

基于 YOLOv11+CBAM 注意力机制的遥感图像滑坡检测

翁 存¹ 陈新房^{1,2} 杨丽佳¹ 王壮志¹ 孙梦雪¹ 吴笛白¹ 沈美怡¹ 汪世伟¹

WENG Cun CHEN Xinfang YANG Lijia WANG Zhuangzhi SUN Mengxue WU Dibai SHEN Meiyi WANG Shiwei

摘要

由地震引发的滑坡是山区强震后常见且危害显著的地质灾害类型，快速、准确获取滑坡的具体分布信息对震后应急响应具有重要的协助意义。针对传统人工存在的解译效率低、主观性强的问题，文章基于深度学习目标检测方法，提出一种融合 CBAM 注意力机制的 YOLOv11 滑坡目标检测模型。该方法通过在 YOLOv11 网络中引入通道注意力与空间注意力模块，自适应增强滑坡关键区域特征表达能力。通过在遥感滑坡数据集上进行对比实验，所提模型在 mAP50 和 mAP50-95 指标上分别达到 0.981 3 和 0.675 4，均优于 Faster R-CNN 及原始 YOLOv11 模型，表现出更高的检测精度与稳定性。研究结果表明，所提方法可有效提升遥感滑坡识别性能，为震后滑坡灾害的定位与快速评估提供可靠技术支持。

关键词

CBAM；目标检测；YOLO；注意力机制

doi: 10.3969/j.issn.1672-9528.2026.07.002

0 引言

地震引发的滑坡是内陆山区强震活动中最常见的地质灾害类型，数量大、规模大。根据《2024 年中国自然资源公报》数据统计，2024 年，我国共发生各类地质灾害 5 719 起，其中滑坡 3 316 起^[1]。地震诱发滑坡易导致堰塞湖形成、道路损毁和桥梁断裂等次生灾害，影响救援人员进入灾区以及获取第一手资料，给震后救援工作带来不利影响^[2]。基于遥感卫星影像建立滑坡识别模型可以提高滑坡空间分布信息识别准确性，有助于震后快速了解灾区情况并开展相关工作^[3]，因此有着十分重要的实际意义。

传统方法主要依靠人工解译和人工经验判断进行滑坡识别，无法达到快速准确识别的要求。而机器学习由于其高效性和鲁棒性，在大规模遥感图像中进行滑坡识别发挥重要作用。如支持向量机^[4]、逻辑回归以及随机森林等分类器被大量应用到滑坡识别当中^[5-6]，但是它们在复杂的遥感环境中对于高层语义信息表示不足，不能很好地应对滑坡大小不同、形状各异以及周围环境的影响等问题。

2012 年，Krizhevsky 等^[7]在 ImageNet 图像识别竞赛中首次取得突破性成果，标志着深度学习在计算机视觉领域的快速崛起。随后的 VGG、ResNet 等网络进一步推动了视觉识别技术的发展，使深度学习方法逐渐成为图像处理领域的

主流方案^{[8][9]}。这些经典卷积神经网络为后续视觉任务奠定了重要的特征提取基础，但其主要面向自然场景的整图分类设计，缺乏对遥感影像中复杂空间结构与多尺度目标的显式建模能力。

近年来，目标检测领域得到迅速发展的深度学习方法也被应用于滑坡检测的研究当中，利用端到端学习完成对滑坡的自动定位及检测，在复杂的遥感自然灾害中提供一种更优的方法。YOLO（you only look once）系列模型逐渐成为高效实时检测方法。而 YOLOv5、YOLOX 和 YOLOv7 等又在此基础上改进网络结构以及训练方式，使得该类模型保持较高准确率的同时，还具有很高的速度以及部署效率，对遥感图像的目标检测起到很好的支持作用。YOLO 系列模型虽然有着很好的实时性和检测效果，但是在复杂的遥感滑坡中也有许多不足之处，例如背景干扰严重、不规则目标建模困难、多尺度特征表示不足，以及对重要部分的关注度较低等问题。

本研究基于 YOLOv11m 模型引入 CBAM 注意力机制来提高模型对滑坡图像特征的学习能力进而提升模型对滑坡的检测效果。

1 模型与研究方法

1.1 YOLOv11

YOLOv11 是基于端到端单阶段的目标检测模型，由三部分组成：主干网络（Backbone）、特征融合网络（Neck）以及检测头（Head）。主干网络通过大量的卷积操作及特征提取模块对输入图片进行学习，在多个尺度上获取从底层纹理信息到高层语义信息的不同层次特征，为后面的检测做准备。

1. 应急管理大学 河北廊坊 065201

2. 河北省高校智慧应急应用技术研发中心 河北廊坊 065201

[基金项目] 廊坊市科学技术研究与发展计划自筹经费项目 (2024011018)

YOLOv11 的颈部加入了多尺度特征融合,在来自不同层次特征图之间进行有效结合,使模型可以学习到更多的尺度信息,以提高对具有较大尺度差异的目标检测精度。而检测头则在这些融合之后的多尺度特征图上同时完成对每个位置的类别预测以及该位置对应的边框偏移量估计,即完成了对目标位置和类别的联合建模。

YOLOv11 在保持较快推理速度基础上,提高对特征表示以及目标检测准确性,具有较好速度与精度折中表现,在复杂地形、干扰严重环境下进行滑坡目标检测方面有着较大优势。

1.2 CBAM

CBAM 结构如图 1 所示。卷积注意力模块(convolutional block attention module, CBAM)是一种轻量化且高效的注意力机制,通过通道注意力与空间注意力的顺序建模,引导网络在特征提取过程中更加关注与滑坡相关的关键语义特征及其空间分布。相较于仅依赖局部感受野的传统卷积操作, CBAM 能够在不显著增加计算开销的前提下,自适应地强化重要特征响应、抑制无关背景干扰,从而提升模型对复杂地形条件下滑坡区域的判别能力。

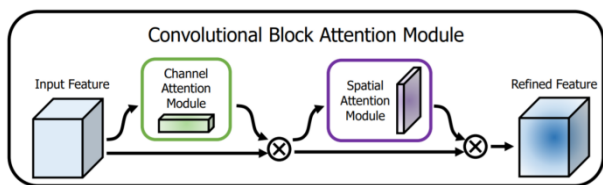


图 1 CBAM 注意力模块示意图

CBAM 中通道注意力模块和空间注意力模块,可以用式(1)~(2)表示:其中给定中间层特征图 $F \in \mathbb{R}^{C \times H \times W}$ 作为输

入, CBAM 依次推断一个 1D 通道注意力图 $M_c \in \mathbb{R}^{C \times 1 \times 1}$ 和一个 2D 空间注意力图 $M_s \in \mathbb{R}^{1 \times H \times W \times 10}$ 。

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

在滑坡目标检测任务中,将 CBAM 引入 YOLOv11 的 Backbone 可有效增强网络对滑坡判别特征的建模能力。CBAM 通过通道与空间注意力的协同作用,对特征提取阶段的关键信息进行自适应强化:

(1) 通道注意力突出与滑坡相关的高判别性特征,抑制背景与光照等干扰;

(2) 空间注意力进一步聚焦形态复杂、尺度变化显著的潜在滑坡区域,提升对小尺度目标和复杂背景的感知能力。

由此,在主干网络关键层嵌入 CBAM 能提升特征表示的判别性,为后续 Neck 与 Head 提供更精细特征输入,提高模型在复杂地形条件下的滑坡检测精度与定位性能。

1.3 总体模型

YOLOv11+CBAM 模型是在 YOLOv11 基础上引入卷积块注意力模块构建而成。该模型整体仍采用 YOLOv11 的单阶段检测框架。在特征提取阶段, YOLOv11 的 Backbone 负责从输入遥感影像中提取多尺度语义特征。为增强模型对滑坡区域关键特征的感知能力,本文在主干网络与特征融合路径中引入 CBAM 模块从而自适应地对特征图进行加权,从而突出与滑坡目标相关的判别性特征,整体模型结构图如图 2 所示。

在模型训练方面,训练轮数设为 250 epochs,输入影像尺寸为 512 px×512 px, Batch size 设为 24。初始学习率设为 lr0=0.003,采用 YOLO 默认动量与权重衰减策略进行训练更

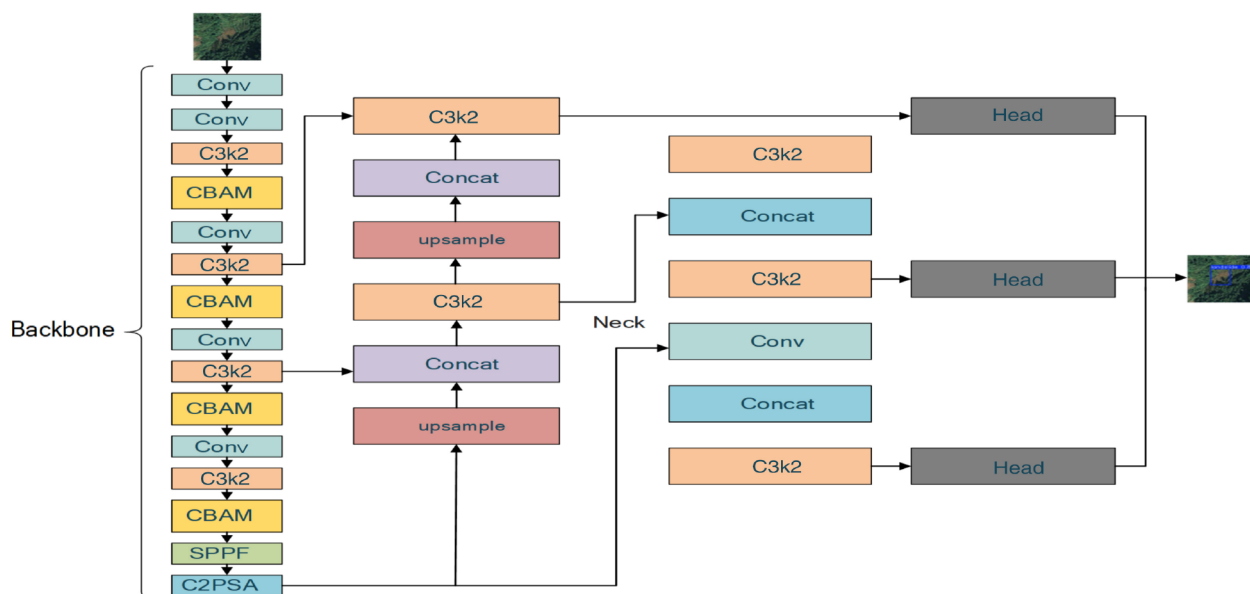


图 2 YOLOv11+CBAM 网络架构图

新。为避免过拟合，训练引入早停机制，当验证指标在 50 个 epoch 内未提升时自动中止训练。

2 实验过程

2.1 实验设置

本研究的实验环境配置为：Windows 11 64 位操作系统，处理器为 AMD Ryzen 7 7800X3D CPU，配备 32 GB 内存，并搭载 NVIDIA GeForce RTX 3070 GPU 作为计算加速设备。实验采用 Python 3.12 作为开发语言，使用 PyCharm 作为主要编程工具。

2.2 数据集

为验证所提出轻量化滑坡检测模型的有效性，本文选取贵州省毕节市公开发布的滑坡数据集开展实验研究。该数据集提供高分辨率遥感影像及对应的滑坡区域精细标注，其中滑坡范围以掩膜（Mask）形式给出，可用于模型的像素级监督训练。数据集共包含 770 个滑坡正样本，样本规模适中，为滑坡检测模型的构建与验证提供了可靠的数据基础。

2.3 数据处理

首先，对输入掩膜图像进行灰度化与阈值分割，生成二值掩膜以区分目标与背景，保证轮廓提取的稳定性。在此基础上，利用 OpenCV 轮廓检测算法提取前景连通区域，并为每个区域计算最小轴对齐外接矩形，获取目标在图像中的边界框位置与尺寸。随后，按照 YOLO 标注规范，将外接矩形的中心坐标及宽高按图像分辨率进行归一化，生成格式为（class_{id}, x_{centers}, y_{centers}, width, height）的检测框标注，所有参数均归一化至 [0, 1] 区间。最终，将检测框信息以文本形式输出，生成与掩膜图像一一对应的 YOLO 标注文件，YOLO 标注文件可视化效果如图 3 所示。



图 3 YOLO 标注文件可视化效果

本研究中，使用基于毕节市滑坡数据集构建的训练集与验证集，训练集比例 70%、验证集比例 20%，测试集比例 10%，为提升模型的泛化能力，训练过程中加入多种图像增强策略，包括色调调整、饱和度增强、亮度增强、垂直翻转、水平翻转以及尺度缩放等，以增强模型对光照变化、地形差异及不同成像条件的适应性。图像增强如图 4 所示。

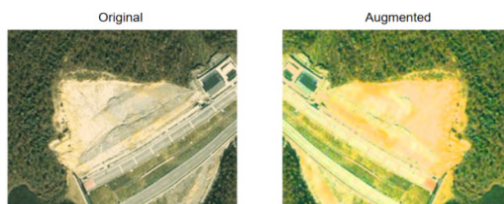


图 4 图像增强处理

2.4 评价指标

为了全面评估滑坡识别模型的性能，本文采用了以下评价指标：精确率、召回率、mAP50 和 mAP50-95。在计算各项指标之前，需先根据预测结果与实际结果构建混淆矩阵，如表 1 所示。

表 1 预测结果与实际结果的混淆矩阵

预测结果	实际为滑坡	实际为非滑坡
预测为滑坡	TP（真正例）	FP（假正例）
预测为非滑坡	FN（假反例）	TN（真反例）

各项评价指标的计算公式为：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{mAP}_{50} = \frac{1}{C} \sum_{c=1}^C \text{AP}_c^{\text{IoU}=0.50} \quad (5)$$

$$\text{mAP}_{50-95} = \frac{1}{10C} \sum_{t \in \{0.50, 0.55, \dots, 0.95\}} \sum_{c=1}^C \text{AP}_c^{\text{IoU}=t} \quad (6)$$

式中：精确率（Precision）表示被预测为滑坡的样本中，实际为滑坡的比例；召回率（Recall）表示实际为滑坡的样本中，被正确预测为滑坡的比例；IoU 阈值表示目标检测中用于判定预测边界框与真实标注框匹配关系的评价准则，当两者的交并比达到或超过该阈值时，预测结果被认定为正确检测。mAP50（mean average precision at IoU=0.5）在 IoU 阈值 0.5 下的平均检测精度，mAP50-95（mean average precision across IoU=0.50~0.95）在多个 IoU 阈值下计算的平均检测精度。

3 实验结果及分析

为系统评估不同目标检测模型在滑坡识别任务中的性能表现，本文选取了 Faster R-CNN、YOLOv11m 以及引入 CBAM 注意力机制的 YOLOv11+CBAM 模型进行对比实验。所有模型均在相同的数据集划分与训练策略下进行训练，训练轮数统一为 250 轮。为提升训练曲线的可读性并减弱小批量训练过程中随机性带来的波动影响，本文对各评价指标曲线进行了滑动平均处理，从而更加清晰地反映模型的整体收

敛趋势。

3.1 检测图对比分析

由图 5 可以看出, Faster R-CNN 滑坡目标上存在一定程度的冗余检测现象,即单一滑坡区域对应多个预测框。这一现象与其两阶段检测范式密切相关:RPN 模块在不同尺度与长宽比的 anchor 上生成高密度候选区域,以提高复杂场景下的目标召回率,但在边界不清晰、目标范围连续扩展的滑坡场景中,不同候选框之间的语义差异较小,导致后续分类与回归阶段难以完全消除冗余预测,冗余检测会增加后处理复杂度,不利于滑坡区域的快速判读与实际应用。

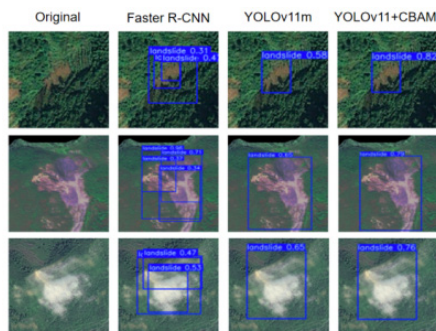


图 5 模型检测效果对比图

相比之下, YOLOv11 系列模型通过端到端训练直接回归目标边界框,在预测阶段引入空间位置与类别置信度的全局竞争机制,从结构层面限制了重复检测的产生,因此在视觉结果上表现为更少的检测框和更清晰的目标定位。

此外,引入 CBAM 后,模型能够在通道与空间维度上自适应强化与滑坡区域相关的判别特征,使网络更加关注具有显著地貌变化的区域,从而进一步提升检测置信度并改善目标定位稳定性。

3.2 精确率与召回率曲线分析

图 6~7 所示的精确率与召回率曲线可以看出, Faster R-CNN 在训练过程中表现出较为平稳的收敛特性,但其最终精度与召回率均相对较低,说明该两阶段检测模型在复杂滑坡场景下的特征表达能力存在一定局限。

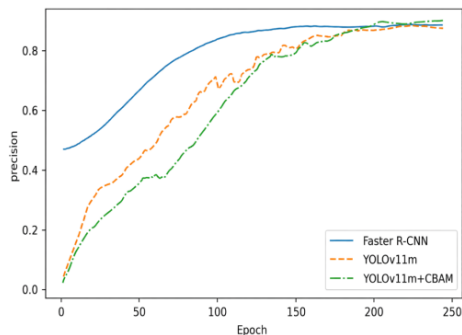


图 6 精确率曲线图

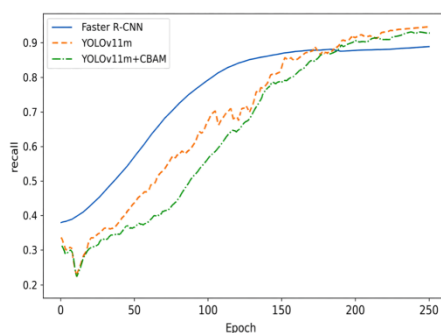


图 7 召回率曲线图

相比之下, YOLOv11m 在中后期训练阶段两项指标均快速提升,尤其在召回率方面取得了显著优势,单阶段检测框架在处理尺度变化大、形态不规则的滑坡目标时具有更强的检测能力。通过引入 CBAM 注意力机制后,模型能在保持较高召回率的同时,精度也得到了进一步的提升。这表明 CBAM 能够有效增强模型对滑坡区域特征的专注能力,抑制复杂背景的干扰,从而减少误检现象。

3.3 mAP50 与 mAP50-95 分析

如图 8~9 所示, mAP50 与 mAP50-95 的变化趋势进一步提供了有力的依据,支持了上述结论。Faster R-CNN 的 mAP 指标随训练轮次稳步提升,但最终的性能上限明显低于 YOLO 系列模型,尤其在 mAP50-95 指标上差距更加显著。YOLOv11m 在 mAP50 指标上已取得较高水平,而 YOLOv11m+CBAM 在 mAP50 与 mAP50-95 两项指标上均实现了进一步提升。

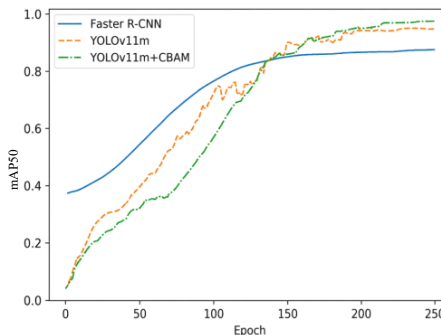


图 8 mAP50 曲线图

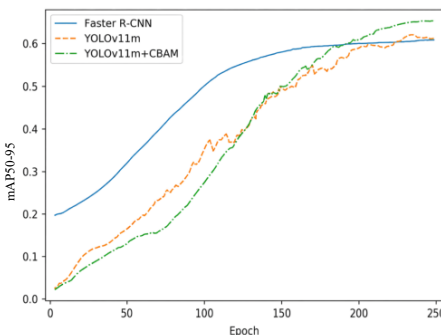


图 9 mAP50-95 曲线图

其中, mAP50-95 的提升尤为明显, 最高提升至 0.675 4, 说明引入注意力机制后, 模型在更严格 IoU 阈值下仍能保持较高检测性能, 显著增强了滑坡目标的定位精度。

3.4 数据分析

表 2 模型评估指标

模型	准确率	召回率	mAP50	mAP50-95
Faster R-CNN	0.908 5	0.864 6	0.862 7	0.611 6
YOLOv11m	0.941 7	0.974 0	0.977 4	0.653 7
YOLOv11+CBAM	0.954 0	0.972 1	0.981 3	0.675 4

根据表 2 的实验结果, 在检测精度方面, YOLOv11+CBAM 的精确率达到 0.954 0, 较 YOLOv11m 提升 0.012 3, 相比 Faster R-CNN 提升 0.045 5。在召回率方面, 其召回率为 0.972 1, 与 YOLOv11m 的 0.974 0 基本一致, 说明注意力机制的引入未对目标检出能力造成负面影响。在综合性能指标上, YOLOv11+CBAM 在 mAP50 和 mAP50-95 两项指标上均取得最优结果。mAP50 提升至 0.981 3, 较 YOLOv11m 和 Faster R-CNN 分别提高 0.006 1 和 0.118 6。在 mAP50-95 指标上提升至 0.675 4, 较两种对比模型分别提高 0.021 7 和 0.167 7, 表明所提出模型在不同 IoU 阈值下具备更稳定的目标定位能力和更强的泛化性能。

4 结论

本文围绕遥感滑坡目标检测中存在的复杂背景干扰和多尺度目标识别难题, 提出了一种基于 YOLOv11 的注意力增强检测模型。通过在网络中引入 CBAM 模块, 模型能够在特征提取阶段自适应聚焦于与滑坡相关的关键通道和空间区域, 从而提升特征表达的判别性。

实验结果表明, 本文所提出模型在多项评价指标上均优于对比模型以及基线模型, mAP50 提升至 0.981 3, mAP50-95 提升至 0.675 3, 尤其本文模型在定位精度和综合检测性能方面表现出更好的稳定性。YOLOv11+CBAM 中将注意力机制引入高效目标检测框架中, 对提升遥感滑坡识别精度具有积极作用, 可为相关灾害监测与应急响应提供有力的技术支撑。

参考文献:

- [1] 中华人民共和国自然资源部. 2024 年中国自然资源公报[R]. 北京: 中华人民共和国自然资源部, 2025.
- [2] 于宪煜, 杨森. 基于 Faster R-CNN 和 Mask R-CNN 的滑坡自动识别研究[J]. 大地测量与地球动力学, 2025, 45(1): 1-12.
- [3] 彭令, 徐素宁, 梅军军, 等. 地震滑坡高分辨率遥感影像识

别[J]. 遥感学报, 2017, 21(4): 509-518.

- [4] Hong Haoyuan, Pradhan B, Jebur M N, et al. Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines[J/OL]. Environmental Earth Sciences, 2016[2026-07-13].<https://link.springer.com/article/10.1007/s12665-015-4866-9>.
- [5] Chang K T, Chiang S H, Hsu M L. Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression[J]. Geomorphology, 2007, 89(3): 335-347.
- [6] Chen Tao, Trinder J C, Niu Ruiqing. Object-oriented landslide mapping using ZY-3 satellite imagery, random forest and mathematical morphology, for the Three-Gorges Reservoir, China[J]. Remote Sensing, 2017, 9(4): 333.
- [7] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[PP/OL]. (2015-04-10)[2026-07-13].<https://doi.org/10.48550/arXiv.1409.1556>.
- [9] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [10] Woo S, Park J, Lee J-Y, et al. CBAM: convolutional block attention module[C]//Computer Vision-ECCV 2018. Berlin: Springer, 2018: 3-19.

【作者简介】

翁存(2001—), 男, 江苏宿迁人, 硕士研究生, 研究方向: 机器学习、计算机视觉、人工智能等。

陈新房(1975—), 男, 山东汶上人, 硕士, 副教授, 研究方向: 图像处理、灾害数据处理、数据存储与分析等。

(收稿日期: 2025-12-19 修回日期: 2026-07-17)