量化；马淦等[8] 针对海面成像过程易受天气、光照、水雾等影响问题，设计出一种动态“复制–粘贴”的数据增强提高海面舰船检测性能；Fan 等[9] 针对多变的光照条件和遮阴等挑战使智能采摘机器人的任务复杂化的问题，提出 Center Net 神经网络为检测框架，引入以灰度为中心的 RGB(red green blue) 颜色空间垂直分解图，利用分组卷积和深度可分离卷积设计了一个包含 8 个瓶颈结构的轻量级特征提取网络 Light-Weight Net 来提高检测精度；刑汇源等[10] 采用 RepVGG(re-parameterized visual geometry group network) 改进特征提取模块，融合高分辨率特征图改进融合网络来提高模型在复杂海面目标检测方法；张国印等[11] 针对复杂场景下行人相互遮挡导致目标跟踪精度低的问题，应用自注意力结构提取行人的外观特征；Lyu 等[12] 构建一个混合骨干网络 CMNet 为模型提供了一个全局视场，增强模型对小目标和模糊目标的探测能力；吴攀超等[13] 针对雾霾环境下交通标志检测的漏检、错检及参数较大等问题改进 YOLOv5 检测模型，通过设计基于颜色衰减先验的自适应伽马变换算法，显著增强了图像预处理能力；赵文清等[14] 通过引入轻量级通道注意力和坐标注意力机制提高背景信息复杂环境下的目标检测精确率；许迪等[15] 为解决小目标检测和复杂天气中目标检测精度低、环境干扰因素大、难以在性能一般的移动设备部署等问题,设计实现一种改进 YOLOv8 安全装备检测 (YOLOv8-deep semantic integration, YOLOv8-DSI)。

虽然上述方法可以在复杂背景环境中达到一些检测效果，但是在复杂环境下进行目标检测不仅要满足高精度的要求，还应该关注在不同数据目标和真实检测任务中模型存在泛化能力不足的问题。特别是在面临真实复杂环境的多变性时，如恶劣天气变化、光照变化和烟雾遮挡等，这些因素会显著降低图像质量，使目标变得模糊、尺寸变小，且背景复杂，严重影响了目标检测模型从图像中提取有效特征的能力。同时在低质量图像和复杂背景下，现有算法往往存在过拟合的风险，导致模型的泛化能力较弱，难以应对多变的环境条件。针对上述模型在复杂环境下进行目标检测出现的问题，文章以多种场景和真实环境中常见目标为基础，基于传统图像增强提出一种新的面向应用环境的图像增强方法，构建目标数据集，为后续的目标检测任务提供了更加真实和准确的图像输入。同时为了解决复杂背景下目标检测过程中常见的漏检、误检率高，精度低，速度慢等问题，文章基于当前主流的回归型目标检测算法 YOLOv8 进行改进提出了一种新的目标检测算法。首先，为了提升网络在复杂环境下对于目标特征的提取能力，在基础模型的压缩–卷积–扩展 (BottleNeck) 处引入通道–空间并行注意力机制模块[16]，增强网络在复杂环境下对目标特征信息与位置信息关注度，使网络学习到的信息更加丰富；其次，为了提升模型对于特征的融合能力与

速度, 引入渐进特征金字塔网络 (adaptive feature pyramid network, AFPN) 进行特征融合来加强非相邻层级的融合效果[17], 加速模型特征融合; 最后为了提高模型的预测能力, 采用 Inner_IoU 损失函数[18] 加速模型收敛, 弥补现有 IoU 损失函数在不同的检测任务中的泛化能力较弱且收敛速度较慢的不足。文章在面向应用环境目标数据集的基础上, 设计失效增强来丰富样本数据, 模拟真实复杂环境, 同时改进模型。在原始数据集和失效数据集上分别进行了训练与测试, 并与现有主流目标检测算法进行了对比分析, 以验证失效增强算法和改进目标检测模型在复杂环境下检测的性能。为了进一步证明改进算法的有效性, 文章还进行了模型拆分的消融实验, 对比不同模型组件的贡献和影响。实验结果表明, 所提出的目标检测算法在自建数据集和迁移数据集上均展现出卓越的检测性能, 尤其是在处理复杂环境中的目标检测任务时表现尤为突出。该算法不仅提高了对军事目标的识别精度, 还在应对复杂背景和多样化目标时展现出强大的适应能力。

1　面向复杂应用环境的失效增强

1.1　复杂环境数据集特点与收集

构建涵盖多种复杂场景条件的数据集, 是实现高效检测算法的关键基础。文章分析检测模型在不同环境下的目标感知特点, 根据实际场景中的典型行动单元构建目标数据集。

在实际应用场景中, 特别是海上环境条件下, 浓雾会导致图像信息严重缺失, 造成目标细节的消失, 使得检测系统出现漏检或错检, 甚至完全无法识别目标的类别和形状。

雨雪天是实际环境中常出现的天气类型, 雨雪天气会降低图像整体能见度外, 雨线和雪花在图像中直接形成干扰条纹和模糊区域, 影响特征提取的连续性与稳定性, 遮挡图像局部信息, 增加系统的误判风险。

阴天由于缺乏强烈的自然光照, 整体图像对比度下降, 目标轮廓不清, 颜色灰暗, 进一步加大检测系统的识别难度。

晴天虽然具备较强的光照条件, 但若目标表面反光严重, 则强光可能导致图像局部区域过曝, 掩盖目标纹理或产生伪影, 从而误导检测系统。

光照条件的变化同样对感知性能造成不可忽视的影响。黎明与黄昏作为昼夜交替的特殊时间段, 光照方向与强度快速变化, 图像中常出现低照度区域与高光反差区域并存的现象。由于光照逐渐减弱, 目标边缘和颜色信息变得模糊, 尤其对于某些颜色与背景相近的目标, 检测难度显著提升。

白天光照通常较为稳定, 但受天气、地形等因素影响依然可能存在局部阴影或强反光区域。

夜间环境中, 整体亮度极低, 可见光成像设备受限严重, 图像信噪比下降, 目标与背景对比度弱, 加之实际黑夜行动中常缺少高亮度照明设备的使用, 导致目标识别效果显著衰减。此外, 夜间背景光源的杂散干扰, 也会进一步加剧识别系统的误判概率。

在噪声干扰方面, 实际应用环境下的图像采集平台通常处于运动状态, 机械振动以及地形颠簸易导致图像抖动、偏移、模糊。同时, 实际应用场景中的爆炸、炮击等会产生强烈的冲击波与电磁干扰, 直接影响传感器工作稳定性, 导致图像中出现大量随机噪声与局部图像块失真。这类高强度噪声不仅削弱了图像的特征表达能力, 也极易引起目标错检、漏检, 显著增加系统的虚警率。

遮挡现象在目标检测任务中尤为普遍, 在实际任务中, 目标常处于动态或半遮挡状态, 尤其是在人员、装备密集, 环境地形复杂的场景中, 遮挡物种类繁多, 如植被、建筑、武器装备等, 均可能对目标造成局部或完全遮挡。这种遮挡不仅削弱了可见区域的特征信息, 还可能导致识别系统错将非目标区域作为检测对象, 影响整体判断。动态遮挡如爆炸烟雾、快速通过的载具等, 还会干扰目标连续跟踪, 造成目标轨迹中断、目标切换频繁等问题, 极大削弱了感知系统在长时间任务中的可靠性。

因此, 在算法设计与系统部署中, 应充分考虑复杂天气、光照、噪声与遮挡条件的影响, 通过引入环境适应性增强机制和复杂场景仿真训练等方式, 有效提升感知系统在实际应用环境任务中的应用性能。文章以 3 种场景(海洋、陆地、天空)和真实环境中常见的 8 类目标为基础, 构建了自建目标数据集, 并根据复杂环境稀缺问题提出失效增强算法。

1.2　失效增强算法

基本模型 YOLOv8 采用了翻转、旋转、剪切、Mosaic[19] 等数据增强策略来提升检测性能。然而真实环境是复杂多变的, 光照、天气、遮挡等众多干扰因素都会影响目标成像质量, 进而影响复杂环境下的目标检测性能, 最终导致视觉软件失效。因此上述数据增强方法对于复杂环境下的目标检测性能提升有限, 为进一步模拟真实复杂环境下的图片成像效果, 提升检测精度, 文章从训

练样本的角度提出了一种简单而有效的数据增强方法，面向应用环境的图像失效增强。

面向应用环境的图像失效增强从天气、光照、场景、干扰和成像等方面进行失效模式分析。首先从复杂环境特点出发，设计了面向应用环境的基础失效增强方法，并根据基础失效组合概率，设计基础失效组合增强。其次，在基础环境失效增强仿真得到的新样本上，进行多失效模式组合增强仿真。最后，输入深度学习模型进行模型训练。算法流程如算法 1 所示。

算法 1　面向失效的图像增强仿真

输入　随机样本

输出　增强样本

1) for　样本 1　in　数据集:

2) if　增强方式==1:

3) 传统图像增强（几何变换）

4) elif　增强方式 == 2:

5) 基础失效增强（随机选择天气+光照状态）

6) elif　增强方式 == 3:

7) 组合失效增强（根据天气组合矩阵模拟状态序列）

8) elif　增强方式 == 4:

9) 多失效模式图像增强（叠加复杂场景下干扰与遮挡）

算法 1 中基础失效增强包括天气、光照、场景因素。根据面向应用环境的失效模式分析得出多种复杂环境下的失效场景。

5 种天气状态：雾天、晴天、阴天、雨天、雪天。

4 种光照状态：白天、黄昏、黑夜、黎明。

3 种场景状态：海洋、陆地、天空。

首先，对于任意一个训练样本建立一个离散的组合模型。由于样本本身包含场景信息，基础失效增强随机在天气集合和光照集合中各选择一次失效增强来保证三者之间的独立性。其具体方法为

$$B = \mathrm{random}(W, L, S) \tag{1}$$

式中：B 表示基础增强结果，W 表示天气，L 表示光照，S 表示场景。

例如，“晴天–白天–陆地”可以为一组基础失效增强，即共有 60 种可能状态。

天气集合中的雾天、晴天、阴天、雨天、雪天并不是相互独立的，例如，如果基础失效增强为晴天，那么其组合状态还有可能是雾天。因此建立一个关于天气的离散状态组合模型。首先需要根据实际场景确定天气状态之间的组合概率，得到一个状态组合矩阵：

$$\boldsymbol{M} = \begin{bmatrix} P_{0,0} & \cdots & P_{0,3} \\ \vdots & \ddots & \vdots \\ P_{3,0} & \cdots & P_{3,3} \end{bmatrix} \tag{2}$$

式中：P_{ij}代表i天气状态下，j天气状态的组合概率。

本文在统计过程中，以某一日的主天气，如“晴”为状态i，统计其当天是否同时出现次级天气状态j，如“雾”，然后对所有观测日进行遍历，得到各状态组合ij的共现频数，最终归一化得到组合概率P_{ij}。例如，统计结果显示在“晴天”日中，有 32% 的概率同时伴有“雾”，因此$P_{晴雾}$=0.32。

基于状态组合矩阵，来模拟天气因素下不同状态组合情况，了解在不同的初始条件下，可能经历的状态序列，从而根据序列完善基础失效组合增强。图 1 为一组基础失效组合增强序列。

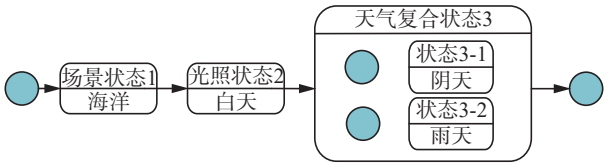


图 1　基础失效组合增强序列样例

Fig. 1　Example of a basic failure combination enhancement sequence

在复杂海面环境下进行目标检测不仅应用环境的复杂多变会导致输入数据的不确定性，而且特殊复杂场景中的噪声和遮挡是远远大于普通场景的。多失效模式组合增强在基础环境仿真得到的样本上结合干扰失效因素和成像失效因素，进行多失效模式组合增强仿真。

文章设置了 4 种图像失效增强启动方式动态调整增强模式，实现传统图像增强的随机仿真、基础失效的随机仿真，基础组合失效的随机仿真、多失效模式组合增强随机仿真的任意选择。同时根据式（1）、（2）实现样本的随机增强，并根据样本数量和复杂场景检测需求动态调整增强样本的数量。

2　改进的 YOLOv8 目标检测算法

为了提升 YOLOv8 模型在复杂环境下进行目标检测的效果，提出了一种灵活渐进式的目标检测方法 AS_YOLO，结构如图 2 所示。针对复杂背景带来的挑战，文章在 Backbone 网络中融合卷积注意力模块 (convolutional block attention module, CBAM) 模块，以获取空间和通道两个维度的信息，从而更好地将背景等因素与目标物体联系起来，提升模型检测精度；此外，AS_YOLO 替换原有的路径聚合网络 (path aggregation network, PANet)[20] 网络结构，采用渐进特征金字塔结构 AFPN

渐进地将高层特征纳入融合过程，来减少语义信息在不同层级间传递过程中丢失或退化；由于目标远近交替，环境复杂多变，AS_YOLO 还引入了 Inner_IoU 机制，Inner_IoU 能够根据不同的检测器和检测任务进行自我调整，展现出极强的泛化能力，使得模型在各种复杂环境中均能保持较高的检测性能。

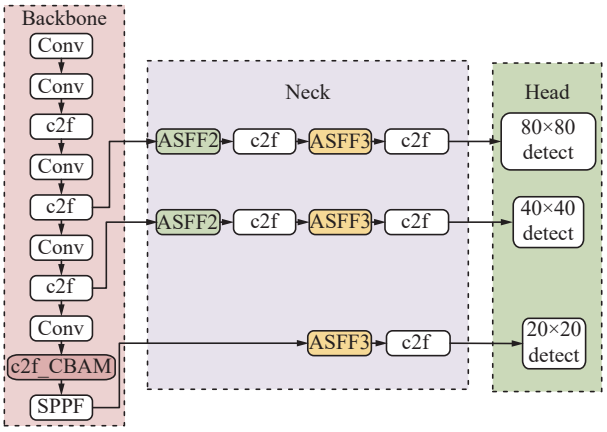


图 2　AS_YOLO 模型结构

Fig. 2　AS_YOLO model structure

2.1　提升复杂背景下的目标感知能力模块

在复杂环境下目标可能与背景有相似的颜色、纹理或形状，同时由于遮挡物和烟雾目标被部分遮挡，导致检测变得困难。因此在复杂背景环境下，目标检测算法需要处理更多的特征，导致模型计算复杂度增加。CBAM（convolutional block attention module）结合了通道注意力机制[21]和空间注意力机制[22]，可以提取更具信息量的特征。CBAM 可以根据每个通道的重要性调整特征图，使网络在复杂背景下更准确地聚焦于目标区域，并抑制背景区域的特征，从而减少背景干扰，提高目标检测的精度。CBAM 可以自适应地学习不同通道和空间位置之间的关系，使网络对不同场景和复杂背景具有更好的适应性，这也增强了网络的鲁棒性。CBAM 结构如图 3 所示。

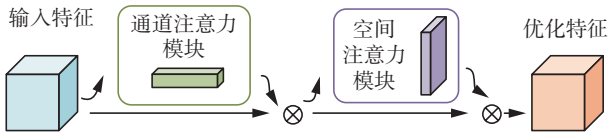


图 3　CBAM 结构

Fig. 3　CBAM structure

CBAM 的大致计算流程为

$$M_c(\boldsymbol{F})=\sigma(\mathrm{MLP}(\mathrm{AvgPool}(\boldsymbol{F}))+\mathrm{MLP}(\mathrm{MaxPool}(\boldsymbol{F})))=\sigma\left(\boldsymbol{W}_1\left(\boldsymbol{W}_0\left(\boldsymbol{F}^{\mathrm{c}}_{\mathrm{avg}}\right)\right)+\boldsymbol{W}_1\left(\boldsymbol{W}_0\left(\boldsymbol{F}^{\mathrm{c}}_{\mathrm{max}}\right)\right)\right)$$

式中：$\boldsymbol{W}_0\in\mathbf{R}^{C/r\times C}$，$\boldsymbol{W}_1\in\mathbf{R}^{C\times C/r}$，Avgpool(·)和 Maxpool(·)分别表示全局平均池化和全局最大池化，MLP(·)为全连接处理，$\sigma(\cdot)$为 sigmoid 函数。

首先，输入一张特征图 $\boldsymbol{F}\in\mathbf{R}^{C\times H\times W}$，经过通道注意力机制 M_c，得到通道注意力图 $M_c(\boldsymbol{F})$。

其次，将得到的 $M_c(\boldsymbol{F})$ 与输入图像进行像素级别的相乘得到 $\boldsymbol{F}'$。$\boldsymbol{F}'$ 继续进入到空间注意力机制 M_s，由公式得到空间注意力图 $M_s(\boldsymbol{F}')$：

$$M_s(\boldsymbol{F}')=\sigma(f^{7\times7}([\mathrm{AvgPool}(\boldsymbol{F}');\mathrm{MaxPool}(\boldsymbol{F}')]))=\sigma\left(f^{7\times7}\left(\left[\boldsymbol{F}^{\mathrm{s}}_{\mathrm{avg}};\boldsymbol{F}^{\mathrm{s}}_{\mathrm{max}}\right]\right)\right)$$

式中：f 表示卷积运算，得到的 $M_s(\boldsymbol{F}')$ 与输入图像 $\boldsymbol{F}'$ 进行像素级别的相乘操作，得到 $\boldsymbol{F}''$。

为了降低模型计算量，选择在主干网络的深层融合 CBAM 模块。相比于深层特征，浅层特征图具有较大的分辨率，因此如果在浅层网络中融合 CBAM 模块，会显著增加模型的计算量。通过将 CBAM 模块集成在主干网络的深层，不仅能够有效减轻计算负担，还能充分利用深层特征的表达能力。融合 CBAM 模块前后的 c2f(cross stage to feature) 结构对比如图 4 所示。

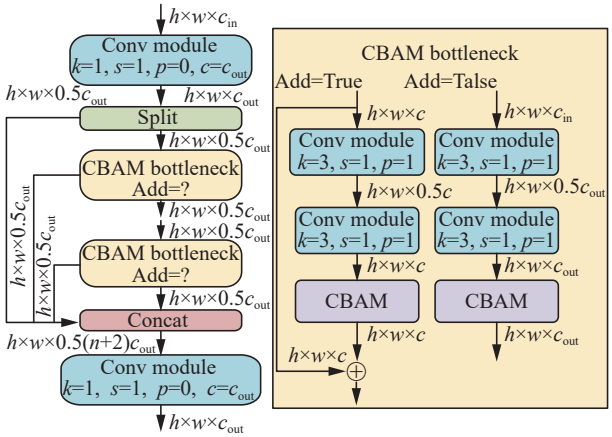


图 4　c2f_CBAM 结构

Fig. 4　c2f_CBAM structure

在 CBAM 模块的实现中，文章引入了两种注意力融合策略，当 Add 为 True 时，表示通道注意力和空间注意力的输出在融合时采用加法的方式，该策略更简洁，并能在不增加参数量的前提下提升注意力感知能力；而当 Add 为 False 时，则采用拼接 (concatenation, Concat) 后接 1×1 卷积的方式进行特征融合，具有更强的表达能力，但同时会增加计算和参数的开销。

此外，在图 4 中的 Concat 操作后，输出通道数为 0.5(n+2)，其中 n 表示中间 CBAM 模块数量，“2”来自于经过初始卷积划分出的两部分特征。

2.2　解决多尺度信息表达退化问题

复杂场景下目标检测任务中的图像信息会受到多种因素的影响导致图像信息呈现多尺度特点。例如成像视角、检测距离、光照条件、伪装手

段等进一步增加了检测任务的难度。为了应对尺度变化的问题，目标检测模型通常采用特征金字塔结构。其中，FPN(feature pyramid network)[23] 是最常用的自上而下特征金字塔结构。Yolov8 在 FPN 的基础上，采用 PANet 为特征金字塔网络添加了一条自下而上的路径，弥补了 FPN 中高层（high-level）特征中低层（low-level）特征细节的不足问题，然而，PANet 采用的自下而上的特征融合途径，虽然在某些情况下可以增强高层特征的表达能力，但模块的不足之处为低层特征的退化。Bi-FPN(bidirectional feature pyramid network)[24](自顶向下和自底而上) 的每个双向路径作为一个特征网络层，而且多次重复同一层，以实现更高层次的特征融合，但 BiFPN 引入了额外的连接和计算，这导致计算复杂度和内存使用的增加。文章使用的 AFPN(asymptotic feature pyramid network) 在结构层级融合策略上进行了重构，不再局限于对称式特征传递，而是采用“自底向上+层间跳连+渐进高层语义引入”的组合方式进行特征的融合。AFPN 通过优化特征层之间的融合机制，显著增强了不同层级之间的信息交互和整合能力。该网络特别强化了非相邻层级之间的特征融合，使得高层特征中的语义信息得以更好地保留，同时有效地维护了低层特征中的细节信息。通过这种设计有效减少了特征在传播和交互过程中的信息损失和退化。此外，AFPN 还通过改进的传播算法和优化的交互策略，最大限度地减少了特征在层级间传递时的质量损失。这种设计不仅提升了网络的整体性能，还增强了模型在处理复杂场景时的稳健性和可靠性。

如图 5（b）PANet 所示，根据 PAFPN 网络结构给出一列多尺度特征 $\boldsymbol{P}^{\text{in}} = (\boldsymbol{P}^{\text{in}}_{l_3}, \boldsymbol{P}^{\text{in}}_{l_4}, \cdots)$，其中 $\boldsymbol{P}^{\text{in}}_{l_i}$ 表示 $\boldsymbol{P}_{l_i}$ 层的特征，使其得到一个变换 f，能够输出一列新特征并对其特征进行融合 $\boldsymbol{P}^{\text{out}} = f(\boldsymbol{P}^{\text{in}})$，其中 $\boldsymbol{P}^{\text{in}}_{l_i}$ 特征层的分辨率为输入图像的 $1/2^i$。例如，输入图像的分辨率是 640×640，则 $\boldsymbol{P}^{\text{in}}_4$ 代表第 4 层特征的分辨率为 40×40（$1/2^4$），而代表第 6 层的分辨率为 10×10。PAFPN 以自顶向下的方式聚集多尺度特征：

$$\boldsymbol{P}^{\text{out}}_6 = \text{Conv}\left(\boldsymbol{P}^{\text{in}}_6\right)$$
$$\boldsymbol{P}^{\text{out}}_5 = \text{Conv}\left(\boldsymbol{P}^{\text{in}}_5 + \text{Resize}\left(\boldsymbol{P}^{\text{out}}_6\right)\right)$$
$$\vdots$$
$$\boldsymbol{P}^{\text{out}}_3 = \text{Conv}\left(\boldsymbol{P}^{\text{in}}_3 + \text{Resize}\left(\boldsymbol{P}^{\text{out}}_4\right)\right)$$

反之则为 PAFPN 以自下向上的特征融合方式。如图 5（d）所示，AFPN 在 Backbone 特征提取中，初始阶段结合不同分辨率的低级特征，逐步引入高级特征，最终融合顶级特征以增强表达能力。为避免层次间目标矛盾，AFPN 采用自适应空间融合 ASFF(adaptively spatial feature fusion)[25]，过滤多级融合特征，保留关键信息。

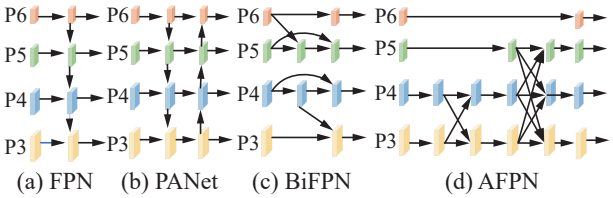


图 5　3~6 层采用不同特征融合网络

Fig. 5　Layers 3-6 use different feature fusion network

如图 6 所示，ASFF 融合了 3 个层次的特征，每一层特征图被输入到对应的 ASFF 模块（ASF_1、ASF_2、ASF_3）中。这些模块会从所有层次的特征图中选择性地融合空间信息，以生成一个优化的输出特征图。结果特征向量表示为 $\boldsymbol{P}^l_{i,j}$，特征向量的线性组合公式为

$$y^l_{ij} = \alpha^l_{ij} \cdot x^{1\to l}_{ij} + \beta^l_{ij} \cdot x^{2\to l}_{ij} + \gamma^l_{ij} \cdot x^{3\to l}_{ij}$$

式中：α^l_{ij}、β^l_{ij}、γ^l_{ij} 表示 3 个层特征在 l 层，(i, j) 位置的空间权重，$\alpha^l_{ij} + \beta^l_{ij} + \gamma^l_{ij} = 1$。

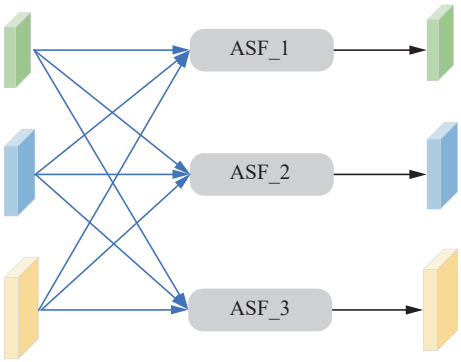


图 6　自适应空间网络结构

Fig. 6　Structure of the adaptive spatial network

2.3　基于 Inner-IoU 的尺度不均目标精确定位

由于目标远近交替，视角、距离多变导致检测目标出现大、中、小不同像素的数据集，同时目标检测涉及到从各种传感器和情报来源获取大量的数据。Inner_IoU 能根据不同检测器与检测任务进行自我调整，具有很强的泛化性，同时 Inner_IoU 更注重框的内部对齐情况，而不仅是外部轮廓的重叠，这使得 Inner_IOU 对于检测小目标或长宽比相差较大的目标更为敏感。Inner-IoU 损失函数通过比例因子 ratio 调节辅助边框的尺寸，从而在回归过程中对不同的样本进行区分时可以通过使用不同尺度的辅助边框来计算损失，有效加速了边框回归的过程。Inner_IoU 的描述如图 7 所示。

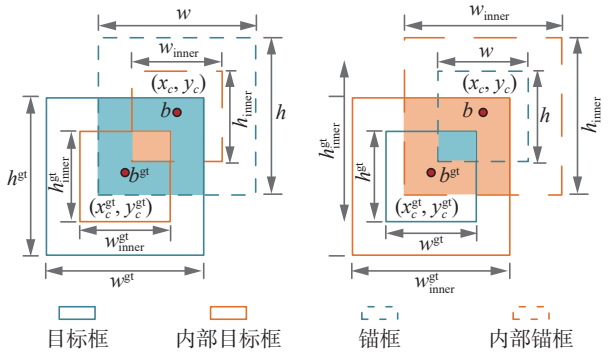


图 7　Inner_IoU 的描述

Fig. 7　Descriptive of Inner_IoU

Inner-IoU 损失函数通过引入辅助边框（根据缩放因子 ratio 调节尺寸）来优化边界框回归。当 ratio 小于 1 时，辅助边框尺寸变小，梯度增大，加速高 IoU 样本收敛；当 ratio 大于 1 时，辅助边框尺寸增大，扩展回归范围，提高低 IoU 样本回归效果。Inner-IoU 损失能够更精准地反映目标位置关系，从而提高检测精度并增强模型对目标的定位能力。该损失函数在不同 IoU 场景下平衡回归效果，并可与 GIoU(generalized intersection over union)、DIoU (distance intersection over union) 等现有 IoU 损失函数结合使用。

3　实验与结果分析

3.1　实验环境与参数设置

实验平台为华硕 DESKTOP-7ANF55M 电脑主机，硬件实验环境为 64 位 Windows11 操作系统，NVIDIA GeForce RTX 4080 GPU。软件环境为 Python3.9 编程环境，使用 PyTorch 学习框架，CUDA 版本为 11.6。

3.2　评价指标

在实验中采用平均精度均值（mean average precision，mAP）和计算量（floating point operations，FLOPs）。其中，mAP 反映了模型在多个阈值下的检测精度，是衡量模型准确性的重要指标。mAP 越高，表明模型的检测效果越好。FLOPs 用于评估模型在处理单张图像时的计算复杂度，FLOPs 越低，意味着模型计算资源需求越少。这些指标共同构成了评估目标检测模型性能的重要依据，综合考量了模型的准确性、效率与实际应用的可行性。mAP 计算公式为

$$m_{\mathrm{AP}} = \frac{1}{N}\sum \mathrm{AveragePrecision}(c)$$

式中：c 为军事目标数据集中的某一类别；N 为数据集中所有类别数量；AveragePrecision(·) 为平均精度，整个数据集的平均精度通过对所有类别的平均值计算得出。

3.3　对比实验

3.3.1　特征融合对比实验

为验证 AFPN 结构在 YOLOv8n 模型中的有效性，文章基于 YOLOv8n 模型进行 AFPN 与其他主流特征融合方式进行对比实验，评估其在特征融合和计算效率上的优势，实验结果见表 1。实验结果表明，AFPN 在 $\mathrm{mAP}_{0.5}$、$\mathrm{mAP}_{0.95}$ 以及 FLOPs 等关键指标上均表现优异。具体而言，由于 FPN 特征融合方式缺乏自下而上的特征融合层，导致其在多尺度特征的整合上存在不足，从而影响了整体的检测精度；而 BiFPN 通过引入权重来平衡不同尺度的特征信息，虽然在一定程度上改善了检测性能，但其复杂的结构导致计算量显著增加，增加了模型的计算开销；相比之下，AFPN 在 $\mathrm{mAP}_{0.5}$、$\mathrm{mAP}_{0.95}$ 以及 FLOPs 上均表现最佳。同时 AFPN 模块不仅保留了 PANet 的高效性，还通过更有效的特征融合策略进一步提升了检测精度，并且在保持较低计算量的同时实现了更高的检测性能，通过上述对比实验验证了 AFPN 模块替换 PANet 结构的有效性。

表 1　特征融合方法对比实验

Table 1　Comparison experiment of feature fusion methods

方法	$\mathrm{mAP}_{0.5}$/%	$\mathrm{mAP}_{0.95}$/%	FLOPs/10^9
FPN	81.0	56.3	8.0
PAFPN	81.5	56.8	8.1
BiFPN	81.8	56.2	8.4
AFPN	**82.1**	**56.9**	**7.2**

注：加粗数值表示该指标的最优结果。

3.3.2　主流算法对比实验

为了全面评估 AS-YOLO 方法在复杂环境下的检测性能，并验证其在实际应用中的优势。通过与当前主流的 YOLOv5n、YOLOv6n[26]、YOLOv7n[27] 和 YOLOv8n 目标检测算法进行性能对比。表 2 中的实验结果显示，AS-YOLO 在 $\mathrm{mAP}_{0.5}$ 和 $\mathrm{mAP}_{0.95}$ 两个指标上分别达到了 83.0% 和 56.9%，相较于 Faster-RCNN 分别提升了 11.4 百分点和 5.4 百分点，相较于 YOLOv5 分别提升了 1.8 百分点和 2.7 百分点，相较于 YOLOv6 提升了 2.8 百分点和 1.1 百分点，相较于 YOLOv8n，$\mathrm{mAP}_{0.5}$ 提升了 1.4 百分点，$\mathrm{mAP}_{0.95}$ 提升了 0.6 百分点。对比结果表明，AS-YOLO 表现出更高的检测精度。在计算量方面，相较于基础模型，AS-YOLO 提升了 0.4%。综合来看，AS-YOLO 不仅在复杂环境下的目标检测精度上优于 YOLO 系列的其他基础主流算法，同时在计算量上也具有显著的优势。

表 2　主流算法对比实验

Table 2　Comparative experiments with popular algorithms

方法	$mAP_{0.5}$/%	$mAP_{0.95}$/%	FLOPs/10^9
Faster-RCNN	77.6	46.0	—
YOLOv5	81.2	54.7	7.1
YOLOv6	80.2	56.3	11.8
YOLOv7	80.7	56.6	20.3
YOLOv8	81.5	56.8	8.1
AS_YOLO	**83.0**	**57.4**	**7.7**

注: 加粗数值表示该指标的最优结果, “—”表示实验未给出该指标数值。

综合来看, AS-YOLO 不仅在复杂环境下的目标检测精度上优于 YOLO 系列的其他基础主流算法, 同时在计算量方面也具有显著优势。AS-YOLO 在各项指标上的表现均证明了其卓越的性能和广泛的应用潜力。

3.3.3　失效增强对比实验

为了验证失效增强对于目标检测任务的有效性, 文章首先对基础模型 YOLOv8 和改进模型 AS_YOLO 进行失效增强策略训练。其次在复杂环境测试集上对基础 YOLOv8、增加失效增强的基础模型 YOLOv8、改进模型 AS_YOLO、增加失效增强的 AS_YOLO 进行测试。实验结果如表 3 所示。

表 3　失效增强对比实验

Table 3　Comparative experiment of failure enhancement

方法	$mAP_{0.5}$/%	$mAP_{0.75}$/%	$mAP_{0.95}$/%
YOLOv8n	77.6	66.0	58.9
YOLOv8n+失效增强	92.4	78.2	70.6
AS_YOLO	81.6	68.3	61.6
AS_YOLO+失效增强	**94.0**	**80.1**	**72.5**

注: 加粗数值表示该指标的最优结果。

测试结果显示, 添加失效增强的基础模型 YOLOv8 在 mAP0.5 上提高了 14.8 百分点, 在 $mAP_{0.75}$ 上提高了 12.2 百分点, 在 $mAP_{0.95}$ 上提高了 11.7 百分点。添加失效增强的 AS_YOLO 模型在 $mAP_{0.95}$ 上提高了 12.4 百分点, 在 $mAP_{0.75}$ 上提高了 11.8 百分点, 在 $mAP_{0.95}$ 上提高了 10.9 百分点。

显然, 没有接受失效增强训练的模型在复杂环境测试集上的检测精度远不如进行失效增强后的模型。结果表明添加失效增强模块使得模型精度得到了全面提升, 证明本文使用, 天气、光照、干扰、遮挡、缩放等失效增强, 模拟了复杂环境的真实场景, 还原了真实复杂场景的干扰, 丰富了图像的细节纹理信息和语义信息, 同时增加了样本数量, 使目标检测模型更加符合复杂场景下海面场景的检测需求, 更具有鲁棒性。

3.4　消融实验

实验以 YOLOv8n 为基线模型, 通过逐步添加各个改进模块来评估其对模型检测性能的影响。首先, 在基线模型上应用 c2f_CBAM 模块, 以增强特征提取的精度; 接着, 添加 AFPN 渐进特征金字塔融合模块, 以优化不同层级特征的融合过程; 最后, 引入 Inner_IOU 损失函数, 进一步提高模型对目标位置的精确度。所有实验均在相同条件下进行, 以公平评估不同模块对检测效果的贡献。

在基线模型上加入 c2f_CBAM 模块后, 模型在复杂环境下的小目标检测精度有所提升, 具体实验结果为表 4。$mAP_{0.5}$ 和 $mAP_{0.95}$ 这两个关键指标分别提高了 0.4 百分点和 0.3 百分点。这一提升表明 c2f_CBAM 模块在增强模型对小目标及其周围背景的特征关注能力方面起到了显著作用, 从而提升整体检测性能。此外, 将 AFPN 模块引入颈部网络后, $mAP_{0.5}$ 和 $mAP_{0.95}$ 分别提高了 0.6 百分点和 0.9 百分点, 进一步提升了模型的整体检测性能。这些结果表明, 所应用的改进模块在提升检测精度方面发挥了积极作用, 尤其是在复杂场景中的小目标检测任务中, 改进效果尤为显著。AFPN 在所有比较方法中具有最低的 FLOPs 主要归功于特征维度的降低, 表明了其在资源利用上的高效性。引入 inner_IoU 后 $mAP_{0.5}$ 和 $mAP_{0.95}$ 相比基线模型分别提升 1.5 百分点和 0.6 百分点。

表 4　消融实验对比

Table 4　Comparison of ablation experiments

方法	$mAP_{0.5}$/%	$mAP_{0.95}$/%	FLOPs/10^9
YOLOv8n	81.5	56.8	8.1
+CBAM	81.0	56.6	8.2
+c2f_CBAM	81.9	57.1	8.1
+AFPN	82.1	57.7	7.2
+c2f_CBAM AFPN	82.4	56.9	7.7
AS_YOLO	**83.0**	**57.4**	**7.7**

注: 加粗数值表示该指标的最优结果。

为了更加直观地突出消融实验中各个模块的有效性, 图 8 为检测结果的可视化。与基线模型 YOLOv8n 相比, 本文所提出的改进方法在检测精度、误检、漏检等方面均表现出一定程度的提升。在 YOLOv8 检测结果图中可以看出, 模型将舰船误检测成了导弹运输发射车, 同时还漏检了小目标的坦克。通过引入空间–通道注意力机制模块, 消除了模型的误检问题。模型也能够捕捉到更多的小目标, 改善基础模型漏检问题。通过改进特征融合模块提升模型在复杂环境下的检测

能力，特别是被遮挡的坦克。联合使用这 3 种改进模块能够进一步显著增强模型的检测精度，特别是在处理背景杂乱或目标尺寸较小的情况下，模型的表现会更加出色。

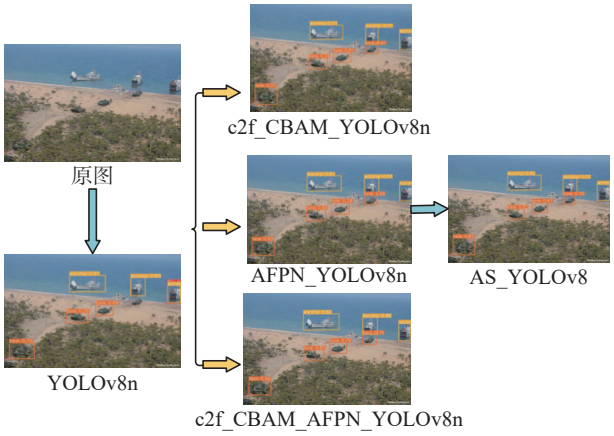
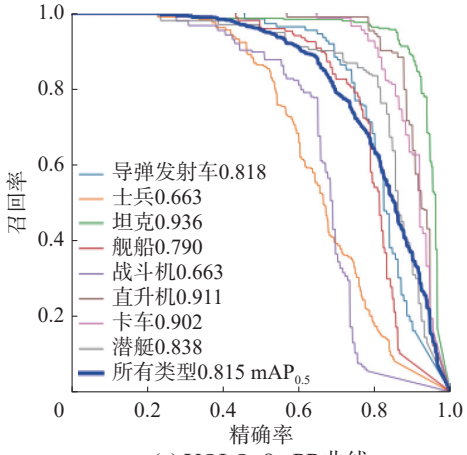


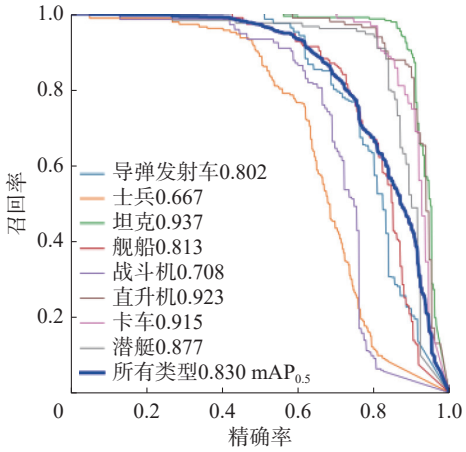
图 8　各模块检测结果

Fig. 8　Test results of each module

图 9 给出了本方法与基线模型在测试集上各类别的 mAP 值和总的 mAP 值，以及相应的 PR 曲线 (precision-recall curve)，曲线的横坐标为精确率，纵坐标为召回率。具体来说，图 9(a) 给出了 YOLOv8n 模型的 PR 曲线，图 9(b) 则给出了改进后的 AS-YOLO 模型的 PR 曲线。从整体趋势来看，AS-YOLO 模型在多数目标类别上均实现了不同程度的性能提升。对于士兵和坦克等目标，改进后的模型在保持较高召回能力的同时，其 AP 值均较 YOLOv8n 有所提升。在舰船和飞机等结构特征复杂、尺度变化较大的类别上，AS-YOLO 的 AP 提升幅度更为明显。对于直升机、卡车及潜艇等类别，AS-YOLO 同样取得了稳定的性能提升。从整体检测性能来看，AS-YOLO 模型的 $mAP_{0.5}$ 由 YOLOv8n 的 81.5 % 提升至 83.0%，进一步验证了所提出改进方法在提升模型整体检测精度和泛化能力方面的有效性。



(a) YOLOv8n PR曲线



(b) AS_YOLOv8n PR曲线

图 9　YOLOv8n 和 AS_YOLO 的 PR 曲线

Fig. 9　PR curve of YOLOv8n and AS_YOLO

这一结果表明，AS-YOLO 在保持高效检测的同时，进一步增强了模型的泛化能力。这些改进源于模型结构优化和特征提取能力的提升，模型能够更好地捕捉和区分细粒度的特征，最终实现了全类别范围内的性能提升。总之，AS-YOLO 在各类目标检测任务中展示了优越的表现，验证了本方法的有效性和实用性。

3.5　模型泛化性分析

为了证明文章方法的泛化能力，本文使用 YOLOv8n 检测器在 AI-TOD 数据集[28] 上进行了对比实验。AI-TOD 包含 28 036 张图片，8 类目标以及 700 621 个目标实例，其中 14 018 张图像作为训练集中另外的 14 018 张图像作为测试集。相较于其他常用的目标检测数据集，平均尺寸为 12.8 像素的 AI-TOD 数据集可以减小计算资源的需求，加快模型训练速度。实验结果如表 5 所示。对于复杂环境下小目标低 IoU 样本，较大尺度的辅助边框的 IoU 梯度的绝对值大于实际边框 IoU 梯度的绝对值。可以看出，应用文章方法后，检测效果有所提高，$mAP_{0.5}$ 提高了 1.5 百分点，$mAP_{0.95}$ 提高了 0.6 百分点。

表 5　不同 ratio 对比结果

Table 5　Compare the results of different ratios　　%

IOU	$mAP_{0.5}$	$mAP_{0.95}$
CIOU	81.5	56.8
Inner_IOU (ratio=1.10)	82.1	57.1
Inner_IOU (ratio=1.13)	81.9	57.1
Inner_IOU (ratio=1.15)	83.1	57.4

3.6　可视化分析

利用梯度加权类激活映射（gradient-weighted class activation mapping, Grad-CAM）[29] 对不同的

网络进行可视化，如图 10 所示，可以发现相较于 YOLOv8n，引入 c2f_CBAM 后，特征覆盖到了待检测目标的更精准的范围，而且最终检测物体的几率也更高，这表明 c2f_CBAM 注意力机制的确让网络学会了关注待检测目标特征信息。AS_YOLOv8 在预测时更多赋予图像不同目标的重要性，更多的关注目标本身的特征信息，在小目标检测的精度方面，本方法有较大的优势。

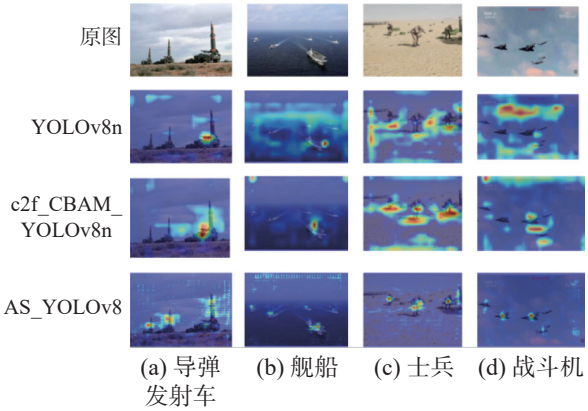


图 10　不同模型的可视化热力图

Fig. 10　Visualization heat maps of different models

3.7　迁移实验

为充分验证本文方法的泛化性和鲁棒性，本文在 WSODD 数据集[30] 上进行迁移实验，WSODD 数据集由海康威视工业拍摄的 7 467 张水面图像组成，每张图像的分辨率为 1 920×1 080。数据集包括海洋、湖泊和河流等广泛的环境，白天、黄昏和夜晚等不同的光照条件和晴天、阴天和雾天等不同的天气条件。数据集中有 14 个类别 (bridge、ship、boat、ball、rubbish、rock、buoy、platform、habor、mast、tree、animal、grass、person)，21 911 个目标实例。文章按照 7:2:1 的比例将图像样本划分到训练集、验证集和测试集。

迁移实验结果如表 6 所示，本文提出的添加失效增强的改进模型 AS_YOLOv8 的 $mAP_{0.5}$ 和 $mAP_{0.95}$ 精度均最佳，分别为 81.4% 和 47.9%，这是由于本文提出的失效增强模拟真实场景可能出现的天气、光照，遮挡等复杂情况，使模型学习到图像更丰富细节纹理信息和语义信息，使目标检测模型更加符合复杂场景下海面场景的检测需求。同时结果表明改进模型 AS_YOLOv8 的精度相较于基础模型 YOLOv8n 和 YOLOv5n 均有提升，表明改进后的模型在精度上优于 YOLO 基础模型。

为了更加直观地说明本文提出的面向应用环境的失效增强的改进模型 AS_YOLO 的有效性，图 11 给出复杂环境下的图像对比结果。

表 6　迁移实验

Table 6　Migration experiments

方法	$mAP_{0.5}$/%	$mAP_{0.95}$/%	FLOPs/10^9
YOLOv8n	73.8	41.5	8.5
YOLOv5n	71.5	40.1	7.4
AS_YOLOv8	75.9	42.8	8.1
AS_YOLOv8+失效增强	**81.4**	**47.9**	**8.3**

注：加粗数值表示该指标的最优结果。

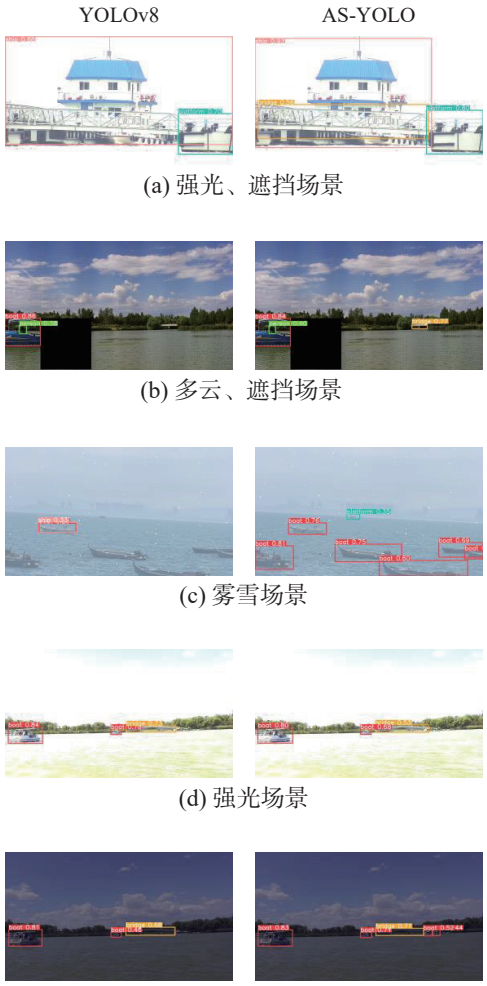


图 11　复杂环境下的图像对比

Fig. 11　Image comparison in complex environments

第一列为基础模型 YOLOv8 的最优模型检测结果，第二列为添加失效增强和改进模型 AS_YOLO 最优模型的检测结果。如图 11(a)所示，由于画面中的桥和背景中的船有重叠部分，导致基础模型出现漏检。图 11(b)、(d)和(e)由于图像分别添加了遮挡、晴天曝光和黎明等干扰，基础模型均出现了漏检现象。而图 11(c)添加了基础失效组合增强，导致模型不仅出现了漏检还出现了错检。而添加了 c2f_CBAM 注意力机制的改进模型能够更好关注目标信息，同时由于 AFPN

结构的引入，强化了非相邻层级之间的特征融合，使得高层特征中的语义信息得以更好地保留，同时有效地维护了低层特征中的细节信息，进一步提高了对复杂场景下目标的特征融合能力。图 11(a)、(c)表明无论是大目标还是小目标改进后的模型都表现出较好的检测效果，这是由于 Inner_IoU 能根据不同尺度下的目标进行自我调整辅助边框。综上，本文提出的方法在复杂环境下进行目标检测有较大的优势。

4 结束语

相较于原始算法，文章所提出的结合改进 YOLOv8 和面向应用环境的失效增强算法在检测精度和计算量上均有提升，特别是针对复杂环境下遮挡、光照变化导致的漏检、误检问题，改进后的模型都有显著提升。同时模型泛化性分析表明文章所提出的方法具有较好的泛化能力，为复杂环境下的目标检测提供参考。后续可以考虑对复杂环境下的目标数据集采用深度学习图像增强技术进行扩充以增加样本丰富性；同时，针对所提出的模型，需细致考虑不同光照条件、天气变化、背景多样性、拍摄视角及距离差异等因素导致的目标图像模糊、形变、尺寸多变等挑战，进一步提升网络精度和检测速度，来满足真实复杂环境下的准确性和实时性。

参考文献:

[1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84−90.

[2] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6): 1137−1149.

[3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779−788.

[4] HUANG Xiaochen, WANG Xiaofeng, TENG Qizhi, et al. Degradation type-aware image restoration for effective object detection in adverse weather[J]. Sensors, 2024, 24(19): 6330.

[5] WANG Zhipan, LIU Di, WANG Zhongwu, et al. A new remote sensing change detection data augmentation method based on mosaic simulation and haze image simulation[J]. IEEE journal of selected topics in applied earth observations and remote sensing, 2023, 16: 4579−4590.

[6] WU Junjun, RAO Yunbo, ZENG Shaoning, et al. Pre-trained SAM as data augmentation for image segmentation[J]. CAAI transactions on intelligence technology, 2025, 10(1): 268−282.

[7] 肖晶晶, 樊博彦, 杨雨婷. 雾环境下的船舶目标检测研究[J]. 重庆理工大学学报 (自然科学), 2024, 38(3): 212−219.
XIAO Jingjing, FAN Boyan, YANG Yuting. Research on ship object detection in foggy environments[J]. Journal of Chongqing University of Technology (natural science), 2024, 38(3): 212−219.

[8] 马淦, 谷雨, 彭冬亮. 结合改进 YOLOv5s 和动态数据增强的海面舰船检测[J]. 计算机工程, 2025, 51(9): 294−305.
MA Gan, GU Yu, PENG Dongliang. Combining improved YOLOv5s and dynamic data augmentation for sea surface ship detection[J]. Computer engineering, 2025, 51(9): 294−305.

[9] FAN Pan, ZHENG Chusan, SUN Jin, et al. Enhanced real-time target detection for picking robots using lightweight CenterNet in complex orchard environments[J]. Agriculture, 2024, 14(7): 1059.

[10] 邢汇源, 崔亚奇, 王子玲, 等. 复杂海况下的海上船舶目标检测算法[J]. 现代防御技术, 2024, 52(6): 88−96.
XING Huiyuan, CUI Yaqi, WANG Ziling, et al. Target detection algorithm for ships at sea under complex sea conditions[J]. Modern defence technology, 2024, 52(6): 88−96.

[11] 张国印, 王传博, 高伟. 抗遮挡的行人多目标跟踪算法[J]. 智能系统学报, 2024, 19(5): 1248−1256.
ZHANG Guoyin, WANG Chuanbo, GAO Wei. Pedestrian multiobject tracking algorithm with anti-occlusion[J]. CAAI transactions on intelligent systems, 2024, 19(5): 1248−1256.

[12] LYU Yunkai, YANG Xiaobing, GUAN Ai, et al. Construction personnel dress code detection based on YOLO framework[J]. CAAI transactions on intelligence technology, 2024, 9(3): 709−721.

[13] 吴攀超, 郑卓纹, 王婷婷, 等. 基于 CF-YOLO 的雾霾交通标志识别[J]. 计算机工程与设计, 2024, 45(7): 2203−2211.
WU Panchao, ZHENG Zhuowen, WANG Tingting, et al. Foggy traffic sign recognition based on CF-YOLO[J]. Computer engineering and design, 2024, 45(7): 2203−2211.

[14] 赵文清, 康怿瑾, 赵振兵, 等. 改进 YOLOv5s 的遥感图像目标检测[J]. 智能系统学报, 2023, 18(1): 86−95.
ZHAO Wenqing, KANG Yijin, ZHAO Zhenbing, et al. A

remote sensing image object detection algorithm with improved YOLOv5s[J]. CAAI transactions on intelligent systems, 2023, 18(1): 86−95.

[15] 许迪, 张淑卿, 葛超. 面向复杂环境的 YOLOv8 安全装备检测[J]. 电子测量技术, 2024, 47(7): 121−129.
XU Di, ZHANG Shuqing, GE Chao. YOLOv8 security equipment inspection for complex environments[J]. Electronic measurement technology, 2024, 47(7): 121−129.

[16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision–ECCV 2018. Cham: Springer, 2018: 3−19.

[17] YANG Guoyu, LEI Jie, ZHU Zhikuan, et al. AFPN: asymptotic feature pyramid network for object detection [C]//2023 IEEE International Conference on Systems, Man, and Cybernetics. Honolulu: IEEE, 2024: 2184−2189.

[18] ZHANG Hao, XU Cong, ZHANG Shuaijie. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[EB/OL]. (2023−11−06)[2025−03−06]. https://arxiv.org/abs/2311.02877.

[19] WANG Weijun, HOWARD A. Mosaic: mobile segmentation via decoding aggregated information and encoded context[EB/OL]. (2021−12−22)[2025−03−06]. https://arxiv.org/abs/2112.11623.

[20] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759−8768.

[21] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial Transformer networks[J]. Advances in neural information processing systems, 2015, 28: 1−9.

[22] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132−7141.

[23] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 936−944.

[24] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10778−10787.

[25] LIU Songtao, HUANG Di, WANG Yunhong. Learning spatial fusion for single-shot object detection[EB/OL]. (2019−11−21)[2025−03−06]. https://arxiv.org/abs/1911.09516.

[26] LI Chuyi, LI Lulu, JIANG Hongliang, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022−09−07)[2025−03−06]. https://arxiv.org/abs/2209.02976.

[27] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7464−7475.

[28] WANG Jinwang, YANG Wen, GUO Haowen, et al. Tiny object detection in aerial images[C]//2020 25th International Conference on Pattern Recognition. Milan: IEEE, 2021: 3791−3798.

[29] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. International journal of computer vision, 2020, 128(2): 336−359.

[30] ZHOU Zhiguo, SUN Jiaen, YU Jiabao, et al. An image-based benchmark dataset and a novel object detector for water surface object detection[J]. Frontiers in neurorobotics, 2021, 15: 723336.

作者简介：

储文娟，硕士研究生，主要研究方向为视觉目标识别与图像处理。E-mail: 2294806304@qq.com。

李震，教授，博士，主要研究方向为可靠性与系统工程。主持国家级项目 2 项、省部级项目 2 项、横向项目 20 余项。E-mail: justlz@just.edu.cn。

黄炜嘉，副教授，博士，主要研究方向为图像处理。主持横向项目 1 项，参研国家自然科学基金面上项目 2 项。E-mail: huangweijia@just.edu.cn。