

基于改进 YOLOv11 的矿下积水检测算法

文武洲¹
WEN Wuzhou

摘要

矿下积水严重威胁煤矿安全，其精准快速检测对保障人员安全至关重要。针对 YOLOv11n 模型在井下复杂环境中特征提取能力弱、小目标检测精度低的问题，文章提出改进模型 MSD-YOLOv11。通过设计多尺度特征提取模块 MSDconv，增强模型对积水特征的捕获能力；同时新增小目标检测层，融合浅层与深层特征，缓解小目标漏检问题。实验结果表明，改进后的模型参数量仅增至 2.81，相较于原始 YOLOv11n，精确率（P）提升至 86.47%，召回率（R）提升至 79.32%，mAP50 与 mAP50-95 分别达到 85.4% 和 49.89%，有效解决了背景干扰大小目标识别难的难题，为矿井积水监测提供了可靠方案。

关键词：积水检测；DWConv；YOLOv11；小目标检测

doi: 10.3969/j.issn.1672-9528.2026.02.047

0 引言

矿下积水检测技术的演进历程反映了工业监测领域从人力依赖向智能化发展的总体趋势。早期阶段主要依靠人工携带便携设备进行巡查，这种方法高度依赖作业人员的经验，且受巷道内视线阻碍、环境复杂等因素限制，容易导致检测延迟，漏检现象较为突出，统计显示漏检率常高于 35%。随

后出现的自动化监测系统，多以湿度、红外等传感器为基础，实现了对特定点的实时监测，但在实际矿井环境中，传感器易受粉尘堆积、电磁信号及温度变化等多重干扰，误报率经常超过 20%，同时也难以实现巷道区域的全面覆盖。

近年来，随着深度学习在计算机视觉领域的突破，基于目标检测的智能识别方法逐渐被引入矿山安全监测中。YOLO 等^[1-3]神经网络系列模型兼具速度与精度优势，能够对矿井全域进行实时、稳定的图像分析，有效解决了传统方

1. 西安工程大学 陕西西安 710600

cephalogram (EEG) motor imagery (MI) signals: a review[J]. Neural computing and applications, 2021,35:14681-14722.

[4]XIONG X, WANG Y, SONG T Y, et al.Improved motor imagery classification using adaptive spatial filters based on particle swarm optimization algorithm[J].Frontiers in neuroscience, 2023, 17: 1303648.

[5]LAWHERN V J, SOLON A J, WAYTOWICH N R, et al. EEGNet: a compact convolutional network for EEG-based brain-computer interfaces[EB/OL].(2018-05-16)[2025-12-30]. <https://doi.org/10.48550/arXiv.1611.08024>.

[6]SCHIRRMESTER R T, SPRINGENBERG J T, FIEDERER L D J, et al.Deep learning with convolutional neural networks for EEG decoding and visualization[J].Human brain mapping, 2017, 38(11): 5391-5420.

[7]DAGDEVIR E, TOKMAKCI M. Optimization of preprocess-ing stage in EEG based BCI systems in terms of accuracy and timing cost[J].Biomedical signal processing and control,

2021, 67: 102548.

[8]刘俊,张道强.脑机接口系统软件平台发展现状与挑战[J].中国生物医学工程学报,2022,41(3):350-360.

[9]李立亭.基于 WPD 和 CSP 的四类运动想象脑电信号分类方法[J].天津理工大学学报,2017,33(1):30-34.

[10]王文智,罗安飞,廖清静.基于 Adaboost 与卷积神经网络的人脸定位研究[J].科技与创新,2023(22):1-4.

【作者简介】

吴叶兰（1970—），女，湖北荆州人，硕士，副教授，研究方向：智能信息处理。

徐梦（2001—），男，湖北荆州人，硕士研究生，研究方向：脑机接口。

秦晴（2002—），女，云南曲靖人，硕士研究生，研究方向：脑机接口。

（收稿日期：2025-12-10 修回日期：2026-02-05）

法漏检与误报率高的问题,显著提升了矿下积水检测的自动化与可靠性。Jafari 等^[4]的研究采用了一种基于实时摄像与计算机视觉的水位监测方法。通过在监测区域部署摄像头,系统持续获取水体视频流,并借助图像处理算法自动检测水面边界。该技术通过将识别到的水体边界与预设的已知参考位置进行比对,从而实现对水位高度的实时计算。胡昊等^[5]采用视频监控技术,构建了道路积水实时识别框架,通过集成 YOLOv5s 目标检测模型提取图像特征,实现了积水区域的动态检测与面积估算。Chen 等^[6]采用 MobileNetV2 网络构建了隧道积水智能识别模型,并以此为基础研发了隧道内涝监测预警平台。该平台可对隧道积水情况进行自动辨识与实时预警,实现了从识别到预警的主动防控功能。Sazara 等^[7]采用预训练的深度学习结合传统图像特征提取方法(如局部二值模式与定向梯度直方图),对图像中的物体进行多类别划分,进而通过语义分割技术实现积水区域的精准检测。Witherow 等^[8]提出了一种基于数字图像处理的积水检测模型。该模型通过预处理与尺度不变特征变换(SIFT)图像配准技术识别道路淹没区域,但在实际应用中易将潮湿路面误判为积水,导致检测精度受到一定影响。在提升小目标检测性能方面,Tang 等^[9]、Zhang 等^[10]、Kang 等^[11]的研究均采用了引入额外微小检测头的方法,以更充分地利用网络中的高分辨率特征信息。

尽管现有方法取得了显著进展,但应用于矿井复杂环境下仍面临挑战。原始 YOLOv11n 模型在井下昏暗、杂乱背景中,存在积水特征提取不充分、小尺寸积水目标漏检率较高等局限性。为此,本文提出一种改进的轻量化检测模型 MSD-YOLOv11,在保持模型高效性的同时,着力提升对矿下积水特征的辨识能力与小目标检测精度,旨在实现更可靠的矿井积水智能监测。

1 MSD-YOLOv11

1.1 MSDConv 模块构建

针对矿下积水目标形态多变、背景复杂及小尺度特征易丢失的问题,本文提出了 MSDConv。该卷积模块通过并行多分支结构提取不同感受野下的特征,并引入深度可分离卷积减少参数量,在保持模型轻量化的同时,增强了对多尺度积水特征的融合与表征能力。

图 1 为 MSDConv 的网络结构。MSDConv 首先使用一个 3×3 标准卷积对输入特征进行初步特征提取,随后通过 3 条并行的卷积路径进一步提取多尺度特征,包括一个 3×3 深度可分离卷积(DWConv)、一个 3×3 标准卷积和一个 5×5 深度可分离卷积。其中深度可分离卷积在保留空间特征

提取能力的同时显著减少了计算量; 3×3 卷积核有助于捕捉局部细节特征,而 5×5 卷积核则能获取更大范围的上下文信息。

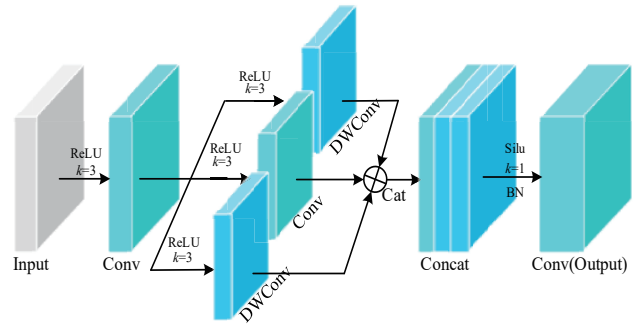


图 1 MSDConv 网络结构

模块的具体运算流程如下:

(1) 初步特征提取: 输入特征首先经过一个 3×3 标准卷积层,进行基础的上下文信息融合。

$$F_{\text{init}} = \text{Conv}_{3\times 3}^{(1)}(X) \quad (1)$$

式中: $\text{Conv}_{3\times 3}^{(1)}$ 表示该第一层标准卷积操作。

(2) 并行多尺度特征提取: 将 F_{init} 送入 3 条并行的路径,分别提取不同感受野下的特征。

$$F_{\text{dw3}} = \text{DWConv}_{3\times 3}(F_{\text{init}})$$

$$F_{\text{conv3}} = \text{Conv}_{3\times 3}^{(2)}(F_{\text{init}}) \quad (2)$$

$$F_{\text{dw5}} = \text{DWConv}_{5\times 5}(F_{\text{init}})$$

式中: $\text{DWConv}_{k\times k}$ 表示卷积核大小为 $k\times k$ 的深度可分离卷积;

$\text{Conv}_{3\times 3}^{(2)}$ 表示第 2 条路径中的标准卷积。

(3) 特征融合与残差连接: 将 3 条路径的输出特征沿通道维度进行拼接(Concat),并与原始输入 X 通过一个捷径连接相加,实现特征重用并促进梯度流动。

$$F_{\text{fuse}} = \text{Concat}(F_{\text{dw3}}, F_{\text{conv3}}, F_{\text{dw5}}) + \mathcal{P}(X) \quad (3)$$

式中: $\mathcal{P}(\cdot)$ 表示一个可选的 1×1 卷积投影操作,用于在输入与拼接特征通道数不匹配时调整 X 的维度。

(4) 归一化、激活与输出变换: 对融合后的特征进行批量归一化(batch normalization, BN),随后通过 SiLU 激活函数引入非线性,最后使用 1×1 卷积将通道数调整至目标输出维度 C_{out} 。

公式(1)~(3)展现了 MSDConv 模块的数据流情况,该设计通过并行异构卷积路径实现了对局部细节与全局上下文的协同感知,深度可分离卷积的运用在维持表征能力的同时控制了参数增长,残差连确保了信息的有效传递。

1.2 MSD-YOLOv11 模型搭建

尽管 MSDConv 模块增强了对多尺度特征的提取能

力，但由于矿下积水目标尺寸差异显著，仅优化卷积层仍难以有效避免小尺寸目标在深层网络中的特征丢失。导致小目标检测性能不足的根本原因在于，随着网络层数的加深与特征图的下采样，包含丰富细节与位置信息的浅层空间特征逐渐被稀释，而主要承载抽象语义信息的深层特征对小目标的表征能力有限。YOLO 系列模型通常采用 P3、P4、P5 三个检测头，其对应的特征图分辨率依次降低，对小目标的感知能力也随之减弱。为此，本文在原有检测头基础上新增了一个面向微小目标的 P2 检测头，对应 160 px×160 px 像素的高分辨率特征图。该特征图源自网络较浅层，能够保留更丰富的空间细节与轮廓信息，从而显著提升模型对小尺度积水目标的感知与定位能力。图 2 为 MSD-YOLOv11 模型结构图。

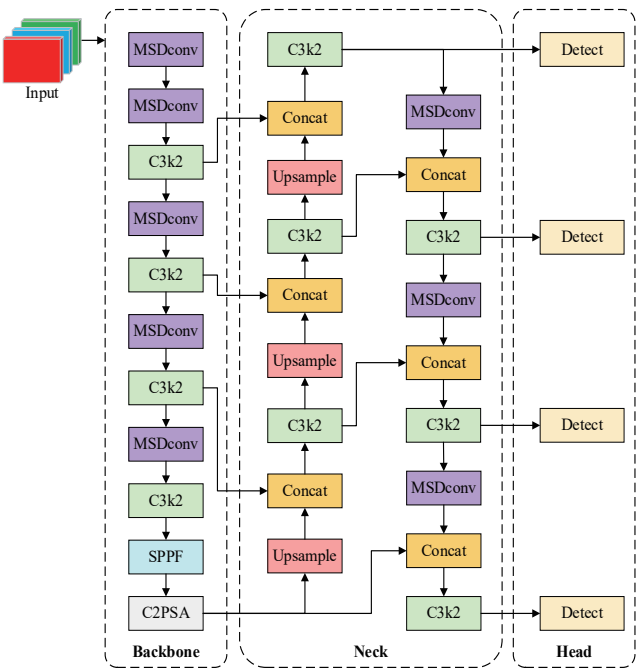


图 2 MSD-YOLOv11 模型结构图

2 实验设计与结果分析

2.1 实验数据

本研究使用自主构建的积水图像数据集。该数据集主要采集自煤矿井下作业现场及部分道路场景下的积水图像，共计 1 437 张。所有图像均通过 LabelImg 工具进行人工标注，统一设定类别标签为“puddle”。数据划分遵循 8:2 的比例，其中训练集包含 1 150 张图像，验证集包含 287 张图像。

2.2 实验配置

本次实验共进行 150 轮训练，初始学习率设为 0.01，其他主要参数配置如表 1 所示。为系统评估模型性能，实验采用精确率（P）、召回率（R）及平均精度均值（mAP）作为评价指标。

表 1 实验参数设置

名称	设置
操作系统	Windows11
中央处理器	12th Gen Intel(R) Core(TM) i7-12700H
图形处理器	NVIDIA GeForce RTX 4090 Laptop GPU
Python	v3. 9. 0
PyTorch	v2. 4. 1
CUDA	v11. 8

训练完成的模型能够对图像及动态视频流中的积水目标进行自动检测与定位，其推理结果如图 3 所示。

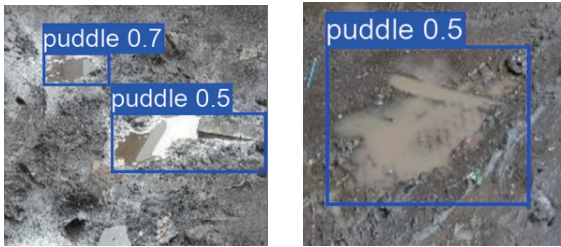


图 3 模型推理结果示例

2.3 评价指标

目标检测模型的性能需通过量化指标客观评估本文采用以下评估指标：平均精度均值（mean Average Precision@0.5，mAP@0.5）、精确率（Precision）、召回率（Recall）、参数量（Params）。

其中精确率表示正确预测框占有所有预测框的比例。其计算公式为：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

式中：TP 为模型正确检测到的目标数量；FP 为模型错误检测到的背景或非目标数量。

召回率表示检测到的目标占真实目标的比例。其计算公式为：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

式中：FN 为模型未检测到的真实目标数量。

平均精度均值结合精确率和召回率来全面反映模型的表现。其计算公式为：

$$\text{mAP@0.5} = \frac{\sum_{i=1}^n \text{AP}(i)}{n} \tag{6}$$

式中：AP 表示 P-R 曲线的面积。

2.4 不同模型对比

不同模型的性能对比如表 2 所示。首先，通过对比 YOLOv5n、YOLOv8n 与基线模型 YOLOv11n 的结果可以看出，基线模型的综合检测性能最优，其 mAP@0.5 达到 0.763，这为后续的改进奠定了良好的基础。

表 2 不同模型实验结果

模型	Precision	Recall	mAP@0.5	mAP50-95	Params/ 10^6
YOLOv5n	0.723	0.683	0.671	0.426	1.9
YOLOv8n	0.741	0.691	0.723	0.431	3.2
YOLOv11n (基线)	0.753	0.702	0.763	0.443	2.58
YOLOv11+MSDconv	0.831	0.754	0.826	0.476	2.72
MSD-YOLOv11	0.854	0.793	0.864	0.489	2.81

在 YOLOv11 中引入 MSDConv 卷积模块后，模型性能实现首次显著提升。该阶段模型的 mAP@0.5 提升至 0.826，较基线绝对值增加 0.063，增幅约为 8.3%。与此同时，模型参数量仅从 2.58×10^6 增至 2.72×10^6 ，增长约 5.4%。这一结果证明，MSDConv 模块通过其多尺度与轻量化设计，能够高效地增强网络对复杂背景下积水特征的提取能力，且计算开销可控。

最后，在融合 MSDConv 模块的基础上，进一步增加专用于小目标检测的 P2 特征层，构成完整的 MSD-YOLOv11 模型。由表 2 可知，该模型的各项指标均达到最优，其中 mAP@0.5 进一步提升至 0.864，召回率显著提升至 0.793。与仅加入 MSDConv 的模型相比，召回率提升超过 5%，这直接归因于 P2 层提供的丰富空间细节信息有效缓解了小目标的漏检问题。最终，MSD-YOLOv11 以 2.81×10^6 的参数量，在精确率、召回率与 mAP@0.5 上相比原始基线模型分别提升了 13.4%、13.0% 和 13.2%，实现了精度与轻量化的较好平衡。

3 总结

本研究针对矿下积水检测中背景复杂、目标尺度多变及小尺寸目标易漏检的难题，提出一种基于 YOLOv11n 的改进检测模型 MSD-YOLOv11。通过设计多尺度深度可分离卷积模块（MSDconv），在控制参数量的同时增强了对复杂环境下积水特征的提取能力，通过新增小目标检测层（P2），融合浅层细节信息，有效缓解了小尺度积水目标的特征丢失问题。实验结果表明，改进模型在仅小幅增加参数量的情况下，精确率、召回率与 mAP@0.5 相比基线模型显著提升，实现了精度与轻量化的较好平衡，为矿下积水实时监测提供了一种高效可靠的技术方案，对保障矿下安全生产具有积极意义。未来研究可进一步优化模型在极端环境下的鲁棒性，并推动其在实际边缘设备中的部署应用。

参考文献：

[1]MURAT A A, KIRAN M S. A comprehensive review on YOLO versions for object detection [J]. Engineering science and technology, an international journal, 2025, 70:102161.

[2]CHEN Y M, YUAN X B, WANG J B, et al. YO-LO-MS: rethinking multi-scale representation learning for real-time object detection [J]. IEEE transactions on pattern Analysis and machine intelligence, 2025, 47(6): 4240-4252.

[3] DOHERTY J, GARDINER B, KERR E, et al. Bi-FPN-YOLO: one-stage object detection integrating Bi-directional feature pyramid networks [J]. Pattern recognition, 2025: 111209.

[4]JAFARI N H, LI X, CHEN Q, et al. Real-time water level monitoring using live cameras and computer vision techniques[J]. Computers & geosciences, 2021, 147: 104642.

[5] 胡昊, 李擎, 马鑫, 等. 基于影像的道路积水监测研究 [J]. 华北水利水电大学学报 (自然科学版), 2023, 44(1): 62–70.

[6]CHEN Y, WANG Y W, LI K Y. Recognition and application of tunnel water accumulation based on computer vision [J].Journal of physics: conference series,2021: 012043.

[7]SAZARA C, CETIN M, IFTEKHARUDDIN K M. Detecting floodwater on roadways from image data with handcrafted features and deep transfer learning [C]// 2019 IEEE Intelligent Transportation Systems Conference (ITSC).Piscataway: IEEE, 2019: 804-809.

[8]WITHEROW M A, SAZARA C, WINTER-ARBOLEDA I M, et al. Floodwater detection on roadways from crowd-sourced images [C]//VI ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing (VipIMAGE). UK: Taylor & Francis Ltd, 2019: 529-540.

[9]TANG S Y, ZHANG S, FANG Y N. HIC-YOLOv5: improved YOLOv5 for small object detection [C]//2024 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2024: 6614-6619.

[10]ZHANG M, YIN L J. Solar cell surface defect detection based on improved YOLOv5 [J]. IEEE access, 2022, 10: 80804-80815.

[11]KANG M, TING C M, TING F F, et al. BGF-YOLO: enhanced YOLOv8 with multiscale attentional feature fusion for brain tumor detection [EB/OL].(2024-10-13)[2025-12-30]. <https://doi.org/10.48550/arXiv.2309.12585>.

【作者简介】

文武洲（2000—），男，陕西西安人，硕士研究生，研究方向：计算机视觉、深度学习。

（收稿日期：2025-12-05 修回日期：2026-02-04）