

基于 DAF-STDC 实时模型的岩屑图像语义分割识别^①



潘显珊¹, 王正勇¹, 罗彬彬², 滕奇志¹, 何小海¹

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 610024)

通信作者: 王正勇, E-mail: 690728634@sina.com

摘要: 岩屑识别是地质勘探录井工作中的重要手段, 为了解决传统人工岩性识别效率低下和常见深度学习网络推理速度慢计算量大等问题, 本文借助于表现较好的短时密集串联网络模型, 提出了一种实时岩屑图像语义分割网络 DAF-STDC. 该网络使用空洞卷积在提取特征过程中保持分辨率, 利用注意力机制帮助模型获取特征图中全局信息, 从而细化岩屑颗粒边缘信息, 使用特征融合模块加强低阶细节特征与高阶语义特征的融合程度, 帮助特征表达. 经实验证明, 改进后的网络模型在精度方面有了大幅提升, 在岩屑录井工作中收集到的 6 类岩屑图像, 制成的数据集 RC_Dataset 上, 平均交并比达到 83.12%, 在保持参数数量的同时, 推理速度和分割精度有了明显提升, 为岩屑录井数字化提供了有效参考.

关键词: 语义分割; 岩屑图像; 岩性识别; 注意力机制; 深度学习

引用格式: 潘显珊, 王正勇, 罗彬彬, 滕奇志, 何小海. 基于 DAF-STDC 实时模型的岩屑图像语义分割识别. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9716.html>

Semantic Segmentation and Recognition of Rock Debris Image Based on Real-time Model DAF-STDC

PAN Xian-Shan¹, WANG Zheng-Yong¹, LUO Bin-Bin², TENG Qi-Zhi¹, HE Xiao-Hai¹

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

²(Chengdu Xitu Technology Co. Ltd., Chengdu 610024, China)

Abstract: Rock debris recognition is an important tool in geological exploration and logging. To improve the efficiency of traditional manual lithology identification and overcome the challenges of slow inference and high computational complexity in common deep learning networks, this study proposes DAF-STDC, a real-time semantic segmentation network for rock debris images based on a well-performing STDC network model. The network uses dilated convolution to maintain resolution while extracting features and utilizes an attention mechanism to help the model acquire global information from the feature map, thus refining the edge information of rock debris particles. It also uses a feature fusion module to enhance the fusion of low-level detail features and high-level semantic features, improving feature representation. Experiments have proved that the improved network model significantly enhances accuracy. The mean intersection over union of DAF-STDC reaches 83.12% on the RC_Dataset which consists of six types of rock debris images collected from exploratory wells. While maintaining the number of parameters, DAF-STDC significantly improves its inference speed and segmentation accuracy, providing an effective reference for the digitization of rock debris logging.

Key words: semantic segmentation; rock debris image; lithology identification; attention mechanism; deep learning

^① 基金项目: 国家自然科学基金 (62071315)

收稿时间: 2024-05-23; 修改时间: 2024-06-17; 采用时间: 2024-07-11; csa 在线出版时间: 2024-10-31

1 引言

油气资源作为社会发展的重要物质基础,人们对于油气资源的需求和要求,随着科技的发展和社会的进步,不断提高.油气勘探和开采技术也获得了广泛关注,掌握高效的油气勘探和开采技术,不仅可以减少人力物力的损耗,还可以提高油气开采的效率.在当前的油气勘探过程中,岩屑录井是岩性鉴定和油气评估最及时和直接的基本技术.钻井时,按照特定的取样间隔和延迟时间进行操作,连续收集和记录被钻头击碎并被钻井液携带上来的岩屑,并对岩屑样本进行系统地描述,保存和识别,通过对岩屑的判断和分析来确定地下岩石的性质,划分岩性和地层,从而判断和评估地层所含油气的情况^[1].传统的岩屑识别方法是人工鉴定识别,由具有专业地质知识的鉴别人员通过肉眼,嗅觉或触觉来进行鉴定,有时也会借用扫描仪或光谱仪等仪器进行辅助.传统的人工岩屑鉴定方法存在着效率低,受主观影响因素较大,对专业人员要求高等问题,无法满足当今的岩屑识别鉴定工业应用的需求.为解决这些问题,岩屑识别和鉴定向着计算机自动化方向发展^[2].

随着计算机的普及和机器学习的发展,各种效果较好的图像分类算法^[3-7]被提出并取得显著的效果,在岩屑样本的识别领域也得到了广泛的应用,许振浩等人^[8]结合深度学习网络 ResNet,提出了一种基于岩石图像迁移学习的岩性智能识别方法,可用于岩石岩性的快速自动识别.谭永健等人^[9]利用 Xception 神经网络,结合迁移学习,实现了岩石图像的自动化分类方法.Wang^[10]基于深度学习方法 VGG16,构建了测井工作中岩屑岩性的智能识别模型,可实现岩石图像的自动识别.Wu 等人^[11]构建了一种基于特征融合的岩屑图像识别模型,将颜色校正后的图像进行特征提取,可以有效提高模型识别准确率.Lu 等人^[12]将岩屑图像的颜色和深度 (RGB-D) 信息进行结合,提出了一种图像融合算法,用于分割多尺度颜色图像.由于岩屑样本可能具有复杂的纹理,颜色和形态特征,需要更为复杂的特征提取和表示方法,基于图像分类算法在细小样本的岩屑识别任务中表现不佳.

对于岩屑录井采集到的岩屑样本通常呈粉末状,颗粒细小,岩屑颗粒边缘以及纹理特征差异较细小,并且成像时一张图片中会含有多种不同岩性的岩屑,因

此逐像素分类的语义分割网络^[13-15]在岩屑识别和鉴定中得到广泛应用.Zhou 等人^[16]提出了一种解决渣土图像的自动分割和估计碎石的大小和形状方法,该方法由具有多尺度输入和边输出的双 UNet (MSD-UNet) 和后处理算法组成.Zhang 等人^[17]提出了一种基于深度学习的岩屑分割方法,利用语义分割网络 MobileNetV3 UNet (M3-UNet) 对岩屑图像的含油区域进行自动分割,进行岩屑含油量的定时定量分析.覃本学等人^[18]针对岩屑图像的特征设计了语义分割网络 Debseg-Net 用于岩屑图像得到岩性鉴定,对于复杂状况下的岩性鉴定泛化性不足.孙杰等人^[19]设计了一种用于岩屑图像实时语义分割的神经网络模型,对于岩屑图像密集的分割仍然存在不足.董帅^[20]提出了一种改进的 MultiRes-UNet++语义分割算法,实现了岩屑图像的自动识别和鉴定,但是对于边界模糊和贴合紧密纹理色泽相近的岩屑群体没有办法做到准确分割.Ogunkeye 等人^[21]利用 Detectron2 框架中最先进的 Mask R-CNN 深度学习算法,捕捉岩石的全局特征,从而仅凭岩石图像就能更高效地识别岩石颗粒.Huo 等人^[22]基于 UNet++ 分割模型的 MultiRes 模块概,提出了基于深度学习的改进端到端 UNet++ 图像语义分割模型,实现岩屑图像的自动分割.

因此,针对岩屑录井工作中的效率低,易受干扰等问题,且为了满足其实际的工作的需求,本文将深度学习主流语义分割网络应用于此项工作中,提出了一种实时岩屑图像分割和识别方法:DAF-STDC (real-time debris segmentation with attention module and feature fusion based on STDC).实验结果表明,本方法识别准确率高,可用于岩屑录井工程中的岩性识别.

2 本文模型

2.1 DAF-STDC 模型框架

由于岩屑识别任务对效率要求较高,为了满足实际工业应用的需求,需要一个精度良好,计算量较小的模型.因此,本文经过多次模型实验,借助表现较好的短时密集串联网络 (short-term dense concatenate network, STDC network)^[23],搭建网络模型与多个语义分割模型在自制岩屑图像数据集 RC_Dataset 上进行对比实验,最终得到在精度和计算量上更具优势的 DAF-STDC 网络模型,其结构如图 1 所示.

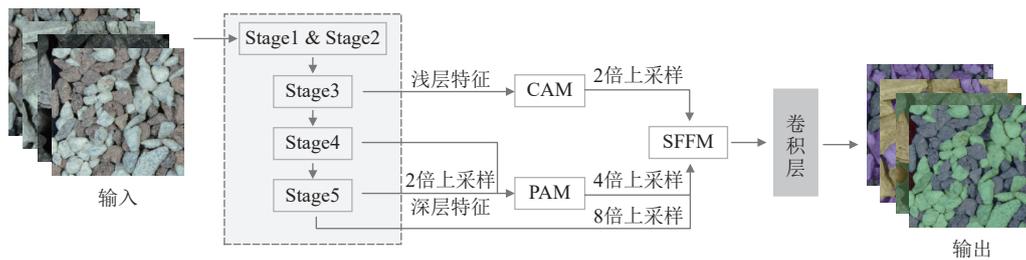


图1 DAF-STDC 网络模型结构

DAF-STDC 使用传统语义分割网络的编解码结构, 编码器部分负责特征提取, 是使用改进短时密集串联模块 (short-term dense concatenate module, STDC module) 构建的网络, 该网络共有 5 个 Stage, 为了减少网络计算量, Stage1 和 Stage2 中只包含 1 个卷积层, 其余 Stage 中包含若干个改进后的 STDC 模块, 其中第 1 个 STDC 模块包含步长为 2 下采样操作, 其余 STDC 模块保持特征图尺寸不变. 该模块通过少量参数就可获得不同大小的感受野以及多尺度信息, 大幅提升了岩屑分割任务中的网络性能. 解码器部分使用双通道注意力机制^[24], 包括位置注意力模块 (position attention module, PAM) 和通道注意力模块 (channel attention module, CAM) 来细化特征表达, 通过计算不同位置和通道之间的相似度和注意力权重, 来增强模型对图像中不同位置的关注, 这使得模型能够更好地获取到岩屑图像的全局上下文信息, 从而帮助计算机视觉任务性能的提升. 同时, 设计了专门的特征融合模块, 旨在利用不同层次的特征图信息来提高模型的分割性能. 最后通过卷积层生成最终的预测图.

2.2 改进的 STDC 模块

STDC 模块的设计和提出是为了解决语义分割算法的复杂度和耗时问题, STDC 模块可以获取到不同的可变化的感受野, 并且使用较少的参数量提取到输入图像的多尺度特征, 模块化后便于在各个网络中集成, 其模块结构如图 2 所示.

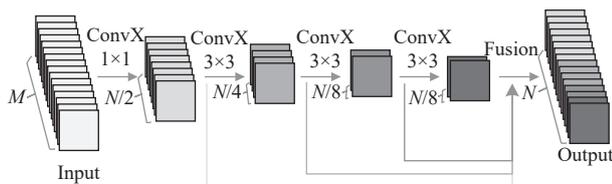


图2 原始 STDC 模块

使用 LN 替换图中的 BN, 在单个样本的特征维度上进行归一化, 减小特征差异, 提高模型的稳定性和泛

化能力, 进而降低过拟合的风险. 并且为了提升感受野, 使用空洞卷积替换内的普通卷积操作. 在 STDC 模块基础上改进得到效果更好的特征提取模块, 如图 3 所示. 图中的 ConvX 表示“卷积+LN+ReLU”操作, Dilated ConvX 表示“空洞卷积+LN+ReLU”操作, 其中输入特征通道数为 M , 最后的输出特征通道数为 N , 除了最后一个卷积模块以外, 该模块内第 i 个卷积模块的输出通道数为 N , 最后一个模块的输出特征通道数与倒数第 2 个模块保持一致. 相较于比较传统的骨干网络来说, 改进 STDC 模块中更关注高阶语义信息的深层的特征通道数量少, 浅层的特征通道数量多. 改进 STDC 模块最终的输出为各模块输出特征的融合, 其计算结果如下:

$$X_{out} = F(x_1, x_2, \dots, x_n) \quad (1)$$

其中, F 表示特征融合函数, x_1, x_2, \dots, x_n 表示 n 个卷积模块的输出, X_{out} 表示改进 STDC 模块的输出. 使用串联 (concatenation) 操作融合 n 个卷积模块的特征, 串联融合方法直接进行通道合并进行系列特征融合, 简单来说就是直接将两个特征进行连接, 若两个输入特征 $F1$ 和 $F2$ 的维数分别为 x 和 y , 输出特征的维数为 $x+y$.

在岩屑语义分割任务中, 针对粉末状岩屑的分割, 具有岩屑粒状小, 形状不规则, 边缘不清晰等特点, 因此分割时全局特征的获取尤为重要. 感受野的大小通常决定了该层卷积模块的上下文信息和全局特征提取的能力, 感受野越大, 从特征图上的点看到输入图像的区域就越大, 特征图所蕴含的语义信息就会更为全面, 层次更高, 从而提升分割结果的准确性. 为了提高模型对全局特征的提取能力, 扩大感受野, 本文使用空洞卷积替代普通卷积如图 4 所示, 其中 (a) 为空洞卷积示意图, (b) 扩张率为 1 的卷积核, (c) 扩张率为 2 的卷积核, (d) 扩张率为 5 的卷积核. 因为在一般网络中为了增加感受野总是要进行降采样导致分辨率降低, 使用空洞

卷积可以扩大感受野同时避免分辨率降低,做到精确定位目标和检测分割大目标.为了避免空洞卷积的网格效应问题,扩张率参数为别设置为1、2、2、5,当设

置不同扩张率参数时,卷积的感受野也会不同,即获取了多尺度上下文信息.但空洞卷积的使用势必会引起算法整体计算量的增加,所以扩张率参数需要权衡.

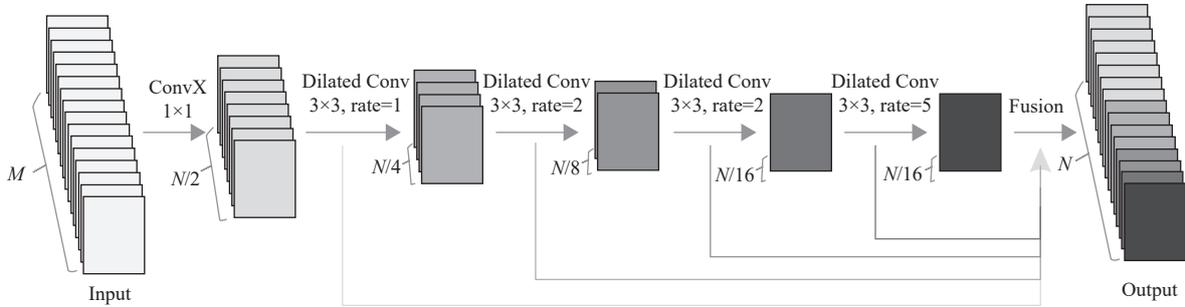


图3 改进后 STDC 模块

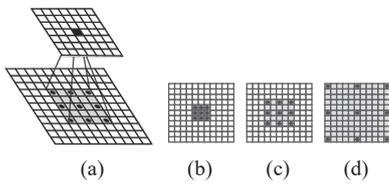


图4 空洞卷积图

2.3 注意力机制

在岩屑图像语义分割算法中,岩屑图像中通常包含了不同尺度的岩屑对象,背景图像以及岩屑的各种纹理和结构,如沉积岩、碎屑、矿物、孔隙等,岩屑图像中的类别则是通过纹理和颜色的一致性进行区分,有时还存在噪声和不完整信息的干扰.通过注意力机制,模型可以有选择性地关注不同区域的特定纹理和结构,从而更好地区分不同的地质类别.通过在模型中加入注意力模块 PAM 和 CAM,其结构如图 5 所示,分别模拟位置和通道维度中的语义相互依赖性,帮助模型更好地处理岩屑图像中的具有空间结构或多通道特征的数据.

为了增强特征的表达能力,使用位置注意力模块关注模型不同位置的特征,并将这些特征融合到全局特征中,有助于网络更准确地识别分割目标岩屑. PAM 利用任意两点特征之间的关联,来帮助各自特征的表达,其结构如图 5(a) 所示.经过骨干网络后提取的特征图 $A \in \mathbb{R}^{C \times H \times W}$ 在经过一次卷积操作后,获得新特征图 B 和 C , $\{B, C\} \in \mathbb{R}^{C \times H \times W}$ 然后将 B 和 C 都重塑为 $C \times N$, 其中 $N = H \times W$. 将 B 进行转置后与 C 相乘,获得的结果再经过 Softmax 操作,得到大小为 $N \times N$ 的特征图 $S \in \mathbb{R}^{N \times N}$, 计算方式如下所示:

$$s_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^N \exp(B_i \cdot C_j)} \quad (2)$$

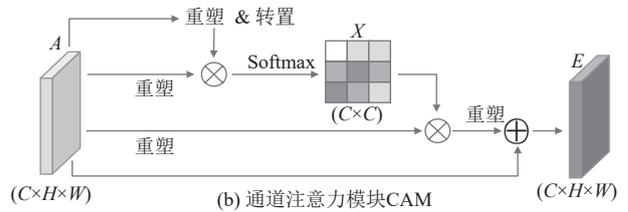
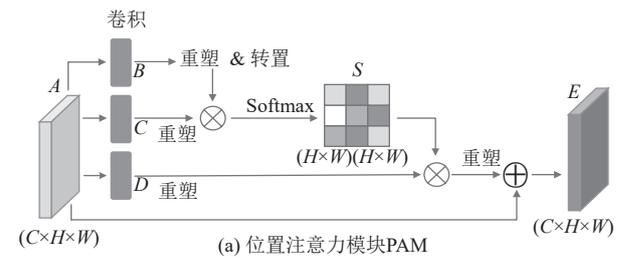


图5 双通道注意力模块

S 中每行的和为 1, s_{ji} 可理解为第 i 位置像素对第 j 位置像素的权重,即所有像素 i 对某固定像素 j 的权重和为 1; B_i 为 B 中第 i 列的元素; C_j 为 C 中第 j 列的元素.同时,将 A 再次经过卷积操作得到特征图 $D \in \mathbb{R}^{C \times H \times W}$, 同样的重塑为 $C \times N$ 的大小.将其与 S 的转置相乘,得到大小为 $C \times N$ 的结果图,再重塑回 $C \times H \times W$ 大小,乘以系数 α .最后将其与 A 相加,获得融合了位置信息的最终特征图 E , 计算公式如下:

$$E_j = \alpha \sum_{i=1}^N (s_{ji} D_i) + A_j \quad (3)$$

其中, E_j 为 PAM 的输出特征; α 是一个需迭代更新的权重参数, 初始值为 0; D_i 为 D 中第 i 列的元素; A_j 为 A 中第 j 列的元素.

通道注意力模块通过建模通道之间的关联, 对特征图的不同通道进行加权来实现对特征的有选择性增强, 提高特征的区分度, 从而增强通道下特定语义响应能力, 其结构如图 5(b) 所示. 与 PAM 模块类似, 只是对特征图 A 没有进行卷积操作, 而是直接对 A 进行操作. 同样的, 将 A 重塑为 $C \times N$ 的大小, 得到特征图 B , 再将 B 与其转置相乘, 再经过 Softmax 操作, 得到大小为 $C \times C$ 的特征图 X . 将 X 的转置与 B 相乘, 再重塑回 $C \times H \times W$ 的大小, 乘以系数 β , 记为 D . 将 A 与 D 相加, 则得到融合了通道信息的最终特征图 E . 计算公式如下所示:

$$x_{ji} = \frac{\exp(A_i \cdot A_j)}{\sum_{i=1}^C \exp(A_i \cdot A_j)} \quad (4)$$

$$E_j = \beta \sum_{i=1}^C (x_{ji} A_i) + A_j \quad (5)$$

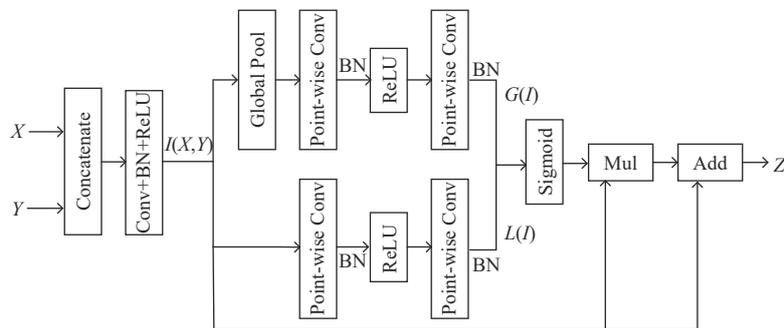


图 6 特征融合模块

给定不同层次的特征, 首先在通道维度 concat 输入的特征图 X 和 Y , 然后经过 BN 层进行归一化来平衡特征的尺度, 得到初始融合特征 $I(X, Y)$. 接下来, 使用归一化后的融合特征, 将其传入两个不同的分支来提取特征, 其中一个分支通过逐点卷积直接提取局部特征 $L(I)$. 计算公式如下:

$$L(I) = \beta(\text{PWConv}_2(\delta(\beta(\text{PWConv}_1(I)))))) \quad (6)$$

其中, PWConv 表示逐点卷积, β 表示 BN 操作, δ 表示 ReLU 操作.

另一个分支将特征池化为一个特征向量, 并计算权重向量, 该权重向量可以对特征重新加权, 这相当于全局特征选择和组合, 从而提取全局特征 $G(I)$, 计算公

其中, x_{ji} 可理解为第 i 通道像素对第 j 通道像素的权重, E_j 为 CAM 的输出特征; β 是一个需迭代更新的权重参数, 初始值为 0; A_i 为 A 中第 i 列的元素.

将两个注意力模块输出的结果进行特征融合, 从而获得更深入地捕获全局依赖关系的特征, 也就是最终的特征, 对像素点进行分类.

2.4 特征融合

特征融合是一种改善模型对图像的语义理解和像素级别分类任务的技术, 其目的是将不同层次或来源的特征信息结合在一起, 利用不同特征图的信息来提高分割模型的性能. 特征融合通常通过不同的操作实现, 例如级联、相加、相乘、池化、上采样等, 具体取决于网络架构和任务要求. 对于岩屑图像的语义分割任务来说, 特征融合使用简单的级联和相加的方式也能取得一定的效果, 但不是最佳的方式, 为了帮助模型更好的理解特征图中的语义和上下文信息, 本文提出了专门的特征融合模块 (special feature fusion module, SFFM), 其结构如图 6 所示.

式相较于 $L(I)$ 多了全局平均池化. 对两个分支的输出结果进行 Sigmoid 运算, 可得到融合权重, 再将初始融合特征与融合权重相乘并进行加权求和, 则得到最终的融合特征 Z . SFFM 的计算公式如下:

$$Z = (\sigma(L(I) + G(I)) \otimes I) \oplus I \quad (7)$$

其中, σ 表示 Sigmoid 操作, \otimes 表示逐元素相乘, \oplus 表示广播加法.

3 实验及结果分析

3.1 数据集制作

本文制作的岩屑数据集 RC_Dataset 初始样本是由合作的石油公司提供, 通过将现场采集到的岩屑经

过清洗整理后得到的大面积粗选岩屑图像,其中所包含的岩屑种类总共有6种,分别为页岩、泥岩、石膏、白云岩、砂岩和石灰岩,具体岩屑种类如图7.

目前,该数据集中共有6450张可用的岩屑图像数据,图像分辨率为512×512.为了后续的岩屑识别工作,

以及模型训练和模型性能优化,将数据集进行划分,为训练、验证和测试3种用途,其中训练集共有4266张岩屑图像,验证集有885张岩屑图像,测试集有795张岩屑图像,如图8所示,展示了部分数据集RC_Dataset图像.

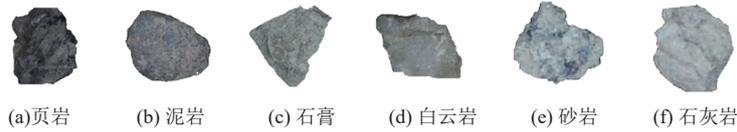


图7 岩屑样本展示

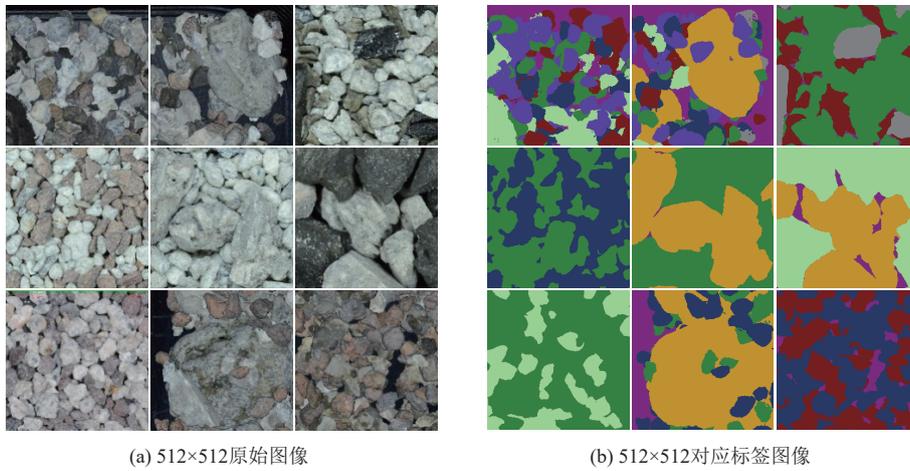


图8 语义数据集

3.2 实验环境和相关参数

本文实验在Linux操作系统上进行,开发环境为PyTorch深度学习框架,CPU型号是Intel i7-9700,GPU型号是NVIDIA GeForce RTX 1080 Ti,内存为16 GB.

在网络训练过程中使用Adam算法有效地优化损失函数,动量为0.9,权值衰减为0.005,batch size设置为4,学习率为0.0001,最大迭代次数是80000次.每个模型训练都是相同的配置条件.

利用损失函数分析模型预测输出结果 \hat{y} 与数据集原始标签 y 的误差,模型训练过程中使用交叉熵(cross entropy)损失函数计算公式如下:

$$l = - \sum_{i=1}^N y_i \log(p_i) \quad (8)$$

其中, y 和 p 表示为真实预测和期望, N 表示为分类标签数.

3.3 评价指标

为了更精准地比较网络模型的分割效果,本文采

用语义分割标准度量方法包括平均像素准确率(mean pixel accuracy, mPA),交并比(intersection over union, IoU)和平均交并比(mean intersection over union, $mIoU$).

mPA 计算分类结果中预测正确的类别的像素数目的比例的平均值.具体计算公式如下:

$$mPA = \frac{1}{K+1} \times \sum_{i=0}^K \frac{P_{ii}}{\sum_{j=0}^K P_{ij}} \quad (9)$$

IoU 计算模型预测结果和真实标签结果的交集占这两个结果并集的比例,具体计算公式如下:

$$IoU = \frac{P_{ii}}{\sum_{j=0}^K P_{ij} + \sum_{j=0}^K P_{ji} - P_{ii}} \quad (10)$$

$mIoU$ 计算所有类别的 IoU 的平均值,具体计算公式如下:

$$mIoU = \frac{1}{K+1} \times \sum_{i=0}^K \frac{P_{ii}}{\sum_{j=0}^K P_{ij} + \sum_{j=0}^K P_{ji} - P_{ii}} \quad (11)$$

其中, K 为分类标签数, $K+1$ 为分类标签数加背景数, P_{ii} 为分类正确像素比例, 表示岩屑图像中像素点实际为 i 类同时被预测为 i 类, P_{ij} 为分类错误的像素比例, 表示岩屑图像中像素点实际为 i 类同时被预测为 k 类。

3.4 实验结果与分析

为了测试本文模型的改进的有效性和性能优势, 本文在自制岩屑数据集 RC_Dataset 上进行大量消融对比实验. 通过表 1 可以发现, 通过在模型中改进 STDC 特征提取模块, 加入注意力机制和特征融合模块, 本文模型的平均交并比相较于原始 STDC 网络有了 7.51% 的提升, 在精度和效率上都具有较好的效果. 实验结果表明, 本文的优化和改进方法是有效的。

表 1 改进策略性能比较

网络模型	$mIoU$ (%)	Params (M)	FLOPs (G)
STDC网络	77.61	8.27	8.44
ResNet+注意力机制+卷积	80.12	47.46	211
STDC模块+注意力机制+卷积	81.54	11.37	20.85
STDC模块+注意力机制+特征融合+卷积 (本文模型)	83.12	18.47	43.16

表 2 测试集下不同模型 IoU , mPA 和 $mIoU$ (%)

类别	U-Net	DANet	BiSeNet-V1	MobileNetV3	STDC	SCTNet	BiSeNet_SPARM_S-iAFF	Ours
背景	46.88	54.88	49.50	42.54	53.52	53.55	57.39	56.63
白云岩	89.06	90.74	90.15	88.69	90.65	92.98	89.06	91.75
泥岩	70.86	88.91	85.99	84.25	88.13	86.33	69.34	89.82
砂岩	66.18	91.57	87.26	89.34	91.56	65.22	86.36	93.24
石膏	63.08	85.36	78.47	79.45	85.42	79.51	57.82	88.30
石灰岩	88.22	93.45	90.94	91.62	92.62	70.28	88.07	94.10
页岩	50.17	55.92	28.60	21.01	41.33	80.98	57.63	67.99
mPA	84.41	90.61	82.01	80.00	85.55	84.19	79.38	94.28
$mIoU$	70.93	80.12	72.99	70.99	75.61	75.55	72.24	83.12

由图 9 可知, 本文提出的 DAF-STDC 整体分割效果优于其余模型, 在图像中使用红色矩形框选出来的颗粒可以明显看出, 其中较大的完整岩屑颗粒都被分割出来, 并且具有较好的岩屑边缘分割效果. 无论是在像素分割准确率还是岩屑边缘的分割, 本文构建的 DAF-STDC 模型具有明显优势, 说明本文的模型具备较强的特征提取能力, 并且注重特征的表达和融合, 对全局信息进行充分利用, 表现出优秀性能. 在面对复杂场景下的岩屑样本时, 同样可以表现出优秀的分割

能力. 本文与图像语义分割中 7 种算法 U-Net^[25], DANet^[24], BiSeNet-V1^[26], MobileNetV3^[27], STDC^[23], SCTNet^[28], 以及岩屑实时网络模型 BiSeNet_SPARM_S-iAFF^[19] 进行对比实验, 其中 U-Net 和 DANet 为高精度模型, BiSeNet-V1, MobileNetV3, STDC 和 SCTNet 为高实时性模型, BiSeNet_SPARM_S-iAFF 为岩屑分割模型. 表 2 实验对比了使用不同模型训练, 在测试集上进行模型预测后的不同类别的 IoU , 以及总的 mPA 和 $mIoU$ 结果, 由表 2 可知 DAF-STDC 模型的 IoU 相较于其余的模型最优结果分别提高了 -0.76%, -1.23%, 0.91%, 1.67%, 2.88%, 0.65%, -12.99%. 在 mPA 度量方法上, DAF-STDC 相较于其他网络模型分别提高了 9.87%, 3.67%, 12.27%, 14.28%, 8.73%, 10.09%, 14.9%. 对于 $mIoU$ 评价指标, DAF-STDC 则分别提高了 12.19%, 3%, 10.13%, 12.13%, 7.51%, 7.57%, 10.88%. 由对比结果可以看出, 在各个评价指标上, 本文构建的 DAF-STDC 网络模型在岩屑分割时的总体表现优于其他对比网络模型, 具有较好的分割效果。

为了进一步直观地展示模型的分割效果, 将模型训练后的最好结果对测试集进行预测, 各个模型的分割结果如图 9 所示, 其中 U-Net 和 MobileNetV3 效果欠佳, 在此不展示。

能力.

表 3 实验对比了不同模型的性能, Params 表示模型训练的参数量, FLOPs 表示模型每秒的浮点运算次数, FPS 表示模型每秒内可以处理的图片数量. 实验结果表明, 本文提出的 DAF-STDC 算法, 在保持模型小参数量同时, 模型精度达到 83.12%, 每秒处理图片数量为 34.15 张. 相较于其余高实时性模型, DAF-STDC 的平均交并比相较于最好效果的原始 STDC 网络有了 7.51% 的提升, 满足工业应用的速率要求, 并且方便在

工业机器上部署和迁移. 相较于其余高精度模型, DAF-STDC 的平均交并比相较于最好效果的 DANet 网络提高了 3%, 具有较好的分割效果. 相较于专用岩屑模型

BiSeNet_SPARM_S-iAFF, 本文网络模型同样具有更好的分割效果. 综合来看, 本文方法可以满足工业应用的实时性和精度要求, 具有优秀的性能.

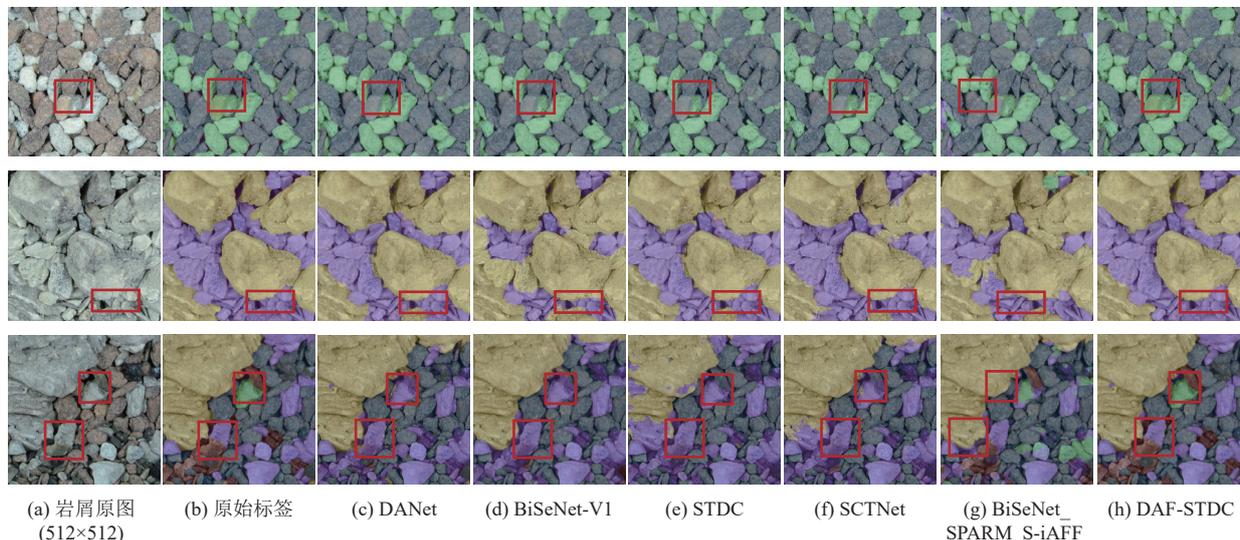


图 9 不同网络分割效果对比

表 3 不同模型性能比较

网络模型	<i>mIoU</i> (%)	Params (M)	FLOPs (G)	FPS
U-Net	70.93	28.98	203	14.62
DANet	80.12	47.46	211	8.92
BiSeNet-V1	72.99	13.31	14.99	86.54
MobileNetV3	70.99	3.28	8.83	49.17
STDC	77.61	8.27	8.44	94.86
SCTNet	75.55	16.41	30.16	50.53
BiSeNet_SPARM_S-iAFF	72.24	13.77	20.33	32.76
Ours	83.12	18.47	43.16	34.15

4 结论

本文提出了一种针对岩屑图像的特征进行快速分割识别的方法, 构建了实时语义分割网络模型 DAF-STDC. 通过引入空洞卷积提取特征避免下采样导致精度下降, 利用双通道注意力模块获取特征图中更为全局的信息, 用特征融合模块加强低阶细节特征与高阶语义特征的融合程度, 在保证模型参数量和计算量的同时, 实现高精度的分割识别. 将该网络应用于岩屑录井时采集到的粉末状岩屑样本的识别与分析, 实验结果表明, DAF-STDC 相较于其余分割网络具有更好的分割效果和泛化能力, 适用于多种岩性混合的复杂工况下的岩性识别, 可以有效提高岩屑录井工作效率, 从而加快录井工作数字化进程.

参考文献

- 1 马陇飞, 萧汉敏, 陶敬伟, 等. 基于深度学习岩性分类的研究与应用. 科学技术与工程, 2022, 22(7): 2609–2617. [doi: 10.3969/j.issn.1671-1815.2022.07.007]
- 2 Zhang SX, Yang YY, Sun FC, *et al.* Application of image sensing system in mineral/rock identification: Sensing mode and information process. *Advanced Intelligent Systems*, 2023, 5(11): 2300206. [doi: 10.1002/aisy.202300206]
- 3 张珂, 冯晓晗, 郭玉荣, 等. 图像分类的深度卷积神经网络模型综述. 中国图象图形学报, 2021, 26(10): 2305–2325. [doi: 10.11834/jig.200302]
- 4 Tan MX, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning*. Long Beach: PMLR, 2019. 6105–6114.
- 5 Han K, Wang YH, Tian Q, *et al.* GhostNet: More features from cheap operations. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 1580–1589.
- 6 Ding XH, Zhang XY, Han JG, *et al.* Scaling up your kernels to 31×31: Revisiting large kernel design in CNNs. *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Orleans: IEEE, 2022. 11963–11975.
- 7 Wang SR, Gao J, Li ZM, *et al.* A closer look at self-

- supervised lightweight vision Transformers. Proceedings of the 40th International Conference on Machine Learning. Honolulu: PMLR, 2023. 35624–35641.
- 8 许振浩, 马文, 林鹏, 等. 基于岩石图像迁移学习的岩性智能识别. 应用基础与工程科学学报, 2021, 29(5): 1075–1092.
- 9 谭永健, 田苗, 徐德馨, 等. 基于 Xception 网络的岩石图像分类识别研究. 地理与地理信息科学, 2022, 38(3): 17–22. [doi: 10.3969/j.issn.1672-0504.2022.03.003]
- 10 Wang HJ. Intelligent identification of logging cuttings based on deep learning. Energy Reports, 2022, 8(S12): 1–7.
- 11 Wu ZY, He JL, Huang C, *et al.* A novel feature fusion-based stratum image recognition method for drilling rig. Earth Science Informatics, 2023, 16(4): 4293–4311. [doi: 10.1007/s12145-023-01132-2]
- 12 Lu B, Zhou JW, Wang QK, *et al.* Fusion-based color and depth image segmentation method for rocks on conveyor belt. Minerals Engineering, 2023, 199: 108107. [doi: 10.1016/j.mineng.2023.108107]
- 13 李新叶, 宋维. 基于深度学习的图像语义分割研究进展. 科学技术与工程, 2019, 19(33): 21–27. [doi: 10.3969/j.issn.1671-1815.2019.33.003]
- 14 徐晓龙, 俞晓春, 何晓佳, 等. 基于改进 U-Net 的街景图像语义分割方法. 电子测量技术, 2023, 46(9): 117–123.
- 15 曹国群, 刘桂雄. 基于双模态融合的线缆图像语义分割方法研究. 电子测量技术, 2023, 46(10): 184–188.
- 16 Zhou XX, Gong QM, Liu YQ, *et al.* Automatic segmentation of TBM muck images via a deep-learning approach to estimate the size and shape of rock chips. Automation in Construction, 2021, 126: 103685. [doi: 10.1016/j.autcon.2021.103685]
- 17 Zhang YZ, Yuan BX, Wang YQ, *et al.* A real-time oil content analysis method of cuttings based on deep learning. IEEE Access, 2022, 10: 132083–132094. [doi: 10.1109/ACCESS.2022.3229760]
- 18 覃本学, 沈疆海, 马丙鹏, 等. 基于 Debseg-Net 的岩屑图像语义分割. 科学技术与工程, 2022, 22(29): 12927–12935. [doi: 10.3969/j.issn.1671-1815.2022.29.028]
- 19 孙杰, 滕奇志, 罗崇兴, 等. 改进 BiSeNetV1 实时模型的岩屑图像识别. 计算机系统应用, 2023, 32(10): 45–53. [doi: 10.15888/j.cnki.csa.009245]
- 20 董帅. 基于深度学习的含油岩屑图像识别研究 [硕士学位论文]. 大庆: 东北石油大学, 2023.
- 21 Ogunkeye F, Shan LQ, Hei X. Rock particle segmentation using mask R-CNN. <https://www.preprints.org/manuscript/202404.0624/v1>. (2024-04-07). [doi: 10.20944/preprints202404.0624.v1]
- 22 Huo FC, Liu KM, Dong HL, *et al.* Research on cuttings image segmentation method based on improved MultiRes-UNet++ with attention mechanism. Signal, Image and Video Processing, 2024, 18(1): 799–808.
- 23 Fan MY, Lai SQ, Huang JS, *et al.* Rethinking BiSeNet for real-time semantic segmentation. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 9716–9725.
- 24 Fu J, Liu J, Tian HJ, *et al.* Dual attention network for scene segmentation. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3146–3154.
- 25 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015. 234–241.
- 26 Yu CQ, Wang JB, Peng C, *et al.* BiSeNet: Bilateral segmentation network for real-time semantic segmentation. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 325–341.
- 27 Howard A, Sandler M, Chen B, *et al.* Searching for MobileNetV3. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 1314–1324.
- 28 Xu ZZ, Wu DY, Yu CQ, *et al.* SCTNet: Single-branch CNN with Transformer semantic information for real-time segmentation. Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver: AAAI, 2024. 6378–6386.

(校对责编: 孙君艳)