文章编号:1007-2780(2026)08-1144-14

基于改进YOLO11n的野外小目标检测算法

王　岑,王彩霞*,刘　鹏
(长春理工大学 电子信息工程学院,吉林 长春 130022)

摘要:针对野外环境中目标尺度小、纹理信息弱、背景干扰复杂以及边界框定位偏差明显等问题,提出一种基于改进YOLO11n的野外小目标检测算法。首先,引入自适应亚像素上采样(ASU)模块,通过亚像素重排和注意力掩码,增强小目标边缘和纹理细节特征;其次,引入多尺度显著性增强模块(MultiSEM),通过多尺度空洞卷积和双池化通道注意力,增强小目标区域响应并抑制复杂背景干扰;最后,构建显著性指数交并比损失函数SE-CIoU,通过引入中心距离指数惩罚和遮挡感知形状约束,提高小目标检测的定位精度和稳定性。实验结果表明,在自建WDD数据集上,改进算法的精度、召回率和mAP@0.5分别为88.8%、91.4%和94.0%,较YOLO11n基准网络算法分别提高0.6%、6.6%和3.3%;在iWildCam2022-Small数据集上进行的泛化实验表明,mAP@0.5达到64.5%,较YOLO11n提高3.7%。所提算法能够在保持轻量化检测框架的基础上,有效增强野外小目标的细节表达、显著性响应和边界框定位能力,提高复杂野外场景下小目标检测精度,并具备一定泛化能力。

关　键　词:野外小目标;目标检测;YOLO11n;亚像素上采样;边界框回归
中图分类号:TP391.4　　**文献标识码**:A　　**doi**:10.3724/CJLCD.2026-0108　　**CSTR**:32172.14.CJLCD.2026-0108

Wild small object detection algorithm based on improved YOLO11n

WANG Cen, WANG Caixia*, LIU Peng

(*School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun* 130022, *China*)

Abstract: To address the problems of small target scale, weak texture information, complex background interference, and obvious bounding-box localization deviation in wild environments, an improved YOLO11n-based algorithm for wild small object detection is proposed. First, an adaptive sub-pixel upsampling(ASU) module is introduced to enhance feature fusion for small objects. Through sub-pixel rearrangement and attention masks, ASU strengthens the representation of edge and texture details, thereby reducing the loss of fine-grained information during upsampling. Second, a multi-scale saliency enhancement module(MultiSEM) is introduced before the detection head. By using multi-scale dilated convolutions, the module extracts spatial features under different receptive fields, and dual-pooling channel attention is further used to enhance the response of small target regions while suppressing interference from complex backgrounds. Finally, a saliency-

收稿日期:2026-06-16;**修订日期**:2026-07-15.
基金项目:吉林省科技发展计划(No. 20250102228JC)
Supported by Jilin Provincial Science and Technology Development Plan (No. 20250102228JC)
*通信联系人,E-mail:wcxhao@sina.com

exponential Complete Intersection over Union loss function (SE-CIoU) is constructed for bounding-box regression. By introducing a center-distance exponential penalty and occlusion-aware shape constraints, SE-CIoU strengthens the position and shape constraints between predicted boxes and ground-truth boxes, thus improving the localization accuracy and stability of small object detection. Experimental results show that, on the self-built WDD dataset, the improved algorithm achieves a precision of 88.8%, a recall of 91.4%, and an mAP@0.5 of 94.0%, which are 0.6, 6.6, and 3.3 percentage points higher than those of the YOLO11n baseline network algorithm, respectively. Generalization experiments on the iWildCam2022-Small dataset show that the mAP@0.5 reaches 64.5%, which is 3.7 percentage points higher than that of YOLO11n. The results indicate that the proposed algorithm can effectively enhance detail representation, saliency response, and bounding-box localization ability for wild small objects while maintaining a lightweight detection framework. Therefore, the improved YOLO11n algorithm improves small object detection accuracy in complex wild scenes and shows certain generalization ability.

Key words: wild small object; object detection; YOLO11n; subpixel upsampling; bounding box regression

1　引　　言

野外小目标检测是目标检测领域中具有代表性的复杂场景任务，在野生动物监测、生态保护、野外安防和无人平台感知等方面具有重要应用价值。与一般场景目标检测相比，野外环境中的目标受拍摄距离、相机视角、运动状态和自然遮挡等因素影响，在图像中呈现出尺度小、像素占比低、边缘轮廓弱和纹理细节不足等特点。同时，野外环境中普遍存在植被、地形、阴影和光照变化等复杂背景干扰，目标与背景之间的显著性差异不明显，容易导致检测模型出现漏检、误检和边界框定位偏差等问题[1-3]。因此，如何合理改进小目标检测模型，提高野外复杂环境中小目标检测的精度和定位稳定性，成为相关工作人员的研究焦点。

目前，学术界针对小目标检测提出了多种改进方法，从特征上采样与细节重建、多尺度特征提取与注意力机制、边界框回归损失函数优化等方面开展了研究[4-5]。其中，特征上采样与细节重建是缓解小目标在特征融合过程中细节丢失问题的重要方法。张志豪等[6]通过改进 YOLO11 网络结构提升了无人机航拍小目标的检测效果；Zhang 等[7]提出 MSD-YOLO11n 以增强多尺度特征表达；曾子豪等[8]提出基于改进 YOLO11n 的导电粒子目标检测算法，提升了细小目标检测效果；杨晓敏等[9]提出面向 SAR 影像超轻量化小目标检测网络，兼顾了检测精度与模型规模；Yan 等[10]提出 TOE-YOLO 以优化微小目标特征提取和检测结构；Liu 等[11]通过学习采样位置改善特征重建过程；Shi 等[12]利用亚像素卷积将通道维度特征重排至空间维度，实现图像细节恢复。特征上采样与细节重建能够在一定程度上缓解小目标特征在特征融合过程中的损失，但在野外复杂背景中，若缺少对目标区域的自适应调节，可能会放大背景纹理噪声，影响小目标有效特征的表达。

多尺度特征提取与注意力机制是提高复杂背景下小目标特征表达能力的重要方法。Ma 等[13]和 Yang 等[14]通过在 YOLOv5s 中引入注意力机制，增强了目标特征提取能力，提高了复杂森林环境下动物目标检测性能；卢汉等[15]针对夜间野生动物目标检测任务提出半监督端到端检测方法，提高了夜间复杂场景下的检测效果；Du 等[16]面向相机陷阱图像中的保护动物检测任务，通过轻量化结构和特征融合改进，兼顾了检测精度和模型复杂度；Zhu 等[17]提出 YOLO-WildASM 算法，通过增加小目标检测层和加权特征融合结构，增强了复杂环境下小目标检测能力；Liu 等[18]通过多尺度卷积、空间注意力和浅层特征融合，实现了无人机野生动物目标的高效检测；Chen 等[19]通过改进 YOLOv8n 的特征交互和轻量化结构，增强了生态场景下目标检测效果；Woo 等[20]提出卷积块注意力模块 CBAM，通过全局平均池化和全局最大池化建模通道注意力，对输入特征进行自适应重标定。尽管多尺度上下文信息和注意力机制有助于增强目标响应，但野外小目标存在纹理弱、对比度低和局部遮挡等问题，普通注

意力机制仍可能受到强背景响应影响，导致小目标区域显著性不足。

边界框回归损失函数优化是提高小目标定位精度的重要方法。小目标像素占比小且边界信息不完整，在遮挡或背景干扰时，预测框容易出现中心偏移和尺度回归不稳定等问题。Zheng等[21]在目标检测和实例分割任务中增强几何因素建模，改进边界框回归过程，提高了目标定位精度。然而，常规边界框回归约束仍难以充分适应野外小目标对中心偏移敏感、遮挡轮廓不完整和尺度变化明显等问题，其定位稳定性仍有提升空间。

综上所述，现有方法虽然能够从不同方面提高小目标检测性能，但野外小目标检测仍面临细节特征弱、背景干扰强和定位稳定性不足等问题。YOLO系列是经典的单阶段目标检测算法[22]，具有端到端检测、推理速度快和部署成本低等优点，被广泛应用于实时目标检测任务。本文选择YOLO11n[23]作为基准算法，并从特征上采样、多尺度显著性增强和边界框回归损失函数3个方面进行改进。首先，引入自适应亚像素上采样（Adaptive Sub-pixel Upsampling，ASU）模块，通过亚像素重排以及注意力掩码，增强小目标边缘和纹理细节；其次，引入多尺度显著性增强模块（Multi-scale Saliency Enhancement Module，Multi-SEM），通过不同膨胀率的空洞卷积提取多尺度空间信息，并结合双池化通道注意力增强小目标相关特征响应，抑制复杂背景干扰；最后，构建显著性指数交并比损失函数（Saliency-Exponential Complete Intersection over Union，SE-CIoU），优化边界框回归过程，提高小目标定位稳定性。以上改进能够提高算法对野外环境下小目标的检测性能。

2 YOLO11n算法

YOLO11n是YOLO11系列[23]中的轻量化目标检测模型，具有参数量小、检测速度快等特点，本文选取YOLO11n作为野外小目标检测的基准算法。YOLO11n网络结构主要由输入端（Input）、主干网络（Backbone）、特征融合网络（Neck）和检测头（Head）4部分组成。输入端主要负责读取图像和对应的标签文件，以及在输入网络之前对图像进行自适应缩放、数据增强和自适应锚框计算等预处理。主干网络用于提取目标的多层特征信息，主要由卷积模块、C3k2模块、SPPF模块和C2PSA模块组成，能够在较低计算量下获得较强的特征表达能力。特征融合网络通过上采样、特征拼接和卷积操作融合不同尺度的特征图，以增强模型对多尺度目标的检测能力，减少小目标在下采样过程中的细节信息丢失。检测头基于融合后的多尺度特征图分别进行边界框回归和类别预测，生成候选目标的位置与类别置信度，并在后处理阶段利用非极大值抑制筛除冗余预测框，得到最终检测结果。YOLO11n在检测速度和模型复杂度之间具有较好的平衡，适合作为野外小目标检测任务的基准模型。

3 改进YOLO11n的野外小目标检测算法

针对野外小目标细节特征易丢失、背景干扰强和定位偏差等问题，本文从特征上采样、多尺度显著性增强和边界框回归损失函数3个方面对YOLO11n进行改进，提出改进YOLO11n的野外小目标检测算法，以提升目标检测性能。改进后的算法网络结构如图1所示。

3.1 ASU模块设计

在野外小目标检测中，目标像素占比低，其边缘、轮廓和局部纹理在深层特征图中响应较弱。YOLO11n颈部网络在特征融合时对深层语义特征进行插值上采样操作，仅依据固定位置关系恢复特征图尺寸，缺少对目标边缘、轮廓和局部纹理的自适应重建能力，导致小目标细粒度特征平滑。为提高特征融合阶段网络对小目标细节信息的保留能力，本文在YOLO11n颈部网络中引入了自适应亚像素上采样模块ASU，其结构如图2所示。

ASU模块由基底分支、细节分支和掩码分支组成。图2中，ASU首先采用1×1卷积对输入特征$X\in R^{C\times H\times W}$进行通道映射，得到共享中间特征$M$。基底分支对输入特征$X$进行上采样，生成尺度放大后的基底特征$B$，用于保留整体语义结构。细节分支先采用分组卷积（Group Conv）对共享中间特征M进行局部特征变换，再结合亚像素重排（Pixel Shuffle）[12]将通道维度中的细节响

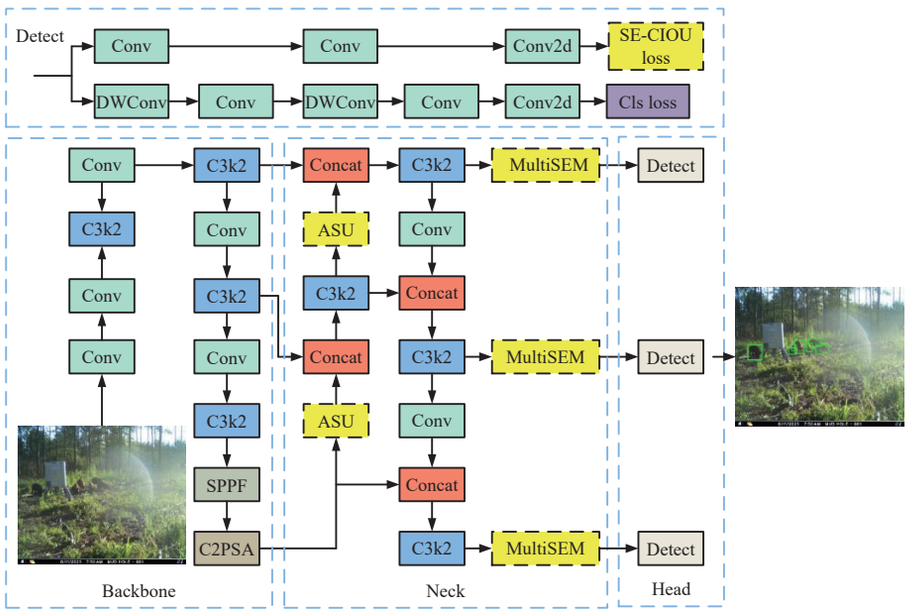


图 1　改进 YOLO11n 算法结构

Fig. 1　Structure of the improved YOLO11n algorithm

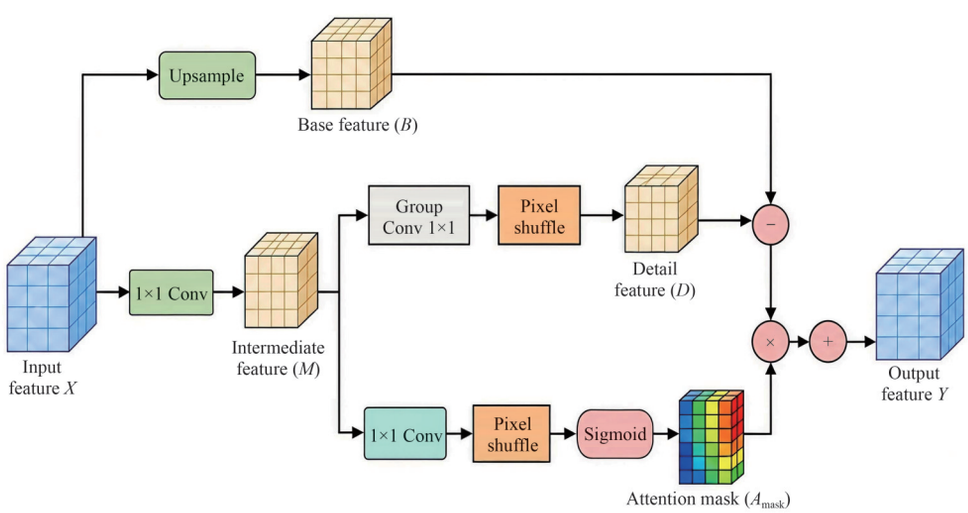


图 2　ASU 模块结构

Fig. 2　Structure of ASU

应映射到空间维度，得到细节特征 D。掩码分支同样以共享中间特征 M 为输入，通过卷积、亚像素重排和 Sigmoid 函数得到注意力掩码 A_{mask}，用于控制不同空间位置的细节注入强度。

最终，ASU 利用注意力掩码 A_{mask} 对基底特征 B 和细节特征 D 进行自适应融合，得到输出特征 Y，如式(1)所示：

$$Y = B + A_{\text{mask}} \odot (D - B), \tag{1}$$

式中：$\odot$ 表示逐元素相乘，A_{mask} 在融合过程中沿通道维度进行广播。由式(1)可知，当 A_{mask} 趋近于 1 时，输出特征更接近细节特征 D，有利于增强小目标区域的边缘、纹理和局部结构信息；当 A_{mask} 趋近于 0 时，输出特征主要保留基底特征 B，从而避免野外复杂背景纹理被过度放大。

3.2　MultiSEM 模块设计

野外场景中存在植被、阴影、地形纹理和遮挡等干扰，小目标与背景之间的特征差异不明显，容易导致前景响应弱化和背景混淆。因此，本文在检测头前引入多尺度显著性增强模块(MultiSEM)。其结构如图 3 所示。

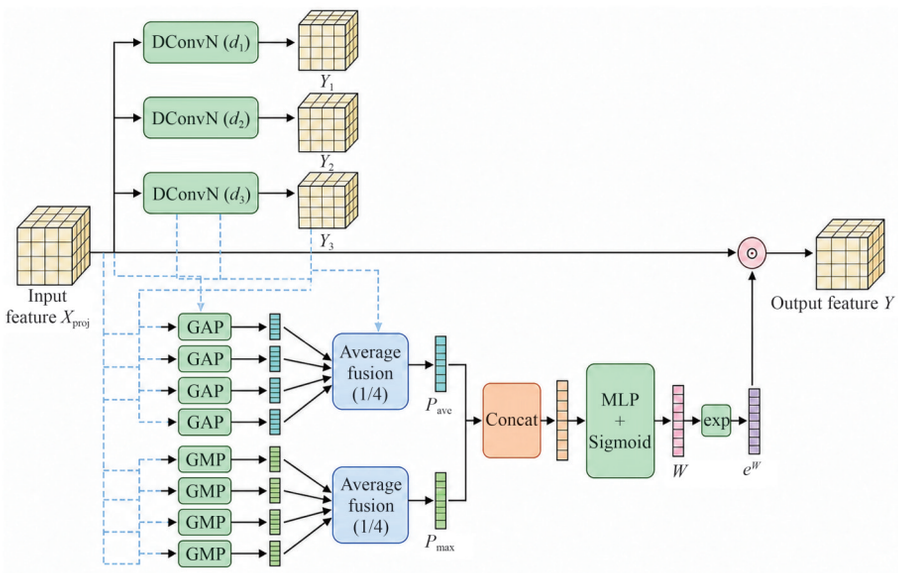


图 3 MultiSEM 结构

Fig. 3 Structure of MultiSEM

图 3 中，首先将 $X_{proj}\in R^{C\times H\times W}$ 分别输入到具有不同膨胀率 d_1、d_2 和 d_3 的 3 个并行空洞卷积分支，提取到不同尺度空间特征 Y_1、Y_2、Y_3。其次，再对输入特征 X_{proj} 以及 3 个尺度特征 Y_1、Y_2、Y_3 进行全局平均池化和全局最大池化。平均池化和最大池化后的输出分别记为 P_{ave} 和 P_{max}，如式(2)和式(3)所示：

$$P_{ave}=\frac{1}{4}\left(\mathrm{GAP}\left(X_{proj}\right)+\sum_{i=1}^{3}\mathrm{GAP}\left(Y_i\right)\right),\quad P_{ave}\in R^{C\times 1\times 1},\tag{2}$$

$$P_{max}=\frac{1}{4}\left(\mathrm{GMP}\left(X_{proj}\right)+\sum_{i=1}^{3}\mathrm{GMP}\left(Y_i\right)\right),\quad P_{max}\in R^{C\times 1\times 1},\tag{3}$$

式中，GAP 和 GMP 分别表示全局平均池化和全局最大池化。其中全局平均池化用于描述通道整体响应，全局最大池化用于突出局部显著响应。

然后，将 P_{ave} 和 P_{max} 在通道维度进行拼接，并输入多层感知机(Multi-Layer Perceptron，MLP)和 Sigmoid 激活函数，生成通道维度的注意力权重 W。为了减弱普通注意力机制对于小目标响应的抑制，对 W 进行指数映射，得到指数增强权重 e^W。最终，将指数增强权重 e^W 作用于输入特征 X_{proj}，得到 MultiSEM 的输出特征 Y，如式(4)所示：

$$Y=X_{proj}\odot e^W.\tag{4}$$

通过上述过程，MultiSEM 能够结合多尺度空间特征，并利用双池化通道注意力生成显著性增强权重，从而增强小目标相关特征响应。

3.3 SE-CIoU 损失函数

YOLO11 采用 CIoU 损失函数[21]，通过引进真实框和预测框的重叠面积、中心点距离和宽高比，来解决目标定位问题，使预测框更接近真实框，其计算公式如式(5)所示：

$$\begin{cases}L_{CIoU}=1-\mathrm{IoU}+\dfrac{d^2\left(b,b^{gt}\right)}{c^2}+\alpha v\\ \mathrm{IoU}=\dfrac{\left|b\cap b^{gt}\right|}{\left|b\cup b^{gt}\right|}\\ v=\dfrac{4}{\pi^2}\left(\arctan\dfrac{w^{gt}}{h^{gt}}-\arctan\dfrac{w}{h}\right)^2\end{cases},\tag{5}$$

式中：d 为真实框与预测框中心点的欧氏距离；b 为预测框；b^{gt} 为真实框；c 为真实框与预测框最小外接矩形的对角线距离；α 为权重参数；v 为度量宽高比一致性；IoU 为真实框与预测框的面积交并比；w^{gt} 为真实框的宽度；h^{gt} 为真实框的高度；w 为预测框的宽度；h 为预测框的高度。

CIoU 通过参数 v 反映真实框与预测框的宽高比差异性，但难以度量预测框与真实框的宽和高之间的真实差异，且有时会阻碍算法优化相似性。为此，本文提出一种面向野外小目标的显著性指数交并比损失函数 SE-CIoU。SE-CIoU 在 CIoU 的基础上提出两个优化方案：

(1)将 CIoU 中的线性距离惩罚改进为指数

型距离惩罚项 R_{dist}：

$$R_{\text{dist}}=1-\mathrm{e}^{-\gamma\frac{\rho^2}{c^2}}, \tag{6}$$

式中，γ 为显著性聚焦因子，用于调节距离惩罚项的变化幅度。该项随归一化中心距离增大而非线性增大，能够在预测框中心偏离真实框较远时提供更强回归约束，并在二者逐渐接近时减小惩罚变化幅度。

（2）引入遮挡感知权重 w_{occ}，根据当前预测框与真实框的重叠程度动态调节形状惩罚强度，如式（7）所示：

$$w_{\text{occ}}=\text{detach}(\text{IoU}), \tag{7}$$

式中，detach 表示对 IoU 进行梯度截断。当目标发生严重遮挡或训练初期定位不准确时，IoU 较低，w_{occ} 接近于 0，此时形状惩罚项被自适应削弱，能够避免模型过早拟合不完整的遮挡轮廓；当预测框逐渐接近真实框时，IoU 升高，w_{occ} 随之增大，长宽比约束逐步恢复，从而提高边界框与真实目标形态的一致性。

综上，SE-CIoU 损失函数计算公式为：

$$L_{\text{SE-CIoU}}=1-\text{IoU}+R_{\text{dist}}+w_{\text{occ}}\cdot\alpha v. \tag{8}$$

4　实验结果与分析

4.1　数据集

为验证改进 YOLO11n 算法在野外小目标检测任务中的有效性，本文构建野外动物检测数据集（Wildlife Detection Dataset，WDD）用于模型训练和测试。该数据集以公开相机陷阱数据集（Caltech Camera Traps，CCT）[24] 为原始数据来源，经图像筛选、类别整理、标注清洗和格式转换后得到。WDD 数据集中的部分样例如图 4 所示。数据集涵盖不同尺度、遮挡程度、光照条件及与背景融合程度较高的目标。考虑到野外目标受拍摄距离、成像视角和动物活动范围影响，目标尺度呈现连续变化，本文将目标框面积占整幅图像面积小于 5% 的目标视为小目标。



图 4　WDD 数据集中的部分图像

Fig. 4　Some images in the WDD dataset

WDD 数据集共包含 5 464 张图像，按照训练集、验证集和测试集进行划分，其中训练集 4 000 张、验证集 928 张、测试集 536 张。数据使用 LabelImg 软件进行标注，并统一转换为 YOLO 格式。数据集包含 Coyote、Deer、Hog、Rabbit 和 Raccoon 5 类野外动物目标，各类别标注数量如表 1 所示。

表 1　WDD 数据集类别分布

Tab. 1　Category distribution of WDD dataset

类别	标签名称	标注数量
郊狼	Coyote	657
鹿	Deer	2 758
野猪	Hog	2 655
兔	Rabbit	267
浣熊	Raccoon	2 210
合计	—	8 547

由表 1 可知，WDD 数据集中共包含 8 547 个目标标注框，其中 Deer、Hog 和 Raccoon 类目标数量较多，Coyote 和 Rabbit 类目标数量较少，数据集具有一定类别不均衡特征。该分布与野外相机陷阱图像中不同动物出现频次差异较大的实际情况相符。

为进一步分析 WDD 数据集中目标尺度分布情况，本文统计了各类别目标框的宽度和高度，结果如图 5 所示。由图 5 可知，WDD 数据集中目标框宽高分布范围较广，说明野外场景中目标尺度变化明显。其中 Rabbit 类目标框尺寸相对集中且整体偏小，Coyote、Deer、Hog 和 Raccoon 等类别目标尺度差异较大。总体来看，WDD 数据集包含小尺度目标、复杂背景、遮挡干扰和类别不均衡等问题，可用于验证所提算法在复杂野外小目标检测任务中的有效性。

4.2　实验环境与训练设置

实验在统一软硬件环境下进行，以保证不同模型和不同模块之间对比的公平性。实验平台采用 Windows 操作系统，GPU 为 NVIDIA GeForce RTX 4070，深度学习框架为 PyTorch，编程语言为 Python。所有模型均采用相同的数据集划分、输入图像尺寸、训练轮数和优化策略，仅改变网络结构或损失函数设置。主要训练超参数如表 2 所示。

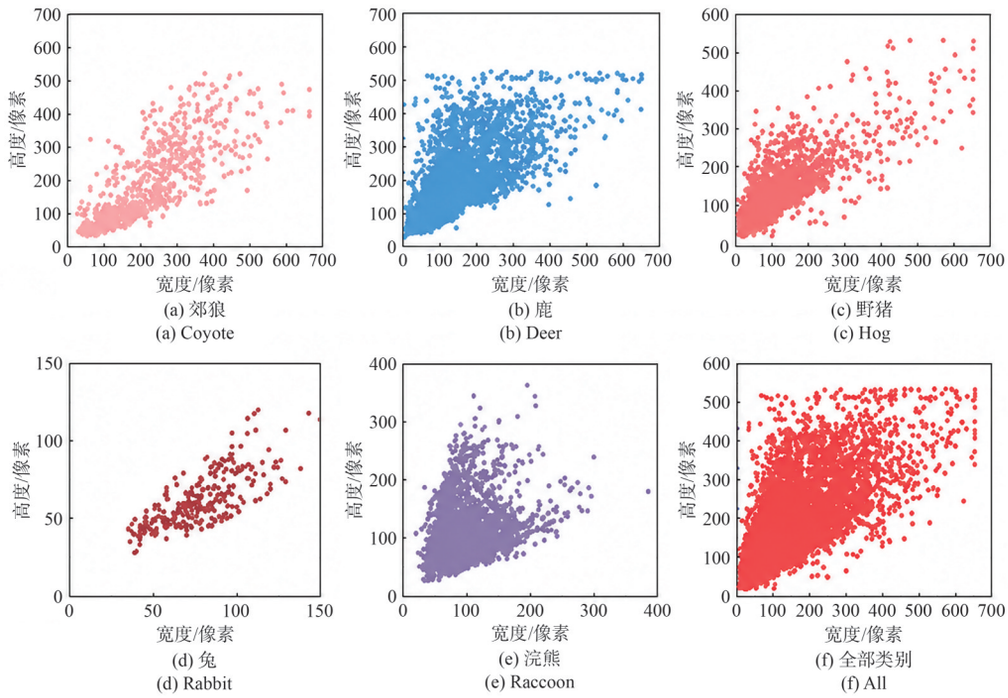


图 5 WDD 数据集中目标框宽高分布

Fig. 5 Width-height distribution of bounding boxes in the WDD dataset

表 2 训练超参数配置

Tab. 2 Training parameter settings

实验参数	参数值
输入图像尺寸	640×640
训练轮数	300
批处理大小	16
优化器	Adam
初始学习率	0.001
最终学习率	0.01
动量参数	0.937
权重衰减	0.000 5

模型训练流程如图 6 所示。首先，将 WDD 数据集划分为训练集、验证集和测试集；然后，加载 YOLO11n 预训练权重并构建改进 YOLO11n 算法模型，在训练集上进行参数学习；训练过程中利用验证集评价模型性能并保存最优权重。若 Precision、Recall 和 mAP@0.5 等指标未达到预期，则调整训练参数继续训练；若模型收敛并达到要求，则在 WDD 测试集和 iWildCam2022-Small 公开子集上进行测试与泛化评估，最终输出检测结果。

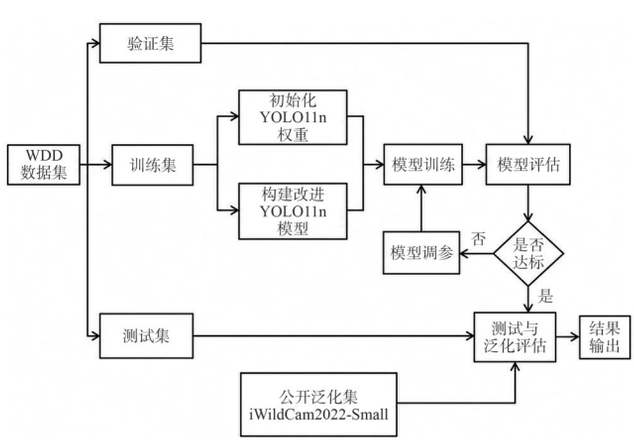


图 6 模型训练流程图

Fig. 6 Flow chart of model training

4.3　评价指标

本文采用精确率(Precision，P)、召回率(Recall，R)、平均精度均值(mean Average Precision，mAP)、每秒处理帧数(Frames Per Second，FPS)、浮点运算量(GFLOPs)和参数量(Params)评价模型性能。其中P、R和mAP@0.5用于评价检测性能，FPS用于评价模型实时性，GFLOPs和Params用于评价模型计算复杂度和规模。各指标计算公式如下：

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\times 100\%\ , \tag{9}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}\times 100\%\ , \tag{10}$$

$$\mathrm{AP}=\int_0^1 P(R)\mathrm{d}R\ , \tag{11}$$

$$\mathrm{mAP}=\frac{1}{N}\sum_{i=1}^{N}\mathrm{AP}_i\times 100\%\ , \tag{12}$$

$$\mathrm{FPS}=\frac{1}{t_{\mathrm{pre}}+t_{\mathrm{inf}}+t_{\mathrm{post}}}\ , \tag{13}$$

式中：TP表示被正确检测的正样本数，FP表示被误检为目标的样本数，FN表示未被正确检测出的真实目标数，N表示目标类别数，mAP@0.5表示IoU阈值为0.5时的平均精度均值。t_{pre}、t_{inf}和t_{post}分别表示单张图像的预处理时间、模型推理时间和后处理时间。

4.4　消融实验

为验证ASU、MultiSEM和SE-CIoU对模型检测性能的影响，本文以YOLO11n为基线模型，在WDD数据集上开展消融实验。实验过程中保持数据集划分、输入图像尺寸、训练轮数和主要超参数一致，仅改变引入的改进模块。实验结果如表3所示，其中“√”表示引入对应模块，“×”表示未引入对应模块。

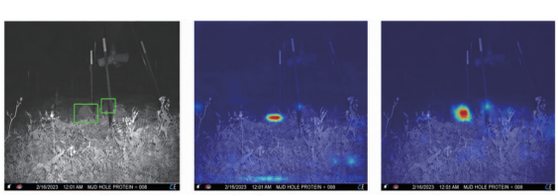
表3　消融实验结果

Tab. 3　Ablation experiment results

YOLO11n	ASU	MultiSEM	SE-CIoU	P/%	R/%	mAP@0.5/%	FPS
√	×	×	×	88.2	84.8	90.7	149.25
√	√	×	×	84.2	88.6	91.4	134.52
√	×	√	×	89.0	84.3	91.8	108.49
√	×	×	√	88.6	87.5	91.9	140.39
√	√	√	×	88.9	88.3	93.2	93.00
√	√	×	√	87.9	89.5	93.7	121.16
√	×	√	√	90.7	89.0	93.9	103.97
√	√	√	√	88.8	91.4	94.0	103.21

由表3可知，单独引入ASU后，R和mAP@0.5分别提高至88.6%和91.4%，说明亚像素上采样能够增强小目标细节表达，提高模型检测精度。单独引入MultiSEM后，P提高至89.0%，mAP@0.5提高至91.8%，说明多尺度显著性增强有助于提高复杂背景下目标区域的判别能力。单独采用SE-CIoU后，R和mAP@0.5分别提高至87.5%和91.9%，表明指数距离约束和动态形状惩罚能够改善小目标边界框回归稳定性。两两组合时，ASU与SE-CIoU组合的mAP@0.5达到93.7%，MultiSEM与SE-CIoU组合的mAP@0.5达到93.9%，说明特征增强与定位约束具有互补作用。3个模块联合后，改进YOLO11n算法的R和mAP@0.5分别达到91.4%和94.0%，较基线模型分别提高6.6%和3.3%，取得最优综合检测性能。

为进一步观察MultiSEM对目标区域响应的影响，提取该模块执行前后的特征热力图，结果如图7所示。未引入MultiSEM时，特征响应在背



(a) 原图　(b) 未引入MultiSEM　(c) 引入MultiSEM
(a) Original image　(b) Before MultiSEM　(c) After MultiSEM

图7　MultiSEM执行前后特征响应热力图对比

Fig. 7　Comparison of feature response heatmaps before and after introducing MultiSEM

景纹理区域存在扩散；引入MultiSEM后，目标区域响应更加集中，背景干扰减弱，表明该模块能够增强模型对野外小目标区域的关注能力。

图8和图9分别给出了YOLO11n与改进YOLO11n在训练过程中的损失曲线和评价指标曲线。由图8可知，两种模型的边界框损失、DFL损失和分类损失均在训练前期快速下降，随后逐渐趋于稳定，说明模型训练过程收敛正常。改进YOLO11n在训练集和验证集上的损失变化趋势一致，未出现明显发散现象，表明所提改进模块未破坏模型的训练稳定性。由图9可知，改进YOLO11n的mAP@0.5、P和R曲线整体高于基线模型，且训练后期波动较小，说明改进YOLO11n算法具有更好的检测精度和收敛稳定性。

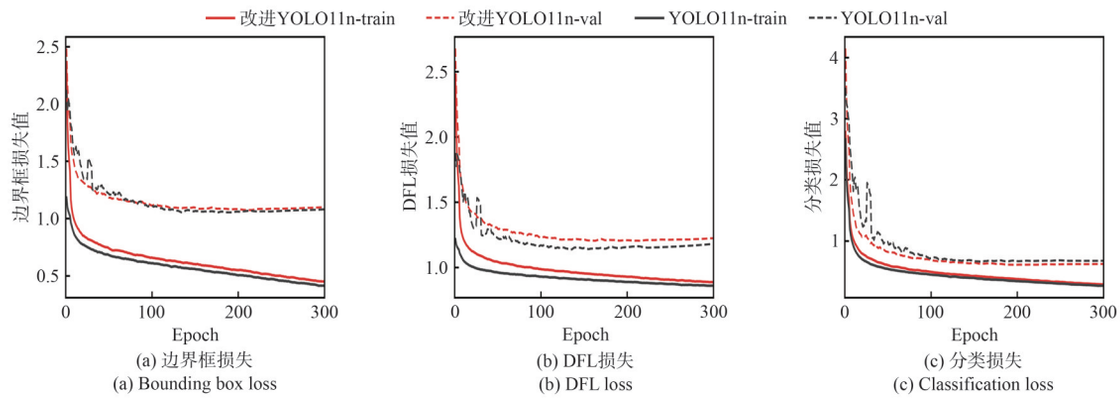


图8 模型损失曲线

Fig. 8 Model loss curves

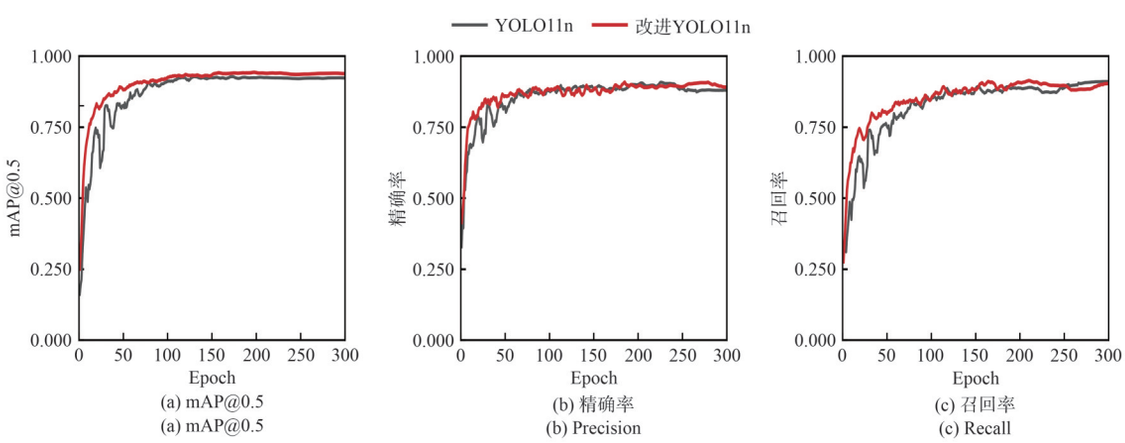


图9 模型评价指标曲线

Fig. 9 Curves of model evaluation metrics

为进一步分析不同类别目标的检测效果，本文对YOLO11n和改进YOLO11n在WDD数据集上的各类别检测结果进行对比，结果如表4所示。改进YOLO11n在5类目标上的AP@0.5均高于YOLO11n基线模型。结合表1中的类别分布可知，Coyote和Rabbit样本数量相对较少，但改进YOLO11n算法仍取得较高检测精度，说明所提方法对少样本和小尺度目标具有一定适应能力。Deer、Hog和Raccoon类别样本数量较多，改进后召回率均有所提升，表明所提改进模块能够有效提升野外小目标的检测性能。

4.5 对比实验

为检验改进YOLO11n算法的综合性能，选取YOLOv8n、YOLO11n、YOLO26n、RT-DETR、Faster R-CNN和RetinaNet进行对比实验。各模型均采用相同的WDD数据集划分、输入图像尺寸、实验环境和训练策略。对比结果如表5所示。

表 4　WDD 数据集各类别检测结果对比

Tab. 4　Detection results of each category on the WDD dataset

种类	YOLO11n			改进 YOLO11n		
	P/%	R/%	AP@0.5/%	P/%	R/%	AP@0.5/%
Coyote	79.8	85.3	87.7	79.5	98.1	95.2
Deer	94.6	79.6	89.1	93.6	84.6	92.0
Hog	89.5	81.1	89.2	87.0	87.0	91.1
Rabbit	90.2	90.9	97.6	100.0	94.3	99.5
Raccoon	86.8	87.3	89.8	83.9	92.8	92.0

表 5　对比实验结果

Tab. 5　Comparison experiment results

模型	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
YOLOv8n	89.2	84.8	91.4	2.9	8.1	200.38
YOLO11n	88.2	84.8	90.7	2.5	6.4	149.25
YOLO26n	84.3	82.6	89.1	2.3	5.2	154.51
RT-DETR	89.0	82.5	87.7	19.2	58.3	73.6
Faster R-CNN	89.7	83.8	92.9	41.4	86.5	31.84
RetinaNet	88.8	81.4	91.7	32.3	97.1	48.92
改进 YOLO11n	88.8	91.4	94.0	2.8	7.1	103.21

由表 5 可知，改进 YOLO11n 算法的 R 和 mAP@0.5 分别为 91.4% 和 94.0%，均高于其他对比模型。其参数量和浮点运算量分别为 2.8M 和 7.1 GFLOPs，仍处于轻量化模型范围。

为进一步分析不同模型在野外复杂场景下的检测性能，本文选取遮挡、低对比度和多目标 3 类复杂场景进行补充测试，结果如表 6 所示。表 6 中的 AP@0.5 表示模型在对应复杂场景子集上的检测结果，平均 AP@0.5 为 3 类场景结果的平均值。由表 6 可知，低对比度场景下各模型的 AP@0.5 整体低于多目标场景，说明当目标与背景在灰度和纹理上差异较弱时，模型更容易受到背景干扰。相较于 YOLO11n 基线模型，改进 YOLO11n 在遮挡、低对比度和多目标场景下的 AP@0.5 均有所提升，表明所提方法能够在一定程度上改善复杂背景下的小目标特征表达和定位稳定性。与 Faster R-CNN 和 RetinaNet 等非 YOLO 类算法相比，改进 YOLO11n 在 3 类复杂场景下也取得了较高检测结果，说明所提算法对复杂野外场景具有较好的适应能力。

表 6　不同模型在复杂场景下的检测结果对比

Tab. 6　Comparison experiment results of different models in complex scenarios

模型	AP@0.5/%			平均 AP@0.5/%
	遮挡	低对比度	多目标	
YOLOv8n	85.8	83.9	90.0	86.6
YOLO11n	85.4	83.5	89.7	86.2
YOLO26n	82.4	80.7	86.5	83.2
RT-DETR	82.0	80.2	85.6	82.6
Faster R-CNN	88.6	85.9	90.7	88.4
RetinaNet	87.4	84.8	89.1	87.1
改进 YOLO11n	90.5	88.4	93.2	90.7

图 10 是不同模型在复杂场景下的检测结果可视化对比。由图 10 可见，在遮挡场景中，目标受到投食器支架、植被等因素影响，部分模型存在漏检或定位不完整现象；在低对比度场景中，目标与周围草丛、阴影区域的灰度和纹理差异较弱，检测难度较高；在多目标场景中，目标数量较多且分布较分散，部分模型容易出现漏检或误检。相比之下，改进 YOLO11n 在上述场景中能够较完整地检出目标，并保持较好的定位效果。

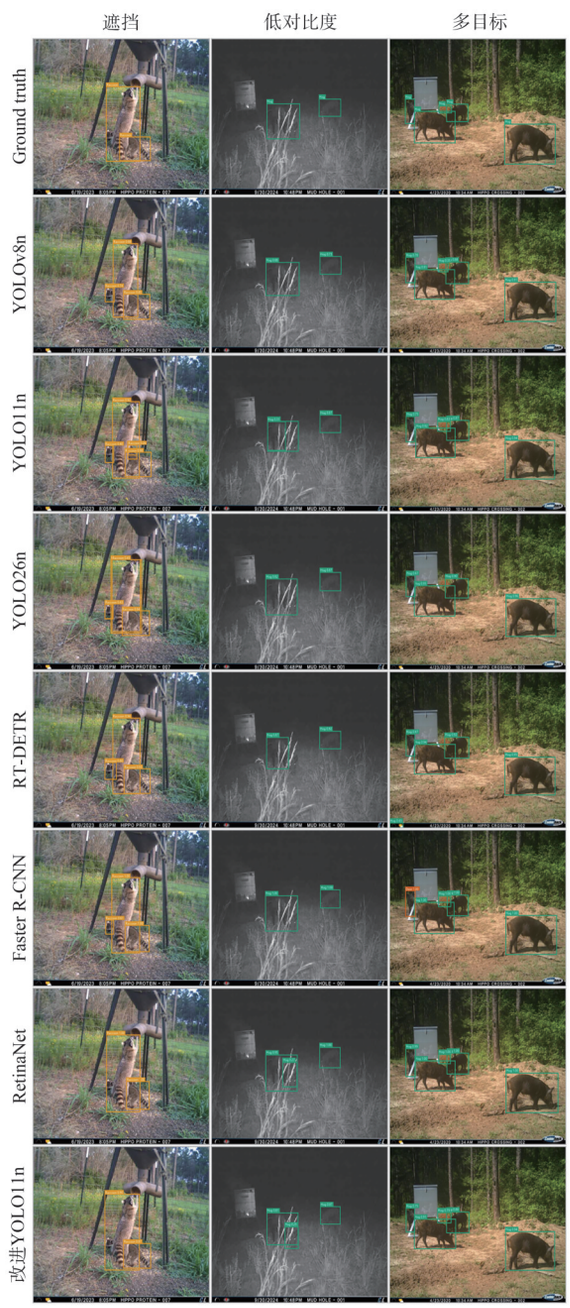


图 10 不同模型在野外复杂场景检测结果可视化对比

Fig. 10 Visualization comparison of detection results for different models in complex field scenarios

4.6 泛化实验

为验证所提算法在公开野外数据集上的泛化能力，本文基于 iWildCam2022 公开数据集[25]构建 iWildCam2022-Small 小目标检测子集。由于 iWildCam2022 原始数据类别数量多、长尾分布明显，且标注格式不能直接用于 YOLO 目标检测训练，本文经筛选与整理后，iWildCam2022-Small 数据集共包含 4 000 张图像。数据集按照 8∶1∶1 的比例划分为训练集、验证集和测试集。同时，为缓解类别不均衡问题，数据构建过程中采用类别配额控制策略，在保证类别语义一致性的前提下，对高频类别进行随机下采样，对低频类别尽量保留，最终形成相对均衡的类别分布。数据集中的部分样例如图 11 所示，各类别图像数量如表 7 所示。在该数据集上，YOLO11n 基准网络算法和所提改进 YOLO11n 算法均采用相同训练策略重新训练，并在测试集上进行性能比较。



图 11 iWildCam2022-Small 数据集中的部分图像

Fig. 11 Some images in iWildCam2022-Small dataset

表 7 iWildCam2022-Small 数据集类别分布

Tab. 7 Category distribution of the iWildCam2022-Small dataset

类别	类别含义	图像数量/张
UngulateA	中小型有蹄类动物	1 010
UngulateB	中大型有蹄类动物	1 009
HerbivoreOthers	其他草食及杂食哺乳动物	972
CarnivoresSmallFauna	食肉、小型哺乳动物及鸟类	1 009

由表 8 可知，在重新训练条件下，所提算法在 iWildCam2022-Small 数据集上的 P、R 和 mAP@0.5 分别为 68.6%、58.3% 和 64.5%，较 YOLO11n 分别提高 1.3%、3.7% 和 3.7%。由图 12 可见，在复杂背景场景下，所提算法能够保持较稳定的检测响应。该结果表明，所提算法不仅在自建 WDD 数据集上有效，在重新构建的公开野外小目标检测子集上也具有一定的泛化能力。

表 8　iWildCam2022-Small 数据集泛化实验结果

Tab. 8　iWildCam2022-Small dataset generalization experiment results

类别	YOLO11n			改进 YOLO11n		
	P/%	R/%	AP@0.5/%	P/%	R/%	AP@0.5/%
All	67.3	54.6	60.8	68.6	58.3	64.5
UngulateA	69.3	58.4	66.1	69.2	61.0	68.8
UngulateB	69.3	66.4	71.8	71.7	68.6	73.8
HerbivoreOthers	69.9	61.4	66.8	71.7	64.3	69.9
CarnivoresSmallFauna	60.8	32.3	38.3	61.9	39.2	45.7
Params/M		2.5			2.8	
GFLOPs		6.3			7.0	

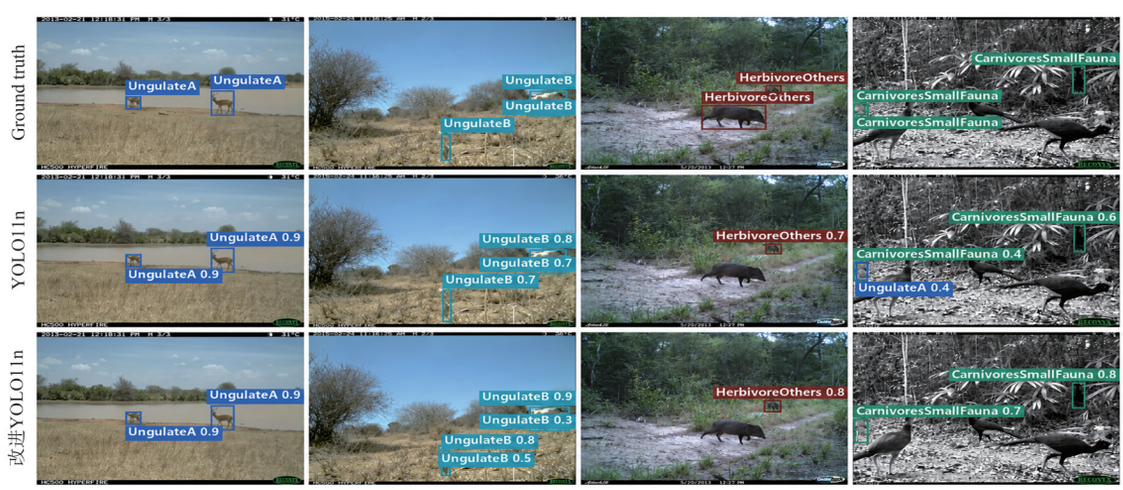


图 12　iWildCam2022-Small 数据集上泛化检测结果可视化对比

Fig. 12　Visual comparison of generalization detection results on iWildCam2022-Small dataset

5　结　　论

基于 YOLO11n 提出一种改进的野外小目标检测算法，从特征上采样、多尺度显著性增强和边界框回归损失函数 3 个方面进行改进。构建 WDD 野外动物检测数据集，对比分析不同目标检测算法的检测性能，开展消融实验研究，验证了各改进模块对提升野外小目标检测性能的有效性：通过引入 ASU 模块，增强小目标边缘和纹理细节特征；通过引入 MultiSEM 模块，增强小目标区域响应和抑制复杂背景干扰；采用 SE-CIoU 损失函数，提高小目标定位精度和稳定性。同时，复杂场景测试结果表明，改进 YOLO11n 在遮挡、低对比度和多目标场景下均具有较好的检测效果，说明所提算法对复杂野外环境具有较强适应能力。在 iWildCam2022-Small 数据集上进行泛化实验，对比了 YOLO11n 和改进 YOLO11n 算法在不同野外场景下的检测性能。结果表明：相较于 YOLO11n 基准算法，本文提出的改进 YOLO11n 算法在野外小目标检测精度和召回率上均有提升，对小尺度、弱纹理、复杂背景和遮挡

条件下的目标具有较好的检测效果,具有一定的复杂场景适应性和泛化能力。后续研究将结合困难样本扩充和知识蒸馏等方法,进一步降低模型复杂度,提高算法在复杂野外环境下的检测性能。

参 考 文 献:

[1] 谢英江. 基于图像融合的野外复杂自然环境人体目标检测算法研究[D]. 北京:军事科学院,2025.
XIE Y J. Research on human object detection algorithm for complex natural environment in the wilderness based on image fusion [D]. Beijing: Academy of Military Science, 2025. (in Chinese)

[2] 王子航,蔡云飞. 基于灰度共生算法的野外环境下高草检测[J]. 计算机与数字工程,2025,53(9):2621-2626,2640.
WANG Z H, CAI Y F. Detection of tall grass in field environment based on gray level co-occurrence algorithm [J]. *Computer and Digital Engineering*, 2025, 53(9): 2621-2626, 2640. (in Chinese)

[3] 江泽涛,施道权,雷晓春,等. 一种基于Night-YOLOX的低照度目标检测方法[J]. 电子学报,2023,51(10):2821-2830.
JIANG Z T, SHI D Q, LEI X C, *et al*. A low-illumination object detection method based on Night-YOLOX [J]. *Acta Electronica Sinica*, 2023, 51(10): 2821-2830. (in Chinese)

[4] MUZAMMUL M, LI X. Comprehensive review of deep learning-based tiny object detection: challenges, strategies, and future directions [J]. *Knowledge and Information Systems*, 2025, 67(5): 3825-3913.

[5] ALDUBAIKHI A, PATEL S. Advancements in small-object detection (2023—2025): approaches, datasets, benchmarks, applications, and practical guidance [J]. *Applied Sciences*, 2025, 15(22): 11882.

[6] 张志豪,厉小润,陈淑涵. 基于改进YOLO11的无人机航拍图像小目标检测算法[J]. 液晶与显示,2025,40(6):915-930.
ZHANG Z H, LI X R, CHEN S H. Small object detection algorithm in UAV aerial images based on improved YOLO11 [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(6): 915-930. (in Chinese)

[7] ZHANG X, DU B Y, JIA Y X, *et al*. MSD-YOLO11n: an improved small target detection model for high precision UAV aerial imagery [J]. *Journal of King Saud University Computer and Information Sciences*, 2025, 37(5): 100.

[8] 曾子豪,刘鹏,邓文娟,等. 基于改进YOLO11n的导电粒子目标检测算法[J]. 液晶与显示,2025,40(12):1853-1867.
ZENG Z H, LIU P, DENG W J, *et al*. Conductive particle target detection algorithm based on improved YOLO11n [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(12): 1853-1867. (in Chinese)

[9] 杨晓敏,杨军. 超轻量化SAR影像小目标检测网络[J]. 光学 精密工程,2025,33(10):1672-1690.
YANG X M, YANG J. Ultra lightweight SAR image small object detection network [J]. *Optics and Precision Engineering*, 2025, 33(10): 1672-1690. (in Chinese)

[10] YAN H M, KONG X B, SHIMADA T, *et al*. TOE-YOLO: accurate and efficient detection of tiny objects in UAV imagery [J]. *Journal of Real-Time Image Processing*, 2025, 22(6): 194.

[11] LIU W Z, LU H, FU H T, *et al*. Learning to upsample by learning to sample [C]. 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris: IEEE, 2023: 6004-6014.

[12] SHI W Z, CABALLERO J, HUSZÁR F, *et al*. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 1874-1883.

[13] MA Z B, DONG Y Q, XIA Y, *et al*. Wildlife real-time detection in complex forest scenes based on YOLOv5s deep learning network [J]. *Remote Sensing*, 2024, 16(8): 1350.

[14] YANG W H, LIU T Y, JIANG P, *et al*. A forest wildlife detection algorithm based on improved YOLOv5s [J]. *Animals*, 2023, 13(19): 3134.

[15] 卢汉,崔博伦,万华洋,等. 半监督式野生动物夜间目标端到端检测[J]. 光学 精密工程,2025,33(5):789-801.
LU H, CUI B L, WAN H Y, *et al*. End-to-end recognition of nighttime wildlife based on semi-supervised learning [J]. *Optics and Precision Engineering*, 2025, 33(5): 789-801. (in Chinese)

[16] DU Z J, WU D S, WEN Q Q, *et al*. An improved lightweight model for protected wildlife detection in camera trap

images [J]. *Sensors*, 2025, 25(23): 7331.

[17] ZHU Y T, ZHAO Y X, HE Y X, *et al*. YOLO-WildASM: an object detection algorithm for protected wildlife [J]. *Animals*, 2025, 15(18): 2699.

[18] LIU C, WANG P, GONG Y P, *et al*. YOLO-WL: a lightweight and efficient framework for UAV-based wildlife detection [J]. *Sensors*, 2026, 26(3): 790.

[19] CHEN L L, LI G, ZHANG S K, *et al*. YOLO-SAG: an improved wildlife object detection algorithm based on YOLOv8n [J]. *Ecological Informatics*, 2024, 83: 102791.

[20] WOO S, PARK J, LEE J Y, *et al*. CBAM: convolutional block attention module [C]. 15th European Conference on Computer Vision. Munich: Springer, 2018: 3-19.

[21] ZHENG Z H, WANG P, REN D W, *et al*. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. *IEEE Transactions on Cybernetics*, 2022, 52(8): 8574-8586.

[22] SAPKOTA R, FLORES-CALERO M, QURESHI R, *et al*. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series [J]. *Artificial Intelligence Review*, 2025, 58(9): 274.

[23] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements [J/OL]. *arXiv*, 2024: 2410.17725.

[24] BEERY S, VAN HORN G, PERONA P. Recognition in terra incognita [C]. 15th European Conference on Computer Vision. Munich: Springer, 2018: 472-489.

[25] KOH P W, SAGAWA S, MARKLUND H, *et al*. WILDS: a benchmark of in-the-wild distribution shifts [C]. *Proceedings of the* 38*th International Conference on Machine Learning*. Online: PMLR, 2021: 5637-5664.

作者简介：



王　岑，女，硕士研究生，2021年于西南科技大学获得学士学位，主要从事深度学习、目标检测方面的研究。E-mail：572438182@qq.com



王彩霞，女，博士，副教授，2009年于长春理工大学获得博士学位，主要从事光电检测、机器视觉与嵌入式图像处理等方面的研究。E-mail：wcxhao@sina.com