

彭小龙, 王正勇, 何小海, 等. 基于改进 YOLOv8 的岩石纹层图像分割研究[J]. 智能计算机与应用, 2025, 15(6): 1-6. DOI: 10.20169/j.issn.2095-2163.24121204

基于改进 YOLOv8 的岩石纹层图像分割研究

彭小龙¹, 王正勇¹, 何小海¹, 滕奇志¹, 龚剑²

(1 四川大学 电子信息学院, 成都 610065; 2 成都西图科技有限公司, 成都 610065)

摘要: 掌握油气岩石纹层的颜色、宽度、面积分布和纹路排列等特征对于油气资源的勘探和开发至关重要。然而, 由于岩石样本纹层的细密和紧凑特性, 人工观测和统计存在较大难度。因此, 岩石纹层图像分割技术成为了分析、评估和统计这些特征的重要工具。岩石纹层图像中的纹层排列紧密, 且颜色差异微小, 使得现有的图像分割技术难以实现精确区分。此外, 岩石纹层的连续性和图像采集设备高分辨率, 导致现有方法无法实现连续分割。为了提高岩石纹层分割的效果, 本文提出了一种基于改进 YOLOv8 的纹层分割算法。该算法的主要内容包括: 以 YOLOv8 网络为基础进行改进, 结合岩石纹层的形状分布特征, 采用自适应形状卷积核替换 YOLOv8 的浅层卷积模块, 以增强模型对纹层形状特征的表达能力; 引入基于 Vision Transformer 的空间注意力融合机制, 高效提取并加权融合岩石纹层样本不同区域的特征, 能够显著提升模型对不同颜色纹层的区分能力。实验结果表明, 针对高密度、颜色差异不明显的纹层时, 相较于传统分割方法和其他分割网络模型, 本文方法展现出更优的指标值, 具有更高的分割准确度和鲁棒性。

关键词: YOLOv8; 岩石纹层图像; 图像分割; 时空注意力机制; 多尺度特征提取

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)06-0001-06

Research on rock layer section image segmentation based on improved YOLOv8

PENG Xiaolong¹, WANG Zhengyong¹, HE Xiaohai¹, TENG Qizhi¹, GONG Jian²

(1 College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2 Chengdu Xitu Technology Co., Ltd., Chengdu 610065, China)

Abstract: Mastering the characteristics of oil rock layering such as color, width, area distribution, and pattern arrangement is crucial for the exploration and development of oil and gas resources. However, due to the fine and compact nature of rock samples' layering, manual observation and statistics are difficult. Therefore, rock layering image segmentation technology has become an essential tool for analyzing, evaluating, and statistically analyzing these features. The layers in rock layering images are closely arranged, and the color differences are minimal, making it difficult for existing image segmentation techniques to achieve precise differentiation. Moreover, the continuity of rock layering and the high resolution of image acquisition equipment lead to the inability of existing methods to achieve continuous segmentation. To enhance the effect of rock layering segmentation, this paper proposes an improved YOLOv8-based layering segmentation algorithm. The main content of the algorithm includes: improving the YOLOv8 network by combining the shape distribution characteristics of rock layering, using adaptive shape convolutional kernels to replace the shallow convolutional modules of YOLOv8, thereby enhancing the model's expressive ability for layering shape features; introducing a spatial attention fusion mechanism based on Vision Transformer, efficiently extracting and weighting the fusion of features from different regions of rock layering samples, which significantly improves the model's ability to distinguish between layers of different colors; experimental results show that this method, when dealing with highly dense and low color difference layering, exhibits superior segmentation performance compared to traditional segmentation methods and other segmentation network models, with higher segmentation accuracy and robustness.

Key words: YOLOv8; rock layer image; image segmentation; spatiotemporal attention mechanism; multi scale feature extraction

0 引言

在地质学研究和油气资源勘探领域, 对岩石纹

层特征的分析至关重要。纹层的形态、颜色、宽度和分布等特征对于揭示地质历史、重建沉积环境以及预测油气藏的潜在分布具有关键作用^[1-2]。然而,

基金项目: 国家自然科学基金 (62071315)。

作者简介: 彭小龙 (2000—), 男, 硕士研究生, 主要研究方向: 图像处理与模式识别。

通信作者: 王正勇 (1969—), 女, 博士, 副教授, 硕士生导师, 主要研究方向: 图像处理与模式识别, 计算机视觉。Email: 690728634@sina.com。

收稿日期: 2024-12-12

哈尔滨工业大学主办 ◆ 学术研究与应用

由于岩石纹层的细微差别和颜色的相似性,传统的人工观测和统计方法不仅效率低下,而且难以确保分析结果的准确性和一致性^[3]。目前,已有研究可知,随着计算机视觉和深度学习技术的不断进步,自动化岩石纹层图像分割技术已成为克服这些挑战的有效工具^[4]。

岩石纹层图像分割的核心目标是将岩石纹层从背景中有效分离,并精确地识别与统计其特征。这项任务的难点在于纹层间的紧密排列、细微的颜色变化以及潜在的噪声干扰。此外,考虑到岩石纹层的连续性和长度,分割算法必须能够处理高分辨率图像,同时保持对细节的高度敏感性。虽然近年来图像分割技术取得了显著的进展,但专门针对岩石纹层分割的算法仍存在着一系列难题亟待解决^[5]。

在现有文献中,多种图像分割技术已被探索用于岩石纹层图像的分析。传统的图像分割算法通常依赖于图像的像素信息,例如聚类算法。这些算法仅基于区域间像素的相似性进行分类,可能导致过分割或欠分割的问题^[6-7]。此外,一些基于阈值的纹层分割方法也被采用。王海涛等学者^[8]提出了一种图像分割技术,该技术选择对灰度阈值最敏感的孔隙度作为优选灰度分割阈值的检验参数。陈然^[9]开发了一种双门限阈值图像分割技术,该技术在爆破岩石图像分割中表现出最佳适应性,原因在于其针对所研究数据集图像的特征提出了先验的表征阈值,并受限于像素颜色分布。

鉴于传统图像分割算法的局限性以及其分割性能未能满足实际应用需求,研究者们开始探索深度学习技术在岩石纹层分割领域的应用潜力。Mask R-CNN^[10-11]是一种较早提出的深度学习图像分割方法,由一个特征提取模块和一个预测头组成。预测头通常用于处理特征提取骨干网络学习到的特征,并利用全连接网络层或其他张量操作技术来获取预测目标的边界框和像素点的类别等结果。这一过程使得模型能够与实际标签进行匹配,并在优化算法的作用下对参数进行更新。尽管 Mask R-CNN 在多个领域的图像分割任务中表现出较好的性能,但其通常使用较大的卷积核来分割较大的目标,这些目标往往具有异质性的轮廓和像素颜色分布,与岩石图像特征有显著不同,但是在岩石颗粒图像分割方面,则已证实 Mask R-CNN 相较于传统方法在分割性能上有了明显提升^[12]。

UNet 架构的图像分割方法在医学影像领域取得了可观成效^[13]。本质上,UNet 属于全卷积神经

网络模型,以其高效的图像特征提取能力而著称。医学图像通常为多通道、低分辨率,经由一些专业医疗设备采集,UNet 通过其胶囊形状的网络结构,在不同抽象层次的网络层间添加残差连接,有效保留了医学图像的关键信息。基于 UNet 的改进模型也被应用于岩石缝隙的分割。通过引入时空注意力机制和多尺度特征提取模块,这些改进模型成功地将 UNet 的应用扩展到了岩石图像分割的其他领域。

在图像分割算法的研究中,针对特定数据集的定制化改进是常见的做法。在特征提取的 Backbone 网络方面,研究者们引入了多种通道注意力机制或时空注意力机制,以促进特征的交互和融合。在预测头的设计上,也根据分割目标的特性进行了适应性调整。YOLOv8^[14]作为一种流行的端到端目标检测和实例分割模型,在多个数据集上展现出了卓越的性能和强大的泛化能力。因此,基于 YOLOv8 模型的岩石纹层实例分割应用,越来越受到学术界的关注。

在现有研究的基础上,本文提出了一种改进的 YOLOv8 算法,专门用于岩石纹层的分割。该算法不仅融合了岩石纹层的形状和分布特征,还集成了基于 Vision Transformer(ViT)的空间注意力机制,以增强模型对图像特征的表达能力,并提升分割性能。这种方法有效地解决了高密度集和颜色差异不明显的岩石纹层分割难题。实验结果证明,与常规分割方法及其他分割网络模型相比,该方法在分割准确度和鲁棒性方面均实现了显著提升。这一成果不仅推动了岩石纹层图像分割技术的进步,而且为地质研究和油气资源勘探领域提供了一种创新的工具。

1 YOLOv8 模型

YOLOv8 不但可以实现目标检测任务,还支持分类、分割、姿态估计等计算机视觉任务。YOLOv8 由特征提取骨干、特征融合和拼接颈部和检测头、分割头等部分组成,每个部分都可以灵活地进行修改以适应特定的任务。在特征骨干结构部分,基于其基本骨干可以设计不同的卷积核和结构,能够对局部细节和全局对比信息进行有效的提取。在分割头部分,可以设定不同的损失函数来针对性地提升特定图像下的模型效果。

图 1 展示了 YOLOv8 的完整网络结构,其主要由 Conv、c2f、SPPF、最近邻采样 UP 和 concat 拼接等模块组成。在左侧的特征提取骨干部分,大致由

Conv 模块和 c2f 组成,从原始分辨率 640×640 不断地下采样,形成类似特征金字塔的结构,最终经过 32 倍下采样之后得到 512 个大小为 2020 的特征图,再经过 SPPF 模块进行卷积和池化操作,综合不同区域的表征。在中间部分为特征融合和拼接部分,目的是将不同抽象层次的特征在通道维度进行拼接之后再进行卷积操作,以此将不同下采样倍数的特征进行融合,以达到对不同像素大小的目标进行有效的检测。在最右侧为 3 个解耦的检测头,分别对特征进行处理然后输出分类损失、检测框损失和分割损失等。

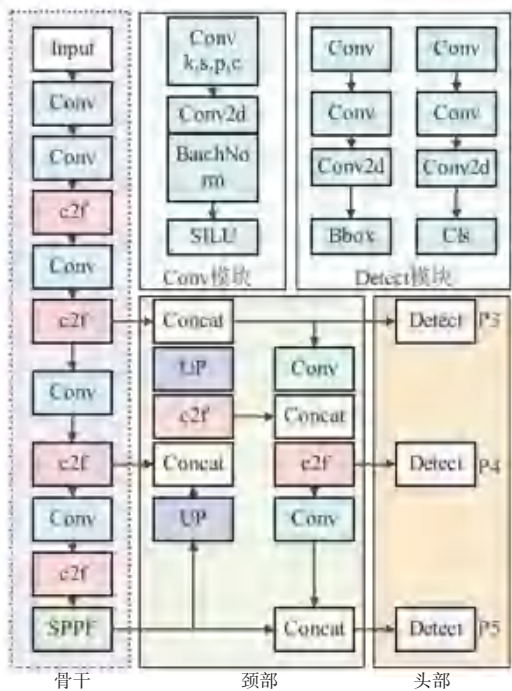


图 1 YOLOv8 网络结构
Fig. 1 YOLOv8 network structure

2 改进的 YOLOv8 网络

2.1 整体网络结构

鉴于岩石纹层图像纹层纹路的紧密排列,白色与灰色纹层难以区分等挑战,传统的 YOLOv8 网络在精确提取目标方面存在局限,常常导致图像的欠分割或过分割现象。考虑到岩石纹路的连续性和方向性特征,传统的正方形卷积核难以有效捕捉这些特征。为了提取纹层的连续性特征,需要在不同区域间进行特征的交换与融合。基于这些考虑,本文将自适应卷积核形状的卷积模块和选择性时空注意力机制整合到原生 YOLOv8 网络中,从而开发出一种新的岩石纹层目标分割模型。本文将 YOLOv8 的特征提取的 Backbone 中的 Conv 模块修改为自适应

形状卷积模块 AdpConv,在 P2 浅层网络进行替换;在 c2f 模块当中引入 Vit 用于融合不同区域的特征。

图 2 详细展示了改进型 YOLOv8 网络的结构。在特征提取骨干的 P2 和 P3 层之前,本研究引入了 AdpConv^[15] 和 c2f-Vit^[16-17] 模块,旨在高效地捕获纹路的方向性和连续性特征。在这些层中,由于下采样倍数较小,感受野能够恰好覆盖纹路的宽度,使得卷积核能够适应纹层纹路的形状,从而允许模型提取纹层纹路的整体特征。同时,c2f-Vit 模块能够整合不同区域的特征图,使得模型能够有效地区分连续性特征以及白色和灰色纹路。此外,颈部的 c2f-Vit 模块能够融合多种尺度的信息,使得模型能够有效地区分不同感受野特征图的连续和非连续纹层特征。

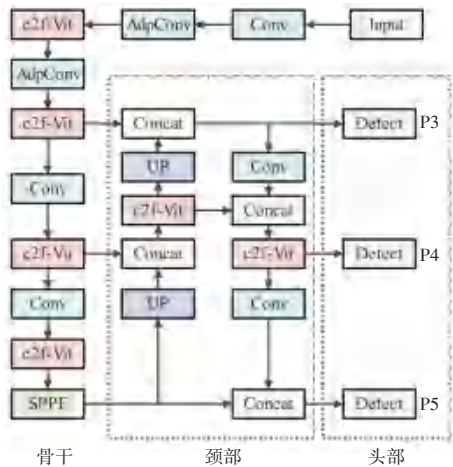


图 2 改进 YOLOv8 完整结构
Fig. 2 Improved YOLOv8 complete structure

2.2 AdpConv 结构

图 3 展示了所提出的 AdpConv 模块的结构。该模块以一个 3×3 的矩形卷积核作为基础,引入了 9 个可训练参数,这些参数用于记录每个卷积位置的偏移方向,从而形成一个 MASK。利用这个 MASK,可以构建出一个新的卷积核,其最大尺寸限制为 4×4 。这个 4×4 的卷积核能够根据不同的 MASK 形成多种不同的形状,以适应数据中的特征。随着梯度下降算法的迭代,这些形状将不断调整以适应数据,最终达到最优的卷积效果。

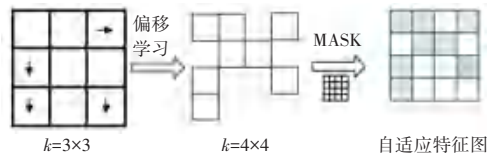


图 3 AdpConv 完整结构
Fig. 3 The complete structure of AdpConv

w_{kernel} 表示卷积核计算得到的该区域的权重, 计算公式如下:

$$w_{\text{kernel}} = F_{H \times W} \cdot W_{4 \times 4} \cdot \text{MASK} \quad (1)$$

2.3 c2f-Vit 结构

图4为c2f-Vit完整结构, 是对原始YOLOv8中c2f模块的改进, 具体为将c2f中的Conv2D卷积层替换为VitBlock网络层, 用于提取图像不同位置之间的特征并学习不同位置之间的权重系数, 提升模型的特征融合能力。具体来说, VitBlock将输入的特征图进行Split和Tokenization, 并利用Transformer编码器融合不同位置的特征图, 因此模型能够对输入岩石纹层图像的不同区域进行差异性学习, 使模型有效区分不同颜色的纹层以及对同一条纹层的连续性特征进行提取。

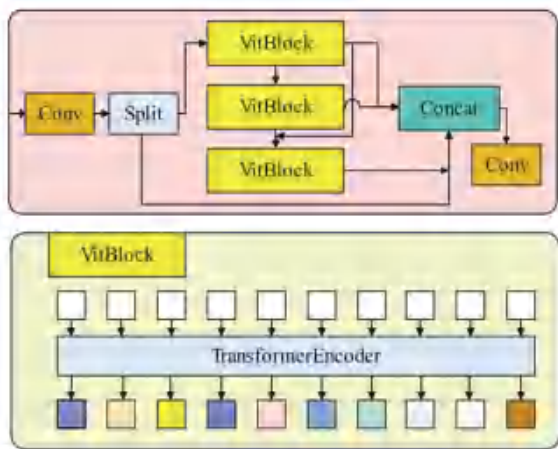


图4 C2f-Vit完整结构

Fig. 4 Complete structure of C2f-Vit

3 实验结果与分析

为了验证改进后的YOLOv8网络的有效性, 本文将该算法与分割网络YOLACT^[18]、UNet、YOLOv8n以及传统阈值分割进行对比, 使用的数据集是岩石纹层切片图像。此外, 为了分析本文算法中提出的改进方法对图像分割结果的影响, 还设计了消融实验来评估算法的有效性。

3.1 实验数据集

为了适应GPU的性能并提高训练效率, 研究从高分辨率岩心裂缝图像中裁剪出3 000张500×500像素的小图。然后, 使用Labelme工具对这些图像进行详细标注, 确保每个裂缝区域都用至少五边的多边形准确标记, 并将标注结果转换为PNG格式。图5展示了具体的原图与标注图, 而图6则提供了岩石纹层及其标注的示例。这样处理后, 得到了一个优化的数据集用于模型训练。



图5 岩石纹层图和标注图

Fig. 5 Rock layered diagram and annotation diagram

此外, 为增强模型对不同视距(感受野)的鲁棒性, 本研究还额外提供了总共200张分辨率为320×320(pixel×pixel)、640×640(pixel×pixel)、1 000×1 000(pixel×pixel)大小的图像。

3.2 环境与参数设置

本文的分割算法实验基于深度学习框架Pytorch2.2.0并结合编程语言Python在工作站上进行环境搭建, 工作站配置如下: 操作系统为Windows10系统, 处理器为AMD Ryzen 5 5600 G CPU, GPU为NVIDIA GeForce GTX1080Ti, 软件环境为CUDA11.2。本文训练时使用SGD优化器优化网络, 初始学率 $lr=0.001\ 667$, $momentum=0.937$, 权重衰减为0.000 5; 经过多次对比实验后, 将训练次数设定为200, 批量大小设置为4, 每次皆对模型参数文件进行保存, 最后在网络内使用最优的参数文件对测试集进行预测。

3.3 实验评价指标

本文将采用MASK的50%均像素精度($mAP50$)、95%均像素精度($mAP90$)、精度(Precision)、召回率(Recall)和PR曲线来对分割效果和网络模型进行评估。本文中选取分割出的纹层作为正例, 背景部分作为反例。在二分类问题中, 混淆矩阵的4个基本元素为:

(1) TP (True Positive)。真正例, 模型预测为正例, 实际是正例。

(2) FP (False Positive)。假正例, 模型预测为正例, 实际是反例。

(3) FN (False Negative)。假反例, 模型预测为反例, 实际是正例。

(4) TN (True Negative)。真反例, 模型预测为反例, 实际是反例。

3.4 改进YOLOv8模型有效性验证

本文针对数据集的特点, 提出了改进YOLOv8的图像分割方法。为了评估本文提出的改进型YOLOv8图像分割方法的有效性和适用性, 进行了

定性分析, 将其与当前主流的深度学习分割网络, 包括 YOLACT、UNet、YOLOv8n, 以及传统的阈值分割方法进行了比较。实验中, 所有参与比较的网络, 包括本文改进的网络, 均在同一数据集上进行训练。训练完成后, 所有网络以及传统阈值分割法均在同一测试集上进行评估, 以确保评价的一致性。分割结果的评估采用统一的标准, 相关结果汇总于表 1, 而图 6 则展示了各算法的分割效果可视化对比。

表 1 不同网络结果对比				
Table 1 Comparison of results from different networks %				
算法	Mask <i>mAP50</i>	Mask <i>mAP95</i>	Mask <i>Precision</i>	Mask <i>Recall</i>
阈值分割	—	—	10.60	12.5
YOLOv8n	26.6	10.1	37.30	35.4
YOLACT	30.5	10.7	37.54	35.8
UNet	29.2	9.8	35.54	34.1
本文算法	32.6	11.6	39.15	37.4

表 1 实验结果表明, 本文研发的算法在关键性能指标上, 如 Mask *mAP50* (32.6%)、Mask *mAP95* (11.6%)、Mask *Precision* (39.15%) 和 Mask *Recall* (37.4%), 均取得了显著的领先优势, 超越了其他现有算法。尽管 YOLOv8n 和 YOLACT 在特定指标上也展现了竞争力, 例如 YOLOv8n 在 Mask *mAP50* 上达到了 26.6%, YOLACT 的 Mask *mAP50* 则为 30.5%, 但总体而言, 两者的性能未能超越本研究提出的算法。UNet 虽然在 Mask *Precision* 和 Mask *Recall* 上保持了一定的平衡, 但在 *mAP* 指标上的表现仍有提升空间。相比之下, 阈值分割方法由于其固有的局限性, 在所有评估指标上的表现均不尽如人意。这些发现证实了本研究算法通过改进模型架构和训练策略, 显著提高了实例分割任务的精确度和稳定性, 为该领域的发展提供了新的视角和解决方案。

表 2 不同模块的消融实验				
Table 2 Ablation experiments of different modules				
算法	Mask <i>mAP50</i>	Mask <i>mAP95</i>	Mask <i>Precision</i>	Mask <i>Recall</i>
YOLOv8n	26.6	10.1	37.30	35.4
YOLOv8n+AdpConv	32.2	11.3	39.10	37.3
YOLOv8n+VitBlock	32.1	11.5	38.80	37.1
本文算法	32.6	11.6	39.15	37.4

分析表 2 中数据, 可以观察到 2 种改进模块均对网络性能产生了积极影响。具体而言, VitBlock

图 6 直观地比较了 6 种不同的分割算法在同一张带有垂直条纹图像上的效果。结果显示, 本文算法、YOLOv8n 和 YOLACT 均能够成功地识别并突出图像中的垂直条纹, 其中本文算法的表现尤为突出。尽管 UNet 在识别条纹方面稍显不足, 但仍能部分地识别出条纹的存在。相比之下, 阈值分割方法由于引入了过多的噪声, 未能有效地捕捉到清晰的纹层轮廓。基于这些观察, 可以推断, 在当前的实验设置下, 本文算法提供了最佳的分割效果, 可能是处理此类图像分割任务的最有效方法。

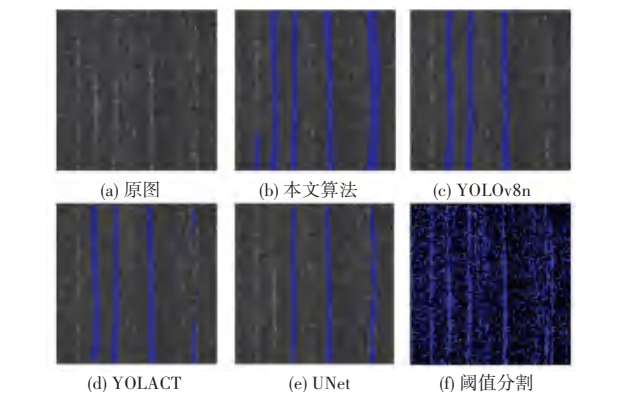


图 6 对比结果可视化
Fig. 6 Visualization of comparison results

3.5 不同模块的消融实验

为了评估本文提出的改进方法对岩石纹层分割效果的具体影响, 设计了一系列消融实验, 皆在岩石纹层数据集上进行评估。这些实验分为 4 个对照组: 第 1 组采用未经修改的 YOLOv8n 网络作为基准; 第 2 组在 YOLOv8n 网络的基础上, 将浅层卷积模块 (Conv) 替换为自适应卷积模块 (AdpConv); 第 3 组在 YOLOv8n 网络中集成了视觉 Transformer (ViT) 至 c2f 模块; 第 4 组则同时集成了 AdpConv 和 ViTBlock 模块 (本文算法)。所有实验均采用统一的训练参数以保证结果的可比性。具体的消融实验结果详见表 2。

表 2 不同模块的消融实验				
Table 2 Ablation experiments of different modules				
算法	Mask <i>mAP50</i>	Mask <i>mAP95</i>	Mask <i>Precision</i>	Mask <i>Recall</i>
YOLOv8n	26.6	10.1	37.30	35.4
YOLOv8n+AdpConv	32.2	11.3	39.10	37.3
YOLOv8n+VitBlock	32.1	11.5	38.80	37.1
本文算法	32.6	11.6	39.15	37.4

模块通过增强模型对不同图像位置及特征图特征之间的建模能力, 显著提升了分割的准确度。

AdpConv 模块通过自适应调整卷积核的形状和尺寸,增强了模型对多尺度和多形状纹层的检测能力,特别是在区分颜色相近的纹层方面表现出色。当这2个模块被集成到网络中时,则共同提升了网络的特征提取精度和鲁棒性,从而证明了改进后的网络在岩石纹层分割任务中具有优越的性能。

4 结束语

本研究致力于解决岩石纹层图像分割中面临的挑战,包括传统阈值分割方法和深度学习方法在精确度上的不足,以及由于背景颜色复杂性和纹层颜色相似性导致的岩石纹层图像细节信息提取困难。为此,研究提出了一种基于改进型YOLOv8的岩石纹层图像纹层分割算法。该算法通过集成视觉Transformer(ViT)增强了网络对输入图像和特征图全局特征的表达能力,并利用可变形卷积核提升了网络对多尺度特征的提取能力,进而提高了分割的准确性。实验对比和验证表明,改进后的网络显著增强了对相似颜色纹层的区分能力,并对图像中的不同层次形状、不同方向的纹层细节信息、边界检测和区域分割均实现了有效提升。

然而,本文的工作还存在一些局限性和可以改进的方面。尽管本研究在岩石纹层图像领域进行了广泛的实验验证,但算法对于其他类型图像数据集的适用性仍需做进一步探究。此外,算法的计算效率有待提升,特别是在处理大规模图像数据集时,以增强其在实际应用场景中的适用性。未来的工作将致力于改进和优化岩石纹层图像分割算法,同时探索更多先进技术以提高分割性能。

参考文献

- [1] 蒋宜勤,周基贤,冯俊羲,等. DBSCAN 算法与数学形态学在岩石薄片图像分割中的应用[J]. 微型机与应用,2016,35(17): 39-41.
- [2] 张中亚. 砂岩薄片图像分割与识别研究[D]. 合肥:中国科学技术大学,2020.
- [3] 吕锦涛. 基于超像素与半监督的岩石图像分割与识别[J]. 工程科学与技术,2023,55(2):171-183.
- [4] 舒小锋,吴晓红,卿粼波,等. 基于改进 U2Net 的岩石薄片图像分割[J]. 计算机系统应用,2024,33(2):159-165.
- [5] 江佳霖. 岩石薄片图像分割识别方法的研究[D]. 武汉:长江大学,2024.
- [6] 黄玲玲. 基于聚类的岩石薄片图像分割方法研究[D]. 焦作:河南理工大学,2022.
- [7] 刘连宏. 密度聚类算法在岩石图像中的研究与应用[D]. 西安:西安石油大学,2021.
- [8] 王海涛,杨叶,耿尊博. 图像分割中灰度阈值对岩石物理属性影响的研究[J]. 地球物理学进展,2018,33(6):2456-2461.
- [9] 陈然. 基于双门限阈值的爆堆岩块图像分割技术及图像识别系统开发[D]. 衡阳:南华大学,2020.
- [10] KABIR H, LEE H S. Mask R-CNN-based stone detection and segmentation for underground pipeline exploration robots [J]. Applied Sciences, 2024, 14(9): 3752.
- [11] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 2961-2969.
- [12] 司晨冉,王仁超,邸阔,等. 一种基于 Mask R-CNN 和分水岭算法的岩石颗粒图像分割方法[J]. 水电能源科学,2020,38(11): 129-132.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional Networks for Biomedical Image Segmentation [C]//18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). Cham: Springer, 2015: 234-241.
- [14] JOCHER G, QIU Jing, CHAURASIA A. Ultralytics YOLO[EB/OL]. [2023]. <https://github.com/ultralytics/ultralytics>.
- [15] ZHANG Xin, SONG Yingze, SONG Tingting, et al. AKConv: Convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters [J]. arXiv preprint arXiv, 2311. 11587, 2023.
- [16] HAN Kai, WANG Yunhe, CHEN Hanting, et al. A survey on vision Transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 87-110.
- [17] KHAN S, NASEER M, HAYAT M, et al. Transformers in vision; A survey[J]. arXiv preprint arXiv, 2101.01169, 2021.
- [18] BOLYA D, ZHOU Chong, XIAO Fanyi, et al. Yolact: Real-time instance segmentation [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2019: 9157-9166.
- [19] LI Jialun, ZHANG Li, WU Zhongchen, et al. Autonomous martian rock image classification based on Transfer deep learning methods[J]. Earth Science Informatics, 2020, 13: 951-963.
- [20] CHEN Cheng, SEO H, JUN C H, et al. A potential crack region method to detect crack using image processing of multiple thresholding[J]. Signal, Image and Video Processing, 2022, 16(6): 1673-1681.