DOI: 10.12398/j.issn.2096-7217.2026.01.007

基于改进RT-DETR的草原鼠洞智能识别与检测模型设计与试验

董振伟, 付学良, 李宏慧, 潘新, 徐喆, 罗小玲*

(内蒙古农业大学计算机与信息工程学院,内蒙古呼和浩特,010018)

摘要:近年来,草原鼠洞的迅速增殖与扩散已成为导致草原退化和生态失衡的重要因素之一。鼠洞的密集分布破坏了草皮结构,削弱土壤稳定性,降低植被覆盖度与牧草产能,严重威胁草原生态安全与畜牧业的可持续发展。针对现有检测算法在小目标识别能力不足、特征提取不充分及复杂背景干扰显著等问题,研究提出一种基于实时检测(real-time detection transformer, RT-DETR)的改进模型——RT-DETR-ECG(efficient global perception and feature selection)。该模型首先引入高效视觉Mamba网络(efficient vision Mamba,EfficientVIM)模块,通过状态空间建模与深浅层特征交互,实现全局与局部信息的高效融合,从而增强小目标特征感知与跨尺度建模能力;其次,构建卷积门控线性单元(convolutional gated linear unit,CGLU)模块,利用卷积与门控机制的协同作用动态抑制复杂背景噪声,有效突出鼠洞边缘与纹理特征;最后,设计小目标加权损失函数(CGLU-Loss),在回归阶段引入中心偏移权重与局部重叠增益项,以强化小尺度样本的梯度贡献并提升定位精度。试验基于2 397张跨季节无人机(UAV)草原鼠洞图像数据集开展,结果表明:RT-DETR-ECG在检测精度、实时性与模型轻量化方面均优于现有算法。在*IoU*阈值为0.5时,模型的平均精度(*mAP*@0.5)达96.3%,较原始RT-DETR提升5.1%,检测速度达91.8帧/s,计算负载降低53.7%。该模型实现了小目标检测的精度提升与实时性能的平衡,为草原鼠洞智能识别与生态监测提供了一种高效可靠的新方法。

关键词:RT-DETR-ECG;目标检测;深度学习;图像识别;无人机影像;啮齿动物洞穴

中图分类号:S24;TP391.4 **文献标识码**:A **文章编号**:2096-7217(2026)01-0063-12

董振伟, 付学良, 李宏慧, 潘新, 徐喆, 罗小玲. 基于改进RT-DETR的草原鼠洞智能识别与检测模型设计与试验[J]. 智能化农业装备学报(中英文), 2026, 7(1): 63-74

DONG Zhenwei, FU Xueliang, LI Honghui, PAN Xin, XU Zhe, LUO Xiaoling. Design and experiment of an intelligent recognition and detection model for grassland rodent burrows based on improved RT-DETR [J]. Journal of Intelligent Agricultural Mechanization, 2026, 7(1): 63-74

0 引言

草原是陆地生态系统的重要组成部分,在维系生物多样性、防止土地沙化和支撑畜牧业可持续发展方面具有重要作用[1]。近年来,草原啮齿动物活动频繁,鼠洞数量急剧增加,成为导致草原退化与生态失衡的重要因素[2-3]。我国草原面积约392 000 khm^2,占国土面积的41%,主要分布于内蒙古、新疆、青海和西藏,是我国畜牧业与生态安全的重要基础。然而,部分退化草原区域鼠洞密度已超过300~500个/hm^2,重度鼠害区甚至达到800个/hm^{2}以上,可导致牧草产量下降20%~50%,严重影响草原生态功能与生产能力。因此,开展草原鼠洞的高精度自动检测与动态监测,对掌握鼠害分布、制定防控措施和推进生态修复具有重要意义[4]。

传统鼠洞监测多依赖人工巡查或浅层特征的机器学习方法[5],如支持向量机(SVM)和Haar级联分类器等。此类方法特征表达能力有限,易受光照、植被和地形变化影响,检测精度和稳定性较低。随着深度学习的发展[6],基于卷积神经网络(CNN)的检测算法(如Faster R-CNN与YOLO系列)显著提升了目标识别性能[7],但在草原复杂场景下的小目标检测仍存在不足。鼠洞尺寸小、边缘模糊、背景复杂,易导致特征丢失与误检;此外,网络结构庞大、计算量高,也限制了其在无

收稿日期:2025-11-02 修回日期:2025-12-15

基金项目:高层次及优秀博士人才引进科研启动项目(NDYB2022-60);内蒙古自治区自然科学基金项目(2023LHMS06020);内蒙古自然基金重点项目(2025ZD012);内蒙古自治区揭榜挂帅项目(2025KJTW0026);内蒙古自治区重点项目(2025KYPT0076)

第一作者:董振伟,男,1998年生,河南商丘人,硕士研究生;研究方向为网络与信息安全。E-mail: dongzw@emails.imau.edu.cn

*通信作者:罗小玲,女,1972年生,内蒙古哲盟科左中旗人,蒙古族,博士,教授;研究方向为图像处理与模式识别。E-mail: luoxl@imau.edu.cn

人机和边缘设备上的实时应用[8]。

针对小目标检测问题,近年来学者提出了多种改进策略。王世勤[9]基于Faster R-CNN开展草原啮齿动物洞穴识别研究,引入特征金字塔结构以增强多尺度感知[10],但特征融合效率有限,浅层与深层信息交互不足,难以保留小目标细节。JIANG等[11]提出多尺度特征融合网络(MFFSODNet),在无人机影像检测中取得一定效果,但模型计算复杂度较高。ALI等[12]和JEGHAM等[13]将注意力机制融入YOLO系列以提升特征选择能力,但在复杂背景下仍存在误检与特征漂移。Transformer架构模型(如Deformable DETR与DETR)具备全局建模优势,但对小尺度目标边缘特征提取仍不充分,且解码过程计算量较大。由此可见,当前小目标检测研究仍面临3大难题:

1)细粒度特征捕获不足:原始RT-DETR采用的ResNet-50主干网络在深层下采样过程中丢失了小目标的边缘与纹理信息,导致对鼠洞等小尺度目标的检测不稳定。

2)跨季节环境适应性弱:草原生态系统受季节影响显著,光照、植被与土壤颜色变化大,模型容易过拟合单一时相特征,跨季节检测性能下降。

3)复杂背景下的误检率较高:草原地表存在阴影、石块及草根等高相似纹理,原模型中仅使用线性前馈网络(FFN)进行特征判别,难以有效区分目标与背景。

针对上述问题,提出一种基于实时检测框架RT-DETR的改进模型——RT-DETR-ECG(efficient global perception and feature selection)。首先,引入高效视觉Mamba网络(efficient vision Mamba,EfficientVIM)模块,实现多尺度特征的高效全局建模;其次,设计卷积门控线性单元(convolutional gated linear unit,CGLU)模块,利用卷积—门控机制抑制复杂背景干扰[14];最后,提出小目标加权损失函数(CGLU-Loss),优化边界框回归过程。

1 材料和方法

1.1 草原鼠洞数据集构建与特征分析

以草原鼠洞为主要识别对象(以下统称鼠洞),影像数据于2024年6月—2025年5月在内蒙古典型草原地区采集,采集区域包括希拉穆仁、辉腾锡勒和格根塔拉草原。数据获取采用DJI Mavic 2 Pro无人机平台,飞行高度约2 m,镜头向下倾斜角约20°,相机分辨率为4 000像素×3 000像素,影像格式为JPEG,如图1所示。分别在春、夏、秋、冬4个时段进行多次飞行采集,共获取有效样本图像2 397张,如图2所示。采集过程中,保持飞行速度约1 m/s,航线间距30 m,以确保地表纹理连续性与鼠洞细节清晰可辨。所采集数据涵盖不同植被密度、地表结构及光照条件,为模型训练与跨季节性能评估提供了多样化的样本支撑。



图1 草原鼠洞影像采集场景与无人机平台

Figure 1 Grassland rodent burrow image acquisition scene and UAV platform

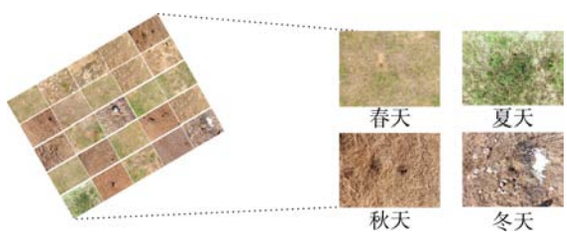


图2 不同季节草原鼠洞样本图像

Figure 2 Sample images of grassland rodent burrows in different seasons

数据统计表明,标准化面积小于0.02的鼠洞被定义为小目标,占比95.93%;86.72%的图像仅包含一个鼠洞,9.45%包含两个,3.83%包含3个及以上,如图3所示。该跨季节、多尺度数据集为模型提供了丰富的特征样本支撑,有助于提升模型在多时空条件下的泛化性能与稳定性。

所构建数据集中,四季影像比例约为春季25.6%、夏季26.8%、秋季28.4%、冬季19.2%,保证了样本的时相均衡性与特征多样性。

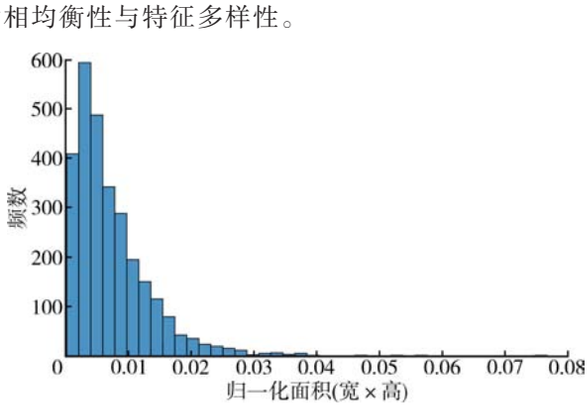


图3 数据集中标准化鼠洞面积的分布

Figure 3 Distribution of standardized rodent burrow areas in the dataset

1.2 数据的增强

草原无人机航拍影像中,鼠洞区域通常存在亮度不足、局部对比度弱、边缘纹理模糊以及易与草丛背景混淆等问题。这些弱纹理小目标在深度网络浅层特征阶段易被背景淹没,从而影响改进RT-DETR模型检

测效果。

为提高模型的泛化能力与跨季节适应性，本文对草原鼠洞数据集进行了系统性数据增强[15]，主要包括几何翻转和亮度－对比度增强两类方法。首先，对原始影像进行随机水平和垂直翻转，引入视角变化与空间对称性，扩展样本分布范围，如图 4(c)、图 4(d)所示。基于前述弱亮度与弱纹理特性，本文在亮度对比度增强中选用了伽马校正(Gamma correction)和限制对比度自适应直方图均衡化(CLAHE)。伽马校正能够对暗部区域进行亮度非线性拉伸，使鼠洞凹陷与周围土壤形成更明显的亮度差；CLAHE 则通过局部对比度提升突出鼠洞的细微纹理变化与洞口边缘结构，提高弱纹理小目标的可分辨性。增强后样本如图 4(e)、图 4(f)所示，分别对应伽马校正与 CLAHE 处理后的结果，验证了该方法在提升鼠洞显著性与场景适应性方面的有效性。

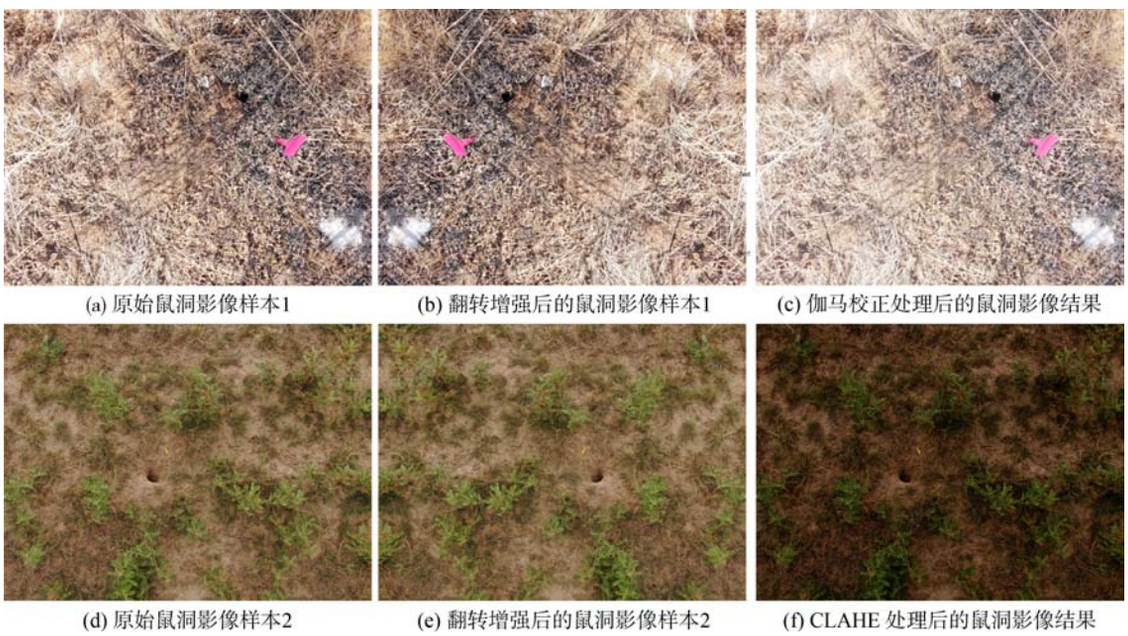


图 4　鼠洞影像数据增强示意图

Figure 4　Schematic diagram of image data augmentation for mouse hole

2　RT-DETR-ECG 模型的改进

如图 5 所示，本研究在 RT-DETR 框架基础上提出了一种改进的草原鼠洞检测模型 RT-DETR-ECG，对整体结构进行了针对性优化。模型主要由特征提取主干(C2f-EfficientVIM backbone)、特征交互融合层(AIFI module)[16]、检测头(RT-DETR decoder)以及小目标加权回归模块(CGLU-Loss)组成。旨在提升复杂草原环境下鼠洞目标的检测精度与跨季节鲁棒性。在特征提取阶段，引入改进的 C2f-EfficientVIM 主干网络，通过深度可分离卷积与并行状态建模机制实现多尺度特征提取与全局－局部信息融合[17]，同时，在主干网络中嵌入 CGLU 模块，用于增强小目标的细粒度特征表达能力。CGLU 由特征变换分支与门控分支构成，其中，特征分支负责生成候选特征映射，门控分支通过卷积操作与非线性激活函数生成通道权重，对不同通道特征进行自适应调制。两路特征在通道维度上进行逐元素融合，从而在突出鼠洞边缘与局部纹理特征的同时，有效抑制植被、阴影及碎石等复杂背景干扰。在特征融合与检测阶段，采用轻量化 AIFI 模块对不同层级特征进行动态加权融合，并将融合后的特征输入 RT-DETR 解码器，通过 Top-K Query Selection 机制提升目标匹配效率与检测稳定性。在回归阶段，引入基于 CGLU-Loss 的加权回归策略，以提高小目标定位精度。整体而言，该模型在保持轻量化结构的同时，实现了检测精度与计算效率的有效平衡。

2.1　C2f-EfficientVIM 主干网络设计

为解决基准 RT-DETR 模型在草原复杂背景下小目标特征提取不足的问题，本文在主干网络中引入 C2f-EfficientVIM 结构，以提升多尺度表征与全局建模能力。传统 ResNet-50 在连续下采样中易造成细粒度特征丢失，尤其在植被密集、鼠洞尺度较小的区域。为此，设计基于跨阶段局部特征融合(cross-stage partial feature fusion，C2f)的改进主干，实现浅层与深层特征的高效融合，增强信息流动与梯度稳定性；同时，引入 EfficientVIM 模块，通过二维状态空间扫描(SS2D)机制[18]，并结合轻量卷积构成混合空间建模(HSM-SS2D)结构，实现了全局信息聚合与计算复杂度的平

衡。兼顾全局依赖建模与计算效率。部分C2f-EfficientVIM块中嵌入轻量卷积门控线性单元,以动态抑制背景噪声、强化目标纹理特征。该结构通过“浅层局部增强—深层全局融合”的协同机制,实现了对草原鼠洞等小目标的高效稳健感知。其整体网络结构如图5所示。

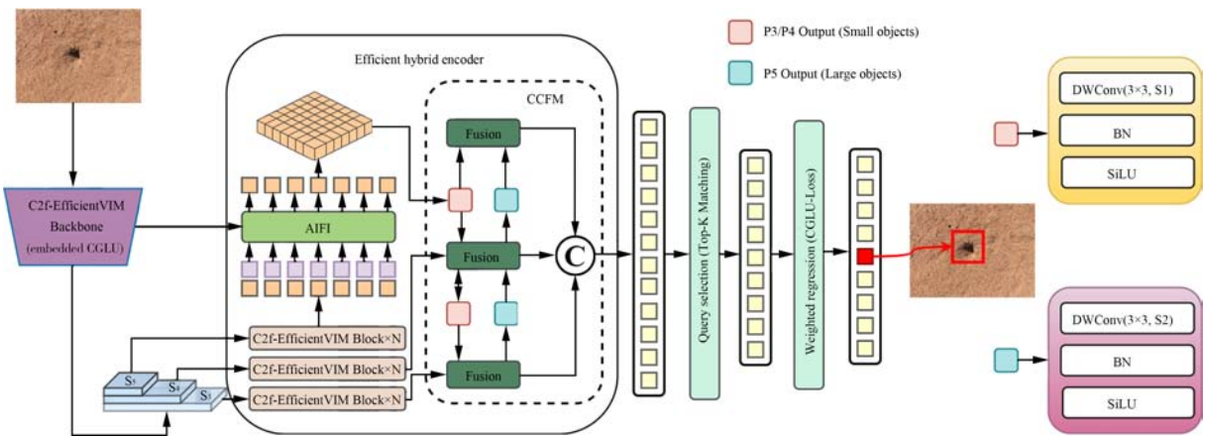


图5 RT-DETR-ECG模型总体架构示意图

Figure 5 Schematic diagram of the overall architecture of the RT-DETR-ECG model

主干网络由3阶段C2f-EfficientVIM模块组成,各阶段通过逐级补丁融合实现空间分辨率压缩与通道维度扩展,从而构建层次化特征表示。每个EfficientVIM模块内部结合轻量级卷积操作与HSM-SS2D单元,在保持计算效率的同时增强对全局上下文与长程依赖关系的建模能力。

多阶段输出特征在经过多尺度特征融合(multi-scale fusion, MSF)后进行统一表达[19],有效整合浅层细节信息与深层语义信息,为后续小目标检测提供稳定且判别性更强的特征支持。其整体结构如图6所示。

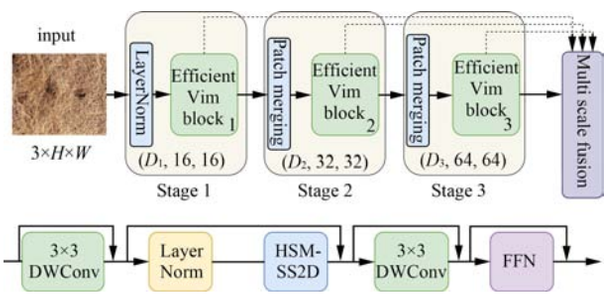


图6 EfficientVIM主干网络的整体结构

Figure 6 Overall structure of the EfficientVIM backbone network

HSM-SSD层通过混合空间建模与选择性扫描机制,实现局部细节建模与长程依赖建模的协同融合。其整体结构由多分支特征映射、局部卷积建模以及状态空间交互模块构成,不同分支特征在后续阶段进行加权融合,形成兼顾局部纹理与全局语义的信息表达。该层的具体结构与特征交互流程如图7所示。

尽管C2f-EfficientVIM在YOLOv8系列中已表现出良好的全局建模与多尺度语义融合能力,但其主要侧重于跨层语义交互,对小目标的边缘与细粒度纹理增强仍存在一定局限。

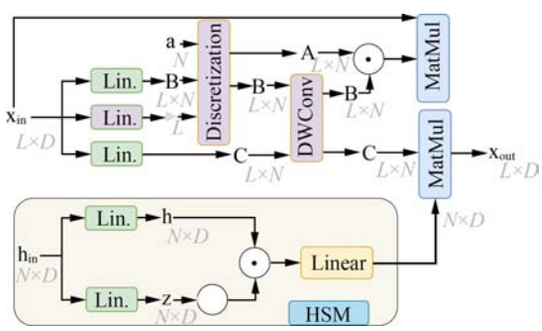


图7 EfficientVIM模块中HSM-SSD层的结构示意图

Figure 7 Schematic diagram of the HSM-SSD layer in the EfficientVIM module

为此,本文引入CGLU模块及相应的CGLU-Loss,以增强局部几何特征建模能力并提升小目标定位敏感性。CGLU与C2f-EfficientVIM在功能上形成互补:前者侧重细粒度特征增强,后者提供稳定的全局与多尺度语义支撑,从而在检测精度与推理效率之间实现较为平衡的整体性能。

2.2 基于卷积门控线性单元的特征增强

在草原鼠洞检测任务中,由于目标尺度较小、边界模糊且与周围土壤和植被纹理相似,传统特征提取模块在复杂背景下往往难以有效区分目标与背景,小目标检测性能受限。为此,本文在主干网络的特征增强阶段引入卷积门控线性单元(CGLU),以提升模型对小目标细粒度特征的感知与判断能力。

通过对特征响应进行自适应调制,CGLU能够在突出鼠洞边缘及局部纹理信息的同时,有效抑制植被、阴影及碎石等复杂背景干扰,从而增强小目标在高相似背景下的可分性。

消融试验结果表明，在主干网络中引入 CGLU 模块后，模型在复杂草原场景下的小目标检测性能得到明显提升。与传统前馈网络（FFN）相比[20]，引入 CGLU 后模型表现出更好的稳定性与鲁棒性，验证了该模块在特征增强方面的有效性。

进一步分析表明，CGLU 与 C2f-EfficientVIM 在功能上形成互补：前者侧重于细粒度局部特征增强，后者提供稳定的全局与多尺度语义建模能力，从而在检测精度与推理效率之间实现较为平衡的整体性能。其具体结构示意如图 8 所示。

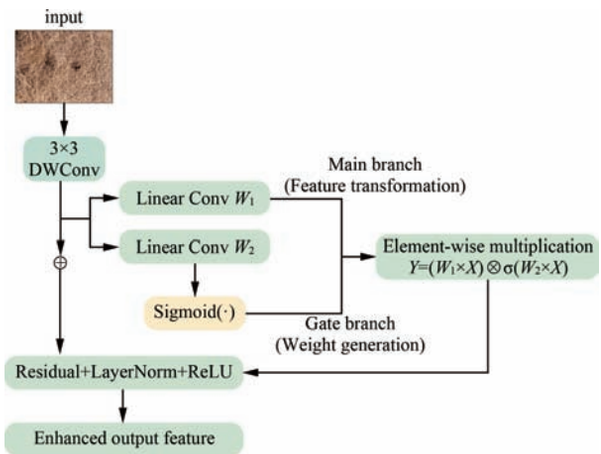


图 8　卷积门控线性单元模块结构示意图

Figure 8　Schematic diagram of the convolutional gated linear unit module structure

2.3　小目标定位的加权损失函数（CGLU-Loss）

由于鼠洞目标开口区域尺度较小、与背景对比度低且边缘纹理模糊，传统检测模型在回归阶段对小目标的定位响应不敏感，易出现预测框偏移与定位不稳定的问题。为此，本文在 RT-DETR 原有回归分支的基础上设计了一种小目标定位的加权损失函数（CGLU-Loss），通过引入针对小尺度目标的加权监督机制，提升预测框中心定位的稳定性与精度。

在回归优化过程中，CGLU-Loss 结合中心偏移加权与局部重叠增益两类策略，对小目标样本实施针对性强化监督。其中，中心偏移加权用于刻画预测框中心与真实目标中心之间的匹配程度，使高质量定位样本在回归损失中获得更大权重；局部重叠增益进一步强调预测框与真实框在局部区域内的空间重叠关系，增强模型对鼠洞边缘区域及局部结构的定位敏感性。两种加权机制相互补充，在避免不可靠样本主导回归过程的同时，提高了小目标定位的鲁棒性。从优化视角看，CGLU-Loss 通过对高质量定位样本梯度的自适应放大，引导模型在回归阶段更加关注小目标的中心一致性与局部结构匹配，从而缓解小目标场景中回归不稳定的问题，并与 CGLU 模块的特征增强机制形成有效协同。

3　试验结果与分析

3.1　试验环境与参数设置

本研究的试验平台基于 Windows 11 64 位操作系统，采用 Intel® Core™ i5-12600KF 处理器（3.7 GHz）与 NVIDIA GeForce RTX 3080 显卡，显存容量为 10 GB。GPU 加速库版本为 CUDA 12.1 与 cuDNN 7.6.5，深度学习框架采用 PyTorch 2.2.2，Python 版本为 3.10.14。训练过程中，选用随机梯度下降法（stochastic gradient descent，SGD）作为优化算法，以加快模型收敛并保证训练稳定性。为防止初始阶段的梯度振荡并提升模型深层权重更新的稳定性，采用 warmup 预热策略，在前 3 轮中将学习率设为 0.000 1 进行平滑启动，随后恢复至初始学习率 0.01。学习率调整方式采用余弦退火衰减（cosine annealing）策略，最终学习率降至 0.000 1。动量因子（momentum）设置为 0.9，权重衰减系数为 0.000 5，最大迭代轮数为 200。输入影像分辨率统一为 640 像素×640 像素，批次大小（batch size）设置为 16，训练时所有模型均从零初始化权重，不采用预训练模型或迁移学习策略。上述设置在确保训练稳定性的同时，验证了 RT-DETR-ECG 模型在草原鼠洞检测任务中高效、稳健的训练性能。

3.2　评价指标

为全面评估改进后 RT-DETR-ECG 模型在草原鼠洞检测任务中的综合性能，本文选取目标检测领域常用的评价指标进行分析，包括衡量模型对鼠洞识别的准确性的精确率（precision，P）、衡量模型对鼠洞识别的完整性召回率（recall，R）、反应总体检测性能的平均精度均值（mean average precision，mAP）、模型参数量（parameters，$Params$）、计算速度（floating point operations）、检测速度（v）及权重文件大小（model size）。具体定义如式（1）—式（6）所示。

$$P=\frac{TP}{TP+FP} \tag{1}$$

$$R=\frac{TP}{TP+FN} \tag{2}$$

$$AP=\int_0^1 P(r)\mathrm{d}r \tag{3}$$

$$mAP=\frac{1}{n}\sum_{i\in\{0,\,1,\,\cdots,\,n\}}P(r) \tag{4}$$

$$v=\frac{1}{T_{\mathrm{avg}}} \tag{5}$$

$$Params_{\mathrm{Conv}}=K_{\mathrm{h}}\times K_{\mathrm{w}}\times C_{\mathrm{in}}\times C_{\mathrm{out}}+C_{\mathrm{out}} \tag{6}$$

式中：TP——被正确检测为鼠洞的正样本数量；

FP——被误检为鼠洞的负样本数量;

FN——未被检测出的鼠洞样本数量;

AP——平均精度。

T_{avg}——单张图像的平均处理时间,s;

$Params_{Conv}$——卷积层的参数量;

K_h——卷积核的高度;

K_w——卷积核的宽度;

C_{in}——输入通道数;

C_{out}——输出通道数。

综合使用上述指标,可对RT-DETR-ECG的检测精度、运算效率及部署性能进行全面对比与验证。

3.3 不同模型间的对比

近年来,在鼠洞小目标或地面洞穴检测方向也出现了一些专门研究,如SUN等[21]利用GF-2卫星影像并基于Faster-RCNN实现布氏田鼠鼠洞识别的研究,ZHAO等[22]提出基于改进YOLOv11的无人机高原鼠兔洞穴自动检测方法,以及KARIMINEJAD等[23]基于多种深度学习模型从UAV影像中实现地表塌陷洞穴检测的工作。

这些方法与本文任务具有较高相关性,但大多未公开训练代码或缺乏统一的评价指标体系,难以在本文数据集上进行严格复现试验。

考虑到试验的公平性与可重复性,本文仍选取当前在小目标检测任务中最具代表性且指标体系一致的深度学习模型(如Faster-RCNN、YOLOv5s、YOLOv11n及RT-DETR系列)作为统一基线模型,并结合文献中相关方法的设计思路进行分析,以增强对比试验的领域背景与说服力。

为验证所提RT-DETR-ECG模型的检测性能与环境适应性,将其与经典两阶段检测模型Faster-RCNN[24]、轻量级一阶段模型YOLOv5s[25-27]、YOLOv11n[28]以及Transformer架构的RT-DETR-L[29]、RT-DETR-r50进行对比试验。

结果表明,所提模型在精确率(P)、召回率(R)和mAP@0.5上分别达到94.2%、92.2%和96.3%,均优于其他模型;其参数量和计算量分别为14.5 M和47.9 GFLOPs,在保持较高精度的同时兼顾了模型轻量化。基于春、夏、秋、冬四季草原鼠洞数据集的测试结果显示,Faster-RCNN与YOLOv5s对部分遮挡鼠洞存在明显漏检,YOLOv11n与RT-DETR-L在岩石阴影区域易产生误判。相比之下,RT-DETR-ECG通过改进特征融合与注意力机制,能有效区分鼠洞与相似背景纹理,表现出更强的鲁棒性与泛化能力。各模型的定量结果见表1,典型检测效果如图9所示[30]。

表1 不同模型试验结果

Table 1 Experimental results of different models

模型	精确率/%	召回率/%	mAP@0.5/%	检测速度/(帧·s^{-1})	参数量/M	计算量/GFLOPs
Fsater-RCNN	78.1	75.7	76.7	12.5	135.6	15.4
YOLOv5s	86.7	84.4	89.1	81.5	7.3	15.9
YOLOV11n	79.1	70.7	75.8	98	19	27
RT-DETR-L	91.2	89.5	91.2	75.8	32	103.4
RT-DETR-r50	90.3	87	90.6	84	29	80
RT-DETR-ECG	94.2	92.2	96.3	91.8	14.5	47.9

3.4 消融试验

为验证改进后的RT-DETR-CGLU模型在鼠洞检测任务中的性能提升效果,本文基于草原鼠洞专属数据集进行了系统的消融试验。试验以原始RT-DETR模型作为基准,采用“逐步添加目标模块”的递进式设计思路,共构建6组试验样本,每次仅新增一个改进模块,从而精准评估各模块对检测精度(mAP@0.5)、检测速度及模型复杂度(参数量)的独立影响。所有试验均在相同数据集与超参数设置下完成,确保结果的可比性与公正性。

从表2的消融试验结果可以看出,在以RT-DETR为基础的各类模型组合中,改进后的RT-DETR-CGLU模型整体表现最为优异。单独引入C2f结构后,mAP@0.5较基准模型有所提升,表明该模块在多尺度特征融合方面具有一定的优势,但同时检测速度略有下降;进一步加入EfficientVIM模块后,mAP@0.5和检测速度均较基准模型有所提高,显示其在全局特征建模与信息聚合上的有效性,但参数量仍然较大。

值得注意的是,当在此基础上引入CGLU模块后,模型性能提升最为显著:mAP@0.5由91.2%提升至94.8%,提升幅度达3.6%,检测速度提升至86.7帧/s,且参数量显著减少至82 M。其门控机制能够在特征提取阶段动态抑制植被、碎石等复杂背景干扰,提升目标特征的纯度与显著性,从而有效降低误检率,验证了

CGLU 模块在强化模型鲁棒性与小目标识别方面的优势。

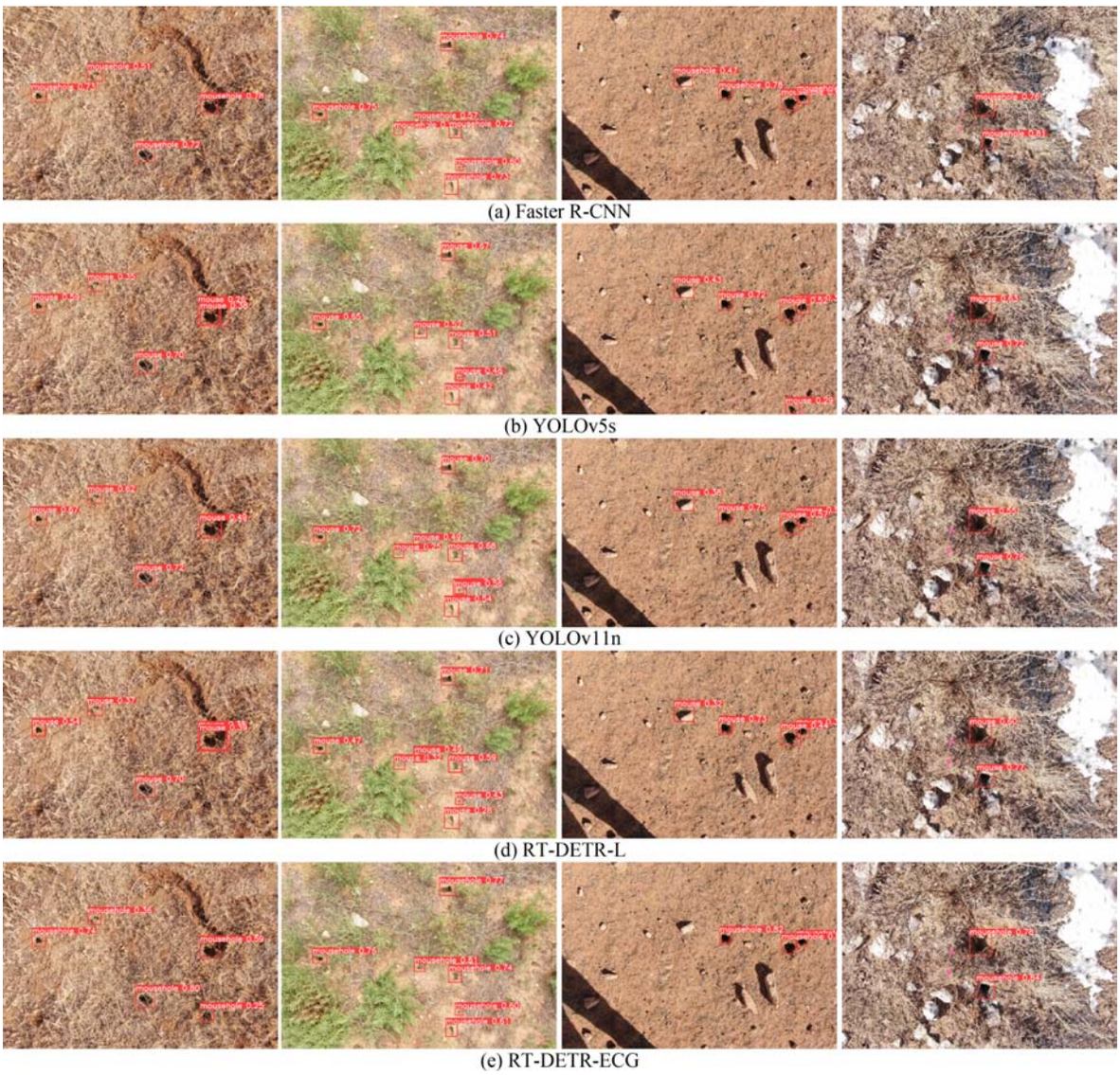


图 9　不同模型的检测效果

Figure 9　Detection results of different models

表 2　消融试验的结果

Table 2　Results of ablation experiments

方法				*mAP*@0.5/%	检测速度/(帧·s^{-1})	参数量/M
RT-DETR	C2f	EfficientVIM	CGLU			
√				91.2	76.7	103.4
√	√			92.9	75.3	98.7
√		√		93.8	90	94
√			√	94.8	86.7	82
√	√	√		91.5	70	60
√		√	√	93	89	79
√	√	√	√	96.3	91.8	47.9

更值得关注的是，当 RT-DETR 与 EfficientVIM、C2f、CGLU 全部结合时，*mAP*@0.5 相比基线模型提升 5.1 个百分点至 96.3%，帧率达到 91.8 帧/s，参数量仅 47.9 M，在检测精度、运行速度和模型轻量化之间实现

了较优的综合性能表现,表明所提出的RT-DETR-ECG模型在多模块协同设计下具有良好的整体性能。

3.5 可视化分析

为直观验证RT-DETR-ECG模型在鼠洞检测任务中的性能优势,对验证集样本进行了可视化分析。借助HiResCAM生成的注意力热图,如图10。可以直观揭示模型在特征提取与决策过程中的响应区域与注意力分布,从而反映不同模块在特征聚焦与信息过滤中的作用机制[31-32]。试验结果表明,改进后的模型在复杂背景下展现出更清晰、稳定的特征聚焦能力。CGLU模块通过卷积—门控机制动态调节特征通道权重,使模型能够自动抑制植被、阴影、碎石等复杂背景干扰,显著降低了误检率。与此同时,检测头中引入的DWConv结构有效扩大了局部感受野,使模型在鼠洞部分被草丛或泥土遮挡的情况下,仍能保持较高的识别精度与稳定性。

综合对比结果表明,RT-DETR-ECG在复杂草原场景下具备更强的鲁棒性与泛化能力,在小目标检测中实现了精度与实时性的双重提升,展现出优异的实用性与应用潜力。

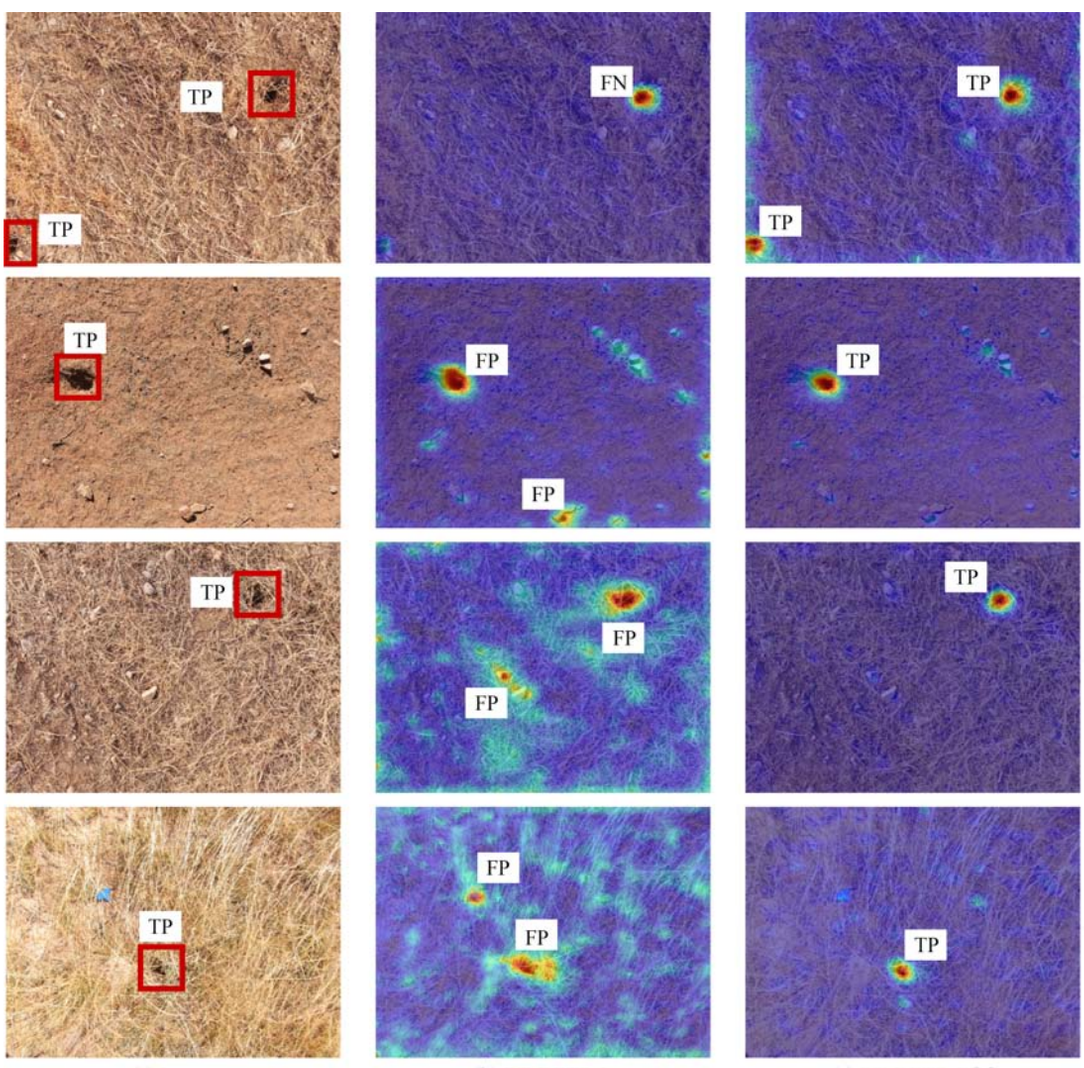


图10 HiResCAM可视化结果

Figure 10 HiResCAM visualization results

3.6 洞穴类型区分与生态学解释

为避免生态学解释上的潜在偏差,需指出的是,本文模型的检测目标为无人机影像中的地表洞口,其形态特征与草原常见的小型啮齿动物(如布氏田鼠、达乌尔鼠等)所形成的洞口较为一致。但在实际草地生态系统中,部分洞口亦可能由其他动物(如野兔、旱獭等)挖掘形成,或由自然风蚀过程导致,存在一定的来源多样性。

亦有学者基于传统计算机视觉与图像处理方法对地表洞口进行识别与分析[33-35],但该类方法在复杂自然背景下对光照变化与纹理干扰较为敏感。本研究所使用的数据来源于草原生态监测部门提供的典型鼠害

监测样区，并在采集前由长期驻地工作人员完成物种确认与栖息地标注，因此研究区域内的主要洞口类型可视为布氏田鼠形成。需要强调的是，本研究方法本质上属于“洞口检测”，尚无法通过影像特征直接区分洞口的物种来源。

基于上述原因，本研究模型的检测结果更适用于洞口空间分布、密度变化趋势、季节性活动动态等生态指标的构建，而不用于鼠类绝对数量的精确估算。未来工作将综合考虑洞口尺寸参数、周边扰动特征、动物行为数据及多模态传感信息，构建更精细的洞口分类模型，以提升对不同小型兽类洞穴的区分精度，从而增强生态监测的科学解释能力。

4　结论

4.1　主要研究贡献

本研究针对草原鼠洞小目标特征弱、背景干扰强和跨季节变化显著等问题，提出了一种基于改进RT-DETR的草原鼠洞智能识别与检测模型RT-DETR-ECG。该模型通过引入EfficientVIM全局－局部混合感知模块、CGLU卷积门控增强模块以及小目标加权损失函数，有效提升了小目标特征表达能力与定位精度。

在自建的2 397张草原无人机影像数据集上，RT-DETR-ECG在检测性能、推理速度和模型复杂度方面均取得了显著改进。试验表明：相较于原始RT-DETR模型，RT-DETR-ECG的*mAP*@0.5提升了5.1%（从91.2%提升至96.3%）；推理速度提高15.1帧/s（由76.7帧/s提升至91.8帧/s）；模型参数量减少55.5 M（由103.4 M降至47.9 M），实现了检测精度、实时性与轻量化的同步优化。此外，消融试验结果表明，单独引入EfficientVIM、CGLU和C2f模块均能够在一定程度上提升模型检测性能。其中，EfficientVIM模块使*mAP*@0.5提升约2.6%，而CGLU模块带来的性能增益更为显著，*mAP*@0.5提升约3.6%。在复杂地表纹理、多种光照条件及不同季节场景下，各改进模块均表现出更强的鲁棒性，展现出良好的草原生态监测与智能巡查应用潜力。

综合来看，RT-DETR-ECG在复杂地表纹理、多种光照条件和不同季节场景下均表现出更强的鲁棒性，能够有效减少鼠洞漏检和误检，具有良好的草原生态监测与智能巡查应用前景。未来研究将进一步探索更高效的注意力结构、跨区域迁移能力增强方法以及与地理信息系统（GIS）结合的三维鼠洞分布建模，以拓展模型在大尺度生态监测场景中的应用潜力。

4.2　不足与展望

尽管RT-DETR-ECG在多项指标上取得了显著提升，但仍存在一定局限性。首先，模型在极端光照、严重遮挡和极度低对比度场景下仍可能出现漏检，说明对极端场景的鲁棒性仍有优化空间。其次，本研究的数据集主要来自固定区域无人机影像，跨区域泛化能力仍需进一步验证。未来将从以下3个方向展开研究。

1）模型压缩与轻量化优化。针对模型参数量较大、计算复杂度较高的问题，后续将研究网络剪枝、量化与知识蒸馏等模型压缩技术，以降低计算开销和硬件依赖，推动模型在无人机及便携式检测终端等边缘设备上的部署应用。

2）多模态融合检测。考虑单一视觉模态在极端光照或遮挡条件下的局限，后续将探索融合红外成像、激光雷达等多源数据的检测框架，提升模型在复杂环境下的检测精度与适应性。

3）实际场景部署与验证。面向草原鼠害防控应用，进一步优化模型的实时推理速度与能耗表现，构建“模型—设备—任务”协同优化体系，推动检测技术从实验室验证向规模化实际应用转化。

参　考　文　献

[1] 王耀彬，王尧，孙伟成，等. 内蒙古国家级自然保护区草原生态系统健康评价[J]. 中国草地学报，2025，47(1)：1－12.
WANG Yaobin，WANG Yao，SUN Weicheng，et al. Grassland ecosystem health assessment in Inner Mongolia′s national nature reserves [J]. Chinese Journal of Grassland，2025，47(1)：1－12.

[2] 郭正刚，庞晓攀. 草原啮齿动物学课程思政建设路径的构建[J]. 草业科学，2025，42(8)：2144－2151.
GUO Zhenggang，PANG Xiaopan. Path of ideological curriculum and politics construction of grassland rodent science [J]. Pratacultural Science，2025，42(8)：2144－2151.

[3] 于媛媛，刘玉富，吴笑天，等. 草原生态系统退化机制与恢复技术研究[J]. 江西农业，2025(9)：190－192.

[4] 花立民，柴守权. 中国草原鼠害防治现状、问题及对策[J]. 植物保护学报，2022，49(1)：415－423.
HUA Limin，CHAI Shouquan. Rodent pest control on grasslands in China：Current state，problems and prospects [J]. Journal of Plant Protection，2022，49(1)：415－423.

[5] 刘楠，郭雪佳，褚东阳，等. 内蒙古典型草原布氏田鼠鼠洞无人机遥感识别研究[J/OL]. 内蒙古农业大学学报（自然科学版），1－13[2025－11－01].
LIU Nan，GUO Xuejia，CHU Dongyang，et al. Research

on UAV remote sensing identification of brandt's vole burrows in typical grasslands of Inner Mongolia [J/OL]. Journal of Inner Mongolia Agricultural University (Natural Science Edition), 1－13[2025－11－01].

[6] 郭秀明, 王大伟, 刘升平, 等. 基于深度学习的近地面草原鼠洞识别计数关键问题研究与应用[J]. 浙江农业学报, 2024, 36(9): 2146－2154.
GUO Xiuming, WANG Dawei, LIU Shengping, et al. Study on key problems for rat hole recognition and count near ground based on deep learning and its application [J]. Journal of Zhejiang Agriculture, 2024, 36(9): 2146－2154.

[7] 姚继林, 刘宏哲, 张铖, 等. 基于多模态融合网络的无人机小目标检测方法[J/OL]. 计算机工程与应用, 1－12[2025－11－01].
YAO Jilin, LIU Hongzhe, ZHANG Cheng, et al. A drone small target detection method based on multi-modal fusion network [J/OL]. Computer Engineering and Applications, 1－12[2025－11－01].

[8] 景婷婷, 曹玉东, 陈鑫, 等. 改进YOLOv11的无人机航拍小目标检测算法[J/OL]. 计算机工程与应用, 1－12[2025－11－01].
JING Tingting, CAO Yudong, CHEN Xin, et al. Improved YOLOv11 algorithm for aerial small target detection in UAVs [J/OL]. Computer Engineering and Applications, 1－12[2025－11－01].

[9] 王世勤. 基于Faster-RCNN深度学习算法的图像识别技术研究[J]. 信息记录材料, 2025, 26(4): 120－122.

[10] 帖军, 龙吕佳, 郑禄, 等. 融合多尺度特征和多重注意力的棉田杂草检测研究[J]. 中国农机化学报, 2025, 46(10): 138－145.
TIE Jun, LONG Lüjia, ZHENG Lu, et al. Research on weed detection in cotton fields integrating multi-scale features and multi-attention [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(10): 138－145.

[11] JIANG L, YUAN B, DU J, et al. MFFSODNet: Multiscale feature fusion small object detection network for UAV aerial images [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1－14.

[12] ALI M L, ZHANG Z. The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection [J]. Computers, 2024, 13(12): 336.

[13] JEGHAM N, KOH C Y, ABDELATTI M, et al. YOLO evolution: A comprehensive benchmark and architectural review of YOLOv12, YOLO11, and their previous versions [J]. arXiv preprint arXiv: 2411. 00201, 2024.

[14] 杨春霞, 吴亚雷, 闫晗, 等. 融合双图卷积与门控线性单元的方面级情感分析模型[J]. 计算机工程, 2024, 50(4): 141－149.
YANG Chunxia, WU Yalei, YAN Han, et al. Aspect-level sentiment analysis model combining double graph convolution and gated linear unit [J]. Computer Engineering, 2024, 50(4): 141－149.

[15] 康斓, 苏志金. 图像数据增强技术原理与发展综述[J]. 信息技术, 2024(9): 176－185.
KANG Lan, SU Zhijin. Principle and prospect of image data enhancement technology [J]. Information Technology, 2024(9): 176 - 185.

[16] 施宇航, 陈琳琳, 郭峰, 等. 基于自适应通道特征交互融合图卷积网络的骨骼行为识别[J/OL]. 山西大学学报(自然科学版), 1－12[2025－11－01].
SHI Yuhang, CHEN Linlin, GUO Feng, et al. Adaptive channel feature interactive fusion network for skeleton-based action recognition [J/OL]. Journal of Shanxi University (Natural Science Edition), 1－12[2025－11－01].

[17] 洪永, 张鑫, 林铭俊, 等. 基于深度可分离卷积与注意力机制的单导联心房颤动轻量级分类网络[J]. 南方医科大学学报, 2025, 45(3): 650－660.
HONG Yong, ZHANG Xin, LIN Mingjun, et al. A lightweight classification network for single-lead atrial fibrillation based on depthwiseseparable convolution and attention mechanism [J]. Journal of Southern Medical University, 2025, 45(3): 650－660.

[18] 夏正龙, 李威龙, 李灿, 等. 基于改进的VisionMamba的低光照图像处理研究[J/OL]. 计算机应用与软件, 1－10[2025－11－01].
XIA Zhenglong, LI Weilong, LI Can, et al. Research on low-light image processing based on improved VisionMamba [J/OL]. Computer Applications and Software, 1－10[2025－11－01].

[19] 赵喆, 李勃, 徐文校, 等. 融合多尺度特征的SAR图像目标检测方法[J/OL]. 电子测量技术, 1－13[2025－11－01].
ZHAO Zhe, LI Bo, XU Wenxiao, et al. Multiscale feature fusion for object detection in SAR images [J/OL]. Electronic Measurement Technology, 1－13 [2025－11－01].

[20] 郭晓新, 李广宇, 杨广奇. 基于前馈卷积神经网络的多模块自监督降噪模型[J/OL]. 吉林大学学报(工学版), 1－9[2025－11－01].
GUO Xiaoxin, LI Guangyu, YANG Guangqi. Multi-modular self-supervised image denoising network based on feed-forward convolutional neural network [J/OL]. Journal of Jilin University (Engineering and Technology Edition), 1－9[2025－11－01].

[21] SUN C, BAO Y, BAO Y, et al. Detection and classification of brandt's vole burrow clusters utilizing GF-2 satellite imagery and faster R-CNN model [J]. Frontiers in

Ecology and Evolution, 2024, 12: 1310046.

[22] ZHAO H, JIA L, WANG Y, et al. Autonomous UAV detection of ochotona curzoniae burrows with enhanced YOLOv11 [J]. Drones, 2025, 9(5): 340.

[23] KARIMINEJAD N, SHAHABI H, GHORBANZADEH O, et al. Evaluation of various deep learning algorithms for landslide and sinkhole detection from UAV imagery in a semi-arid environment [J]. Earth Systems and Environment, 2024, 8(4): 1387－1398.

[24] 邓炜航，赵胜雪．基于改进Faster RCNN算法的马铃薯叶片病害识别[J/OL]．农机化研究，1－8[2025－11－01].
DENG Weihang, ZHAO Shengxue. Potato leaf disease recognition based on improved Faster RCNN algorithm [J/OL]. Journal of Agricultural Mechanization Research, 1－8 [2025－11－01].

[25] 陈文轩，杨风暴，李菠，等．改进YOLOv5s的轻量化无人机航拍小目标识别算法[J/OL]．电子测量技术，1－12[2025－11－01].
CHEN Wenxuan, YANG Fengbao, LI Bo, et al. Improve the small target recognition algorithm of YOLOv5s for lightweight UAV aerial photography [J/OL]. Electronic Measurement Technology, 1－12[2025－11－01].

[26] 陈家政，付根平，黄伟锋，等．基于改进YOLOv5s的云机器人视觉导航方法[J]．智能化农业装备学报(中英文)，2025，6(3)：98－110.
CHEN Jiazheng, FU Genping, HUANG Weifeng, et al. Visual navigation method based on improved YOLOv5s and cloud robot [J]. Journal of Intelligent Agricultural Mechanization, 2025, 6(3): 98－110.

[27] 吴擎，韦润轩，周乐，等．基于改进YOLOv5s的轻量化鲜茶叶识别方法[J]．智能化农业装备学报(中英文)，2025，6(1)：1－14.
WU Qing, WEI Runxuan, ZHOU Le, et al. Lightweight fresh tea leaf recognition method based on improved YOLOv5s [J]. Journal of Intelligent Agricultural Mechanization, 2025, 6(1): 1－14.

[28] 刘方轩，李为相，肖迪，等．改进YOLOv11n的输电线路缺陷检测算法[J/OL]．计算机工程与应用，1－18[2025－11－01].
LIU Fangxuan, LI Weixiang, XIAO Di, et al. Improved YOLOv11n algorithm for defect detection in transmission lines [J/OL]. Computer Engineering and Applications, 1－18[2025－11－01].

[29] 陈仁祥，付铃华，杨黎霞，等．改进RT-DETR的航拍图像机动车违停检测[J]．中国惯性技术学报，2025，33(9)：864－870.
CHEN Renxiang, FU Linghua, YANG Lixia, et al. Improving RT-DETR for detecting vehicle parking violations inaerial images [J]. Journal of Chinese Inertial Technology, 2025, 33(9): 864－870.

[30] 董明书，陈俐企，马川义，等．沥青路面内部裂缝雷达图像智能判识算法研究[J]．山东大学学报(工学版)，2025，55(3)：72－79，87.
DONG Mingshu, CHEN Liqi, MA Chuanyi, et al. Deep learning-based intelligent judgment for radar detection of pavement cracks [J]. Journal of Shandong University (Engineering Edition), 2025, 55(3): 72－79, 87.

[31] 杨瑞珍，周云，张阳，等．基础设施日光热成像无损检测研究综述[J/OL]．工程力学，1－22[2025－11－01].
YANG Ruizhen, ZHOU Yun, ZHANG Yang, et al. Review on solar loading thermography in non-destructive testing of infrastructures [J/OL]. Engineering Mechanics, 1－22[2025－11－01].

[32] 张双双，刘琦，张济国．红外热成像设备目标检测原理及影响因素分析[J]．中国安全防范技术与应用，2024(4)：43－52.

[33] 高振，卢彩云，李洪文，等．基于计算机视觉的种子分布信息检测关键技术研究现状与趋势[J]．智能化农业装备学报(中英文)，2023，4(3)：50－60.
GAO Zhen, LU Caiyun, LI Hongwen, et al. Research progress and the prospect of crucial technology of seed spacing information detection based on computer vision [J]. Journal of Intelligent Agricultural Mechanization, 2023, 4(3): 50－60.

[34] 刘畅，孙雨，杨晶，等．基于3C-YOLOv8n和深度相机的葡萄识别与定位方法[J]．智慧农业(中英文)，2024，6(6)：121－131.
LIU Chang, SUN Yu, YANG Jing, et al. Grape recognition and localization method based on 3C-YOLOv8n and depth camera [J]. Smart Agriculture, 2024, 6(6): 121－131.

[35] 徐辉煌，吴敏，王梦露．基于图像处理的菊花收缩率及色泽变化研究[J]．智能化农业装备学报(中英文)，2024，5(4)：24－38.
XU Huihuang, WU Min, WANG Menglu. Study on the kinetics of shrinkage and color changes of chrysanthemum based on image processing [J]. Journal of Intelligent Agricultural Mechanization, 2024, 5(4): 24－38.

Design and experiment of an intelligent recognition and detection model for grassland rodent burrows based on improved RT-DETR

DONG Zhenwei, FU Xueliang, LI Honghui, PAN Xin, XU Zhe, LUO Xiaoling*

(*College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot* 010018, *China*)

Abstract: Abstract: In recent years, the rapid proliferation and expansion of rodent burrows have become one of the major causes of grassland degradation and ecological imbalance. The dense distribution of burrows destroys turf structure, weakens soil stability, reduces vegetation coverage and forage productivity, and seriously threatens the ecological security and sustainable development of animal husbandry in grassland areas. To address the problems of insufficient small-target recognition, inadequate feature extraction, and significant background interference in existing detection algorithms, this study proposes an improved real-time detection model, RT-DETR-ECG (efficient global perception and feature selection), based on the Real-Time Detection Transformer (RT-DETR) framework. The model first introduces the Efficient Vision Mamba (EfficientVIM) module, which achieves efficient fusion of global and local features through state-space modeling and multi-scale feature interaction, thereby enhancing the perception of small targets. Second, it constructs a Convolutional Gated Linear Unit (CGLU) module, which dynamically suppresses complex background noise via the synergy of convolution and gating mechanisms, effectively highlighting the edges and textures of burrows. Finally , it designs a small-object weighted loss function (CGLU-Loss), which introduces center-offset weighting and local-overlap gain terms in the regression stage to strengthen gradient contributions from small-scale samples and improve localization precision. Experiments based on a dataset of 2 397 cross-season UAV images of grassland burrows demonstrate that RT-DETR-ECG outperforms existing models in detection accuracy, real-time performance, and lightweight efficiency. With an *IoU* threshold of 0.5, the model achieves a mean average precision (*mAP*@0.5) of 96.3%, 5.1% higher than the original RT-DETR, with a detection speed of 91.8 f/s and a 53.7% reduction in computational load. These results verify that RT-DETR-ECG effectively balances detection accuracy and efficiency, providing a reliable and efficient approach for intelligent grassland rodent burrow detection and ecological monitoring.

Keywords: RT-DETR-ECG; object detection; deep learning; image recognition; UAV imagery; rodent burrow

Foundation items: High-level and Excellent Doctoral Talent Introduction Scientific Research Start-up Project (NDYB2022-60); Natural Science Foundation of Inner Mongolia Autonomous Region (2023LHMS06020); Key Project of Natural Science Foundation of Inner Mongolia (2025ZD012); "Revealing the List and Taking the Lead" Project of Inner Mongolia Autonomous Region (2025KJTW0026); Key Science and Technology Project of Inner Mongolia Autonomous Region (2025KYPT0076)

First Author: DONG Zhenwei, E-mail: dongzw@emails.imau.edu.cn

* **Correspondence Author:** LUO Xiaoling, E-mail: luoxl@imau.edu.cn