基于轻量级卷积神经网络的岩石图像岩性识别方法

刘善伟[1]，马志伟[1]，魏世清[1]，魏忠勇[2*]

1. 中国石油大学（华东）海洋与空间信息学院，山东 青岛 266580；
2. 山东省煤田地质局物探测量队，济南 250104

第一作者，E-mail：shanweiliu@163.com

* 通信作者，E-mail：187304503@qq.com

2024-06-24 收稿；2024-09-25 修回；2024-09-29 接受

基金项目：中石油重大科技项目(ZD2019-183-006)

中图分类号：TP183；TU45　　文章编号：2096-8523(2026)01-0360-11　　doi: 10.19509/j.cnki.dzkq.tb20240348

摘　要：岩性识别是油气勘探和开发过程中的重要环节，对于油气勘探定位、储层评价以及储层模型建立具有重要的指导意义。但传统的人工岩性识别方法耗时耗力，经典的深度学习模型虽然识别精度高，但模型的参数量较大，为了提高模型识别精度，同时降低模型的参数量，使模型适用于岩性实时识别工作，首先收集了白云岩、砂岩等 8 种岩石共 3016 张岩石图像构建岩性识别数据集，然后以轻量型卷积神经网络 ShuffleNetV2 模型为基础网络，提出了一种 Rock-ShuffleNetV2 岩性识别模型(RSHFNet 模型)。模型中将混合注意力机制模块(convolutional block attention module，简称 CBAM)以及多尺度特征融合模块(multi-scale feature fusion module，简称 MSF)融入基础网络中以加强模型的特征提取能力，提升模型识别性能，并优化模型中 ShuffleNetV2 单元的堆叠次数以减少模型参数量。结果表明：与基础模型相比，RSHFNet 模型的准确率达到了 87.21%，提高了 4.98%；同时，模型参数量与浮点运算量分别降低到了 869 702 个，0.93×10^8，分别是基础模型的 0.67，0.63 倍，模型参数量明显降低；并且 RSHFNet 模型的综合性能明显优于现有的卷积神经网络。RSHFNet 岩性识别模型具有较高的识别精度和较好的泛化能力，同时更加的轻量化，为实现野外实时的岩性识别工作提供了新思路。

关键词：岩性识别；ShuffleNetV2 网络；混合注意力机制模块；多尺度特征融合模块；卷积神经网络

随着社会快速发展，世界各国对能源的消耗也在与日俱增。现阶段，虽然各种新能源技术不断被提出，但其仍处在探索阶段，目前工业生产与社会生活的能源主要还是依靠石油与天然气等化石能源[1]。在油气勘探过程中，准确地确定地层岩性是精确评估油气储层规模的基础，所以岩性识别对油气勘探具有重要意义[2-3]。

岩性是指岩石的颜色、成分、结构、构造等特征的总和，岩性识别是通过一定的方法来认识和区别岩性的过程，传统的岩性识别方法主要依靠人工识别，通过肉眼观察、薄片鉴定以及实验分析等方法确定岩性，但是此类方法耗时耗力，要求研究人员拥有丰富的专业知识以及专业的仪器设备，成本高昂，而且岩性识别结果受人的主观因素影响[4-8]。随着机器学习技术的发展，机器学习技术也被应用于岩性识别，岩石图像岩性识别的过程本质上是图像分类的过程，故已经有学者利用多层感知机(multilayer perceptron，简称 MLP)、K 近邻算法、支持向量机(support vector machine，简称 SVM)以及决策树等算法实现了岩石图像的分类，但这些机器学习分类方法都存在分类准确率低、分类过程繁琐等缺点[9-13]。近年来，以深度学习为代表的人工智能方法在计算机视觉领域取得了长足发展。深度学习方法可以通过多层神经网络模拟人类大脑，从大规模数据中自

著录信息：刘善伟，马志伟，魏世清，等. 基于轻量级卷积神经网络的岩石图像岩性识别方法[J]. 地质科技通报，2026，45(1)：360-370.
LIU Shanwei，MA Zhiwei，WEI Shiqing，et al. Rock image lithology recognition method based on lightweight convolutional neural network[J]. Bulletin of Geological Science and Technology，2026，45(1)：360-370.

© Editorial Office of Bulletin of Geological Science and Technology. This is an open access article under the CC BY-NC-ND license.　https://dzkjtb.cug.edu.cn

动学习复杂的非线性表征，为图像分类[14-16]、目标检测[17]等各种视觉任务提供了新的解决方案，在岩石图像岩性识别任务中具有强大的应用潜力。

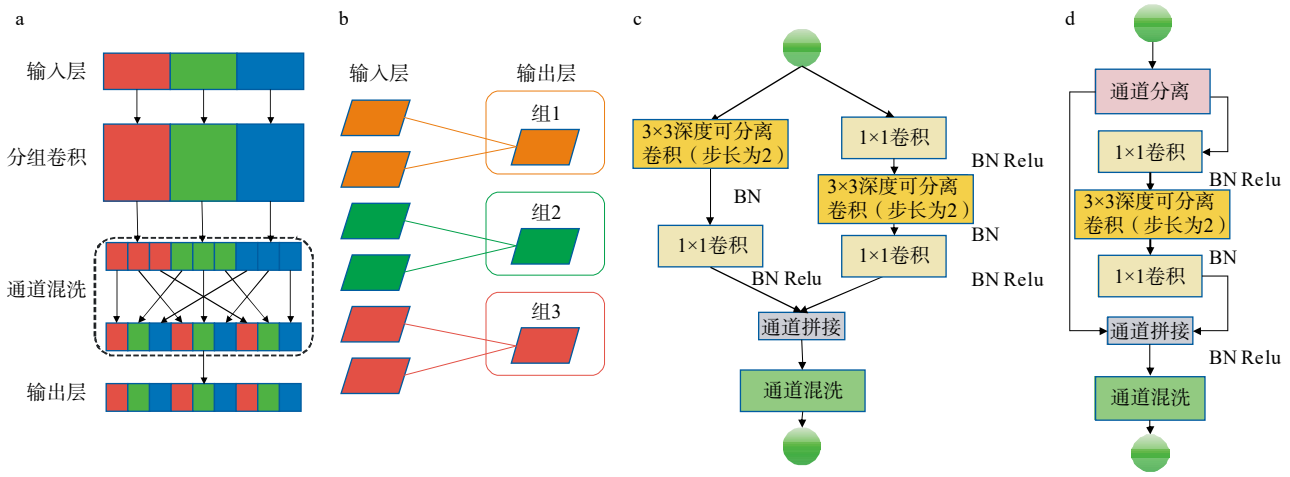
目前，国内外已经有很多学者基于深度学习方法开展了岩性识别研究，程国建等[18]、李娜等[19]使用自构建的卷积神经网络，均实现了对3类砂岩图像的有效识别，但其所采用的图像均为岩石显微镜图像，制作此类图像需将岩石制作成岩石薄片放置在显微镜下拍摄，不适用于野外实时岩性识别，且仅面向砂岩，研究对象相对单一。谭永健等[20]、RAN等[21]、马泽栋等[22]基于在嵊州、巢湖等地野外考察采集的岩石图像，利用Xception、ResNet等经典卷积神经网络，实现了对多种岩石图像的识别，且识别准确率均超过了95%。经典卷积神经网络虽然具有较高的识别精度，但其模型的参数量较大，在目标识别的过程中需要消耗大量的计算资源，难以将其部署至计算资源有限的移动设备中来满足野外实时岩性识别的要求。此外，野外所采集的岩石图像为同一地区拍摄，同类别岩石的颜色与纹理等特征不会有较大差异，然而受制于不同地区的地质条件与气候等因素，同类型的岩石在不同的地区，其颜色与纹理等宏观特征会有较大的差异，而不同类型的岩石也会有相近的性状，所以实际情况下的岩性识别更加具有挑战性。综上所述，利用深度学习图像分类技术在岩性识别工作中已经取得了良好的成果，但为了实现野外实时岩性识别，仍需要解决如下问题：模型参数量大、岩石图像来源单一，适用范围有限。

针对上述问题，本研究以ShuffleNetV2轻量级卷积神经网络[16]为基础，提出了一个识别精度更高、模型更加轻量化的岩性识别模型RSHFNet。该模型通过引入多尺度特征融合模块(MSF)和混合注意力机制模块(CBAM)来增强网络模型的特征提取能力，并优化网络模型深度以减少模型的参数量。同时本研究还收集了8类具有代表性的岩石图像构成岩性识别数据集，这些岩石图像大部分来自于国家岩矿化石标本资源共享平台，岩石图像的采集地为内蒙古、西藏、新疆与河北等不同地区。

1 网络模型构建

本研究提出的RSHFNet模型以ShuffleNetV2作为基础网络。ShuffleNetV2是一个轻量高效的卷积神经网络，通过在ShuffleNetV2单元中进行通道混洗(图1a)与分组卷积(图1b)操作，保持模型精度的同时大幅降低了模型参数量与计算量。ShuffleNetV2网络的主要结构是下采样单元(图1c)与基本单元(图1d)，输入的图像首先进行一次卷积操作与最大池化操作，再经过由下采样单元与基本单元堆栈构成的3个特征提取层，在这3个特征提取层中下采样单元与基本单元的个数分别为(1,3)、(1,7)、(1,3)，最后通过全连接层输出分类结果，从而实现图像分类的任务。

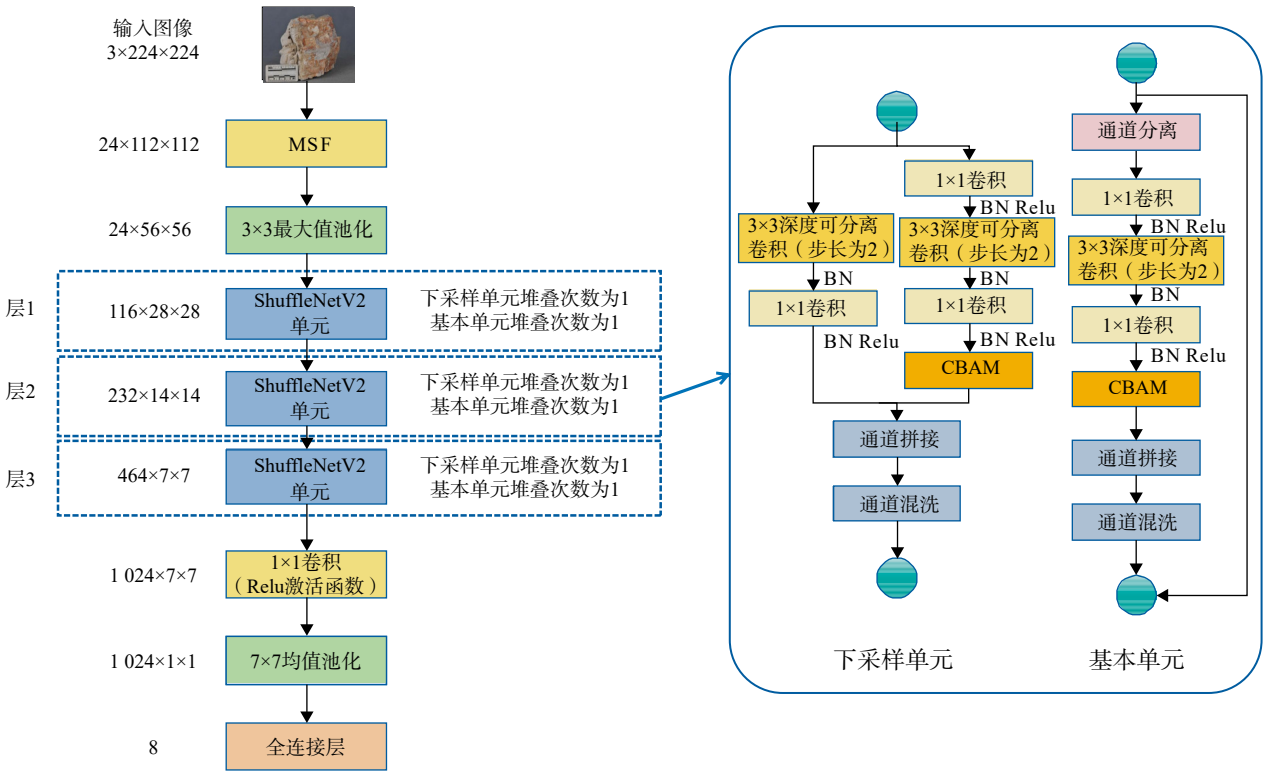
针对现有网络模型参数量大，对岩石图像识别准确率低的问题，本研究根据岩石图像的特点，以ShuffleNetV2网络为基础，提出了一种适用于岩石图像岩性识别的RSHFNet模型，其网络结构如图2所示。RSHFNet模型首先将基础ShuffleNetV2网络中的第一个卷积层替换为MSF，以使得模型在网络



a. 通道混洗；b. 分组卷积；c. 下采样单元；d. 基本单元；BN. 只进行批量化处理；BN Relu. 先进行批量化处理，再用Relu激活函数处理；下同

图1 ShuffleNetV2核心结构

Fig. 1 Core structure of ShuffleNetV2



MSF. 多尺度特征融合模块; CBAM. 混合注意力机制模块; 下同

图 2　RSHFNet 模型结构

Fig. 2　Structure of RSHFNet model

初始阶段提取更多的有效信息; 然后又将 CBAM 融入到 ShuffleNetV2 单元中, 使得模型能够更好地捕捉岩石的重要特征, 以实现更准确的识别; 最后对 ShuffleNetV2 的模型结构进行轻量优化, 减少 3 个特征提取层中基本单元的堆叠次数以减少模型参数量。

（1）多尺度特征融合模块（MSF）

在模型的初始阶段, 基础网络 ShuffleNetV2 只将图像进行一次卷积操作与一次最大池化操作, 然而在图像输入的初始阶段, 图像特征最为丰富, 对图像只进行一次卷积与池化操作并不能有效地提取特征。因此, 本研究结合 GoogLeNet 模型利用多尺度的卷积操作来提取不同大小特征的思想[23], 设计了一种更加轻量化的多尺度特征融合模块(MSF)来替换 ShufflNetV2 基础网络的第一个卷积层。该模块将输入图像分别进行一次 3×3 卷积、5×5 卷积以及 1×1 卷积并平均池化操作, 最后将输出特征在通道维度上进行拼接(图 3)。由于 MSF 模块中采用了不同尺度的卷积核, 可以从图像中提取不同尺度的特征信息, 包括全局特征与局部特征, 其中全局特征可以帮助模型获取更广泛的背景知识, 有助于整体特征的把握, 并帮助减少局部特征可能带来的歧义, 而

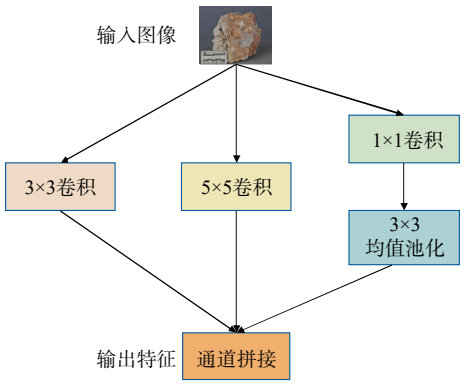


图 3　多尺度特征融合模块（MSF）结构

Fig. 3　Structure of multi-scale feature fusion module

局部特征能有效地捕捉细节, 提高模型对细微特征的感知力。通过融合全局特征与局部特征, 模型可以更加全面地理解输入的图像, 能够更好地适应不同尺度和不同复杂度的输入数据, 进而提高模型的特征提取能力和识别性能。

（2）混合注意力机制模块（CBAM）

为了使分类结果更加准确, 在网络模型中添加注意力机制模块是一种常用的方法[24]。注意力机制模块旨在关注感兴趣的信息, 能够自适应地提取图像中的重要信息。在深度学习卷积神经网络中, 注意力机制模块可以分为 2 类: 通道注意力和空间注

意力。通道注意力模块能够确定不同通道之间的权重关系，增强关键通道的权重并抑制作用较小的通道。空间注意力模块是确定空间领域中不同像素之间的权重关系，增强关键区域像素的权重，并减少不重要区域的权重。注意力机制模块(CBAM)是空间通道混合注意力模块，与仅关注通道特征的注意力模块(SE)相比，可以获得更好的结果。输入特征先经过通道注意力模块，然后再通过空间注意力模块，最后进行输出，实现从通道与空间两方面来增强感兴趣的区域[25](图 4)。由于卷积神经网络在进行特征提取时会产生大量的特征图，其中除了有用信息外，还包含噪声，因此模型需要在大量的特征中找出对当前任务最有用的部分，而卷积神经网络的操作是固定的，不会动态调整特征的重要性。而 CBAM 的通道注意力模块与空间注意力模块分别对通道维度与空间维度上的特征进行选择性关注，从而增强有效信息，并抑制无关信息。输入特征在经过通道注意力模块时，CBAM 会为每个通道分配权重，之后通过多轮迭代计算，自适应地调整每个通道的权重值，既提高有效通道的权重值又降低无用通道的权重值，使得网络模型能够更好地选择和放大对分类结果有帮助的通道特征。在确定好重要的通道之后，空间注意力机制会进一步关注特征图像中的重要像素，并对每一个像素分配一个权重，最后通过多轮的迭代计算，自适应地调整每个像素上的权重值，使得网络模型能够提取出对当前分类任务最具有辨别力的像素。CBAM 从空间与通道 2 个不同的维度对特征进行优化，确保模型不会因为某一维度的误导信息而产生偏差，进而提高模型的特征提取能力与分类性能。由于本研究所使用的数据集为实拍标本图像，图像中岩石标本虽然占据了较大区域，但图像中包含了刻度尺、不同颜色背景等干扰信息，在识别过程中会对模型产生一定的负面影响。因此，本研究将混合注意力机制模块(CBAM)融入 ShuffleNetV2 单元中，使得模型在每个阶段都能够从空间与通道两方面加强对图像岩石区域的关注与细节捕捉能力，从而提高模型对不同岩石的区分能力。

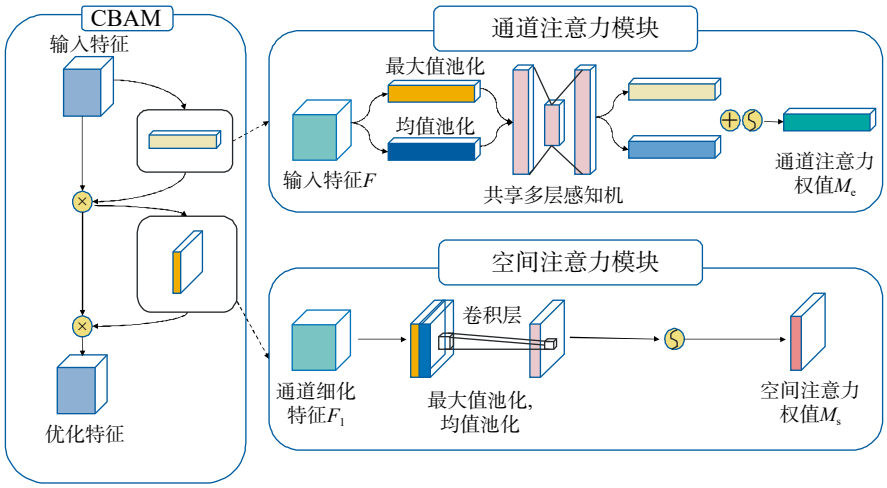


图 4　混合注意力机制模块（CBAM）结构

Fig. 4　Structure of convolutional block attention module

（3）网络模型深度优化

网络模型的轻量优化是通过优化模型的结构与参数，使其在不明显减低性能的前提下减少模型的参数量，从而提升模型在资源受限设备上的推理效率，使其适合应用于实际场景。对于卷积神经网络，通常情况下网络的层数越多，其特征提取能力越强，能够从图像中提取更多具有判别性的信息，有助于提高网络的分类识别性能。但是随着网络层数的增加，模型的参数量与复杂度也越来越大，需要越来越多的计算资源。此外，增加网络深度并不意味着总能提升网络的性能，过深的网络可能会增加网络模型过拟合的风险，导致网络的泛化能力下降[26]。因此，本研究将基础网络 3 个特征提取层中 ShuffleNetV2 基本单元的堆叠次数都优化为 1 次，使模型更加轻量化并降低网络模型过拟合的风险。

2　数据集建立

油气的储集层通常是具有较高的孔隙度和渗透性的岩石，常见的储集层岩石包括碳酸盐岩、砂岩，页岩与泥岩[27-29]。其中碳酸盐岩包括白云岩、石灰

岩与方解石，为增加数据集岩石图像的类别，将作为碳酸盐岩矿产的菱镁矿与菱锰矿也作为岩性识别的种类。泥岩本身的孔隙度和渗透性较低，但泥岩中往往富含有机质且具有较高的含油气饱和度，有助于油气的富集和保存[29]。故本研究选择此 8 类岩石图像构建岩性识别数据集。

实验图像来自“国家岩矿化石标本资源共享平台”(http://www.nimrf.net.cn/)与网络搜索等途径，得到白云岩(图 5a)、方解石(图 5b)、菱锰矿(图 5c)、菱镁矿(图 5d)、石灰岩(图 5e)、砂岩(图 5f)、页岩(图 5g)与泥岩(图 5h)8 类共 3 016 张岩石图像，其中包括白云岩 429 张、方解石 515 张、菱镁矿 56 张、菱锰矿 75 张、石灰岩 285 张、砂岩 618 张、页岩 521 张、泥岩 517 张。将数据集以 8∶2 的比例划分为训练集与测试集，训练集图像共 2 414 张，测试集共 602 张。从图 5 中也可以看出，本研究制作的岩性识别数据集存在不同类别岩石特征相近，同类别岩石特征又不尽相同的情况，总体识别难度较大。



a. 白云岩；b. 方解石；c. 菱锰矿；d. 菱镁矿；e. 石灰岩；f. 砂岩；g. 页岩；h. 泥岩

图 5　岩性识别数据集示例图像

Fig. 5　Example images of lithology identification dataset

关于训练集数据增强，由于数据集中部分种类的岩石图像样本数量较少，易导致网络提取不到足够的特征，使网络模型出现泛化能力差、网络过拟合等问题，影响图像识别精度。通过对训练集图像数据增强可以提高网络的泛化能力与识别能力[30]。本研究使用的数据增强方法包括：图像翻转、图像旋转、图像平移、亮度调整、色度调整、对比度调整、锐度调整以及裁剪拼接(图 6)。原始训练集中菱镁矿与菱锰矿样本数量较少，而其余 6 类岩石样本数量较多，因此对属于弱样本的菱镁矿与菱锰矿进行多倍数的数据增强，而对其余 6 类岩石样本则进行少倍数的数据增强。数据增强后训练集图像总量达到了 8 780 张(表 1)。

3　实验与结果分析

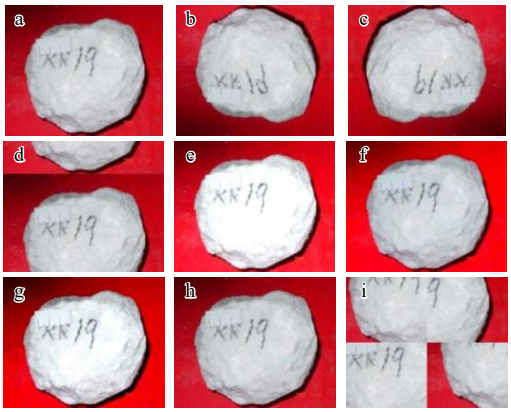
3.1　评价指标

为了有效评价网络的性能，使用准确率(*Accuracy*)、精确率(*Precision*)、召回率(*Recall*)以及 *F*1 分数(*F*1-score)作为评估标准，其计算公式如下：

$$Accuracy=\frac{TP+TN}{TP+FP+TN+FN}\times 100\%, \quad (1)$$

$$Precision=\frac{TP}{TP+FP}\times 100\%, \quad (2)$$

$$Recall=\frac{TP}{TP+FN}\times 100\%, \quad (3)$$



a. 原始图像；b. 图像翻转；c. 图像旋转；d. 图像平移；e. 亮度调整；f. 色度调整；g. 对比度调整；h. 锐度调整；i. 裁剪拼接

图 6　数据增强后图像效果

Fig. 6　Effects of data augmentation on images

表 1　数据增强前后训练集分布

Table 1　Distributions of training set before and after data augmentation

岩性	原始训练集数量/张	增强后训练集数量/张	岩性	原始训练集数量/张	增强后训练集数量/张
白云岩	343	1 372	石灰岩	228	1 368
方解石	412	1 236	砂岩	495	1 485
菱镁矿	45	360	页岩	417	1 251
菱锰矿	60	466	泥岩	414	1 242

$$F1\text{-score}=\frac{2\times Recall\times Precision}{Recall+Precision}\times 100\%,\quad (4)$$

式中：*TN* 为真阴性，表示预测为阴性的阴性样本的数量；*TP* 为真阳性，表示预测为阳性的阳性样本的数量；*FN* 为假阴性，表示预测为阴性的阳性样本的数量；*FP* 为假阳性，表示被预测为阳性的阴性样本的数量。

此外还使用模型参数量与浮点运算量来衡量模型的复杂度，模型参数量与浮点运算量越低就表示需要越少的计算资源，越容易将模型部署到移动设备。

3.2　实验环境与超参数设置

本实验硬件设备 CPU 的型号为 Intel(R)Core(TM)i9-10980XE，显卡的型号为 NVIDIA GeForce RTX 2080 TI。在 Pycharm 集成环境下，使用 Pytorch 2.1.2 搭建深度学习框架。网络训练过程中采用随机梯度下降(stochastic gradient descent，简称 SGD)优化器进行参数更新，初始学习率为 0.01，每完成 2 轮的训练后更新一次学习率，损失函数为交叉熵函数(cross entropy)，训练轮数为 60，一次训练使用的样本数量(*Batchsize*)设置为 32。

3.3　实验结果分析

（1）数据增强对实验结果的影响

为了验证数据增强是否能够提升网络模型的性能，分别对原始数据集以及经过平移、翻转、亮度调整等数据增强方法扩充后的数据集进行训练，实验中保证网络模型结构、实验环境以及超参数等其他因素相同。对比原始数据集，ShuffleNetV2 模型对数据增强后的数据集进行训练，在测试集上的准确率达到了 82.23%，提高了 5.65%；精确率达到了 82.21%，提高了 12.28%；召回率达到了 81.42%，提高了 12.57%；*F*1 分数达到了 81.54%，提高了 12.44%(表 2)。所以，通过数据增强的方法扩充原始数据集，能够平衡各类别岩矿图像的分布，有效地避免过拟合现象，提高模型的泛化能力与识别准确率。

表 2　数据增强前后实验结果

Table 2　Experimental results before and after data augmentation

数据集	准确率/%	精确率/%	召回率/%	*F*1分数/%
原始数据集	76.58	69.93	68.85	69.10
数据增强后的数据集	82.23	82.21	81.42	81.54

（2）注意力模块对模型性能的影响

为了验证注意力模块对岩石图像识别是否有效，本研究以数据增强后的数据集为基础，将通道注意力模块机制(SE)、混合注意力模块机制(CBAM)、通道注意力模块机制(ECA)以及坐标注意力模块机制(CA)分别引入基础网络中，并与基础网络性能相对比。由实验结果可知：引入 SE 注意力机制模块的网络模型的识别准确率提高了 0.66%；引入 ECA 注意力机制的网络模型的准确率提高了 1.09%、精确率提高了 1.59%、*F*1 分数提高了 0.61%；引入 CA 注意力机制的网络模型的准确率提高了 0.83%、精确率提高了 0.81%；引入 CBAM 注意力机制的网络模型的准确率提高了 3.82%、精确率提高了 3.88%、召回率提高了 3.33%、*F*1 分数提高了 3.58%(表 3)。这说明针对本研究的数据集，引入注意力模块机制并不一定对模型的性能都有积极的影响。实验结果表明，混合注意力模块机制(CBAM)对于网络模型的性能有着最佳的影响，在增加少量模型参数量的前提下能够较大地提高模型的识别准确率、精确率、召回率与 *F*1 分数。

表 3　引入不同注意力模块后的实验结果

Table 3　Experimental results after introducing different attention modules

模型	准确率/%	精确率/%	召回率/%	*F*1分数/%	模型参数量/个	浮点运算量/10^8
基础网络	82.23	82.21	81.42	81.54	1 295 037	1.47
基础网络+SE	82.89	81.30	80.21	80.19	1 303 064	1.48
基础网络+ECA	83.32	83.80	81.21	82.15	1 260 268	1.48
基础网络+CA	83.06	83.02	79.76	80.93	1 325 720	1.48
基础网络+CBAM	86.05	86.09	84.75	85.12	1 304 648	1.48

（3）消融实验

为了进一步验证本研究所提出的每个改进点都有助于提高模型的性能，将基础网络与上文提出的各优化后的模型进行消融实验对比测试。其中基础网络+CBAM 表示在基础网络中引入 CBAM 注意力模块、基础网络+MSF 表示在基础网络中引入 MSF 多尺度特征融合模块、基础网络-Lighter 为对基础网

络进行深度优化后的模型。实验结果表明，混合注意力机制模块(CBAM)、多尺度特征融合模块(MSF)的采用和模型深度优化对模型的性能都有积极影响。相比基础网络，本研究设计的 RSHFNet 模型的准确率达到了 87.21%，提高了 4.98%；精确率达到了 86.37%，提高了 4.16%；召回率达到了 86.68%，提高了 5.26%；*F*1 分数达到了 86.39%，提高了 4.85%。同时，模型参数量与浮点运算量降低到了 869 702 个与 0.93×10^8，分别是基础网络的 0.67，0.63 倍(表 4)。这充分证明了 RSHFNet 岩性识别模型的优越性以及在岩性识别问题上的可行性。

表 4　消融实验结果

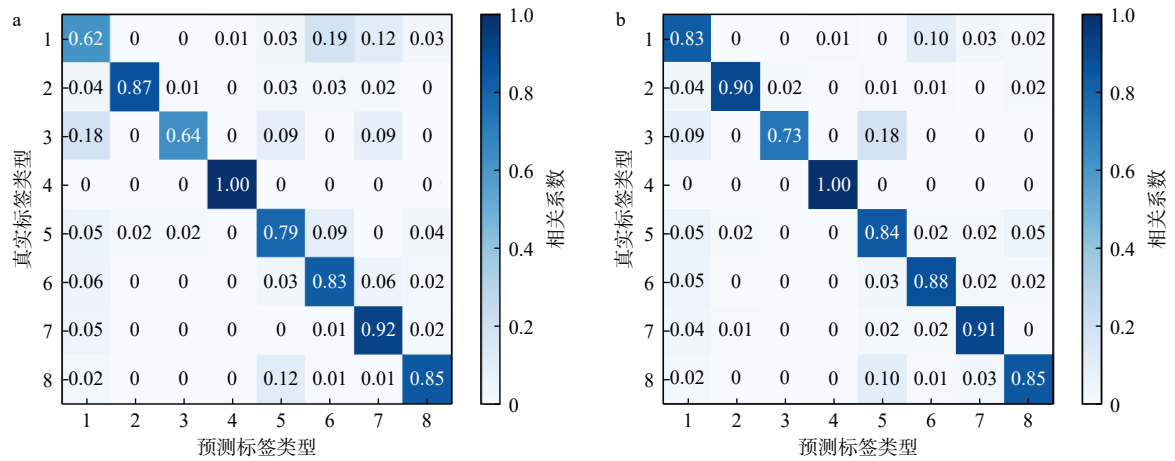
Table 4　Results of ablation experiment

模型	准确率/%	精确率/%	召回率/%	*F*1分数/%	模型参数量/个	浮点运算量/10^8
基础网络	82.23	82.21	81.42	81.54	1 295 037	1.47
基础网络+CBAM	86.05	86.09	84.75	85.12	1 304 648	1.48
基础网络+MSF	83.39	82.84	83.31	82.90	1 260 404	1.51
基础网络-Lighter	85.71	87.41	85.89	83.38	850 848	0.90
RSHFNet	87.21	86.37	86.68	86.39	869 702	0.93

为了能够更加直观地展示模型改进前后对各类岩性识别的性能，绘制了基础网络与 RSHFNet 模型的混淆矩阵图(图 7)，可以看出，基础网络的识别结果中有 19% 的白云岩被识别成了砂岩，18% 的菱镁矿被识别成了白云岩(图 7a)，表明基础网络在岩性识别任务上存在较大的误判。而改进后的模型即 RSHFNet 模型，对岩矿石图像岩性识别总体的准确率推高了 4.98%，其中对白云岩识别的准确率提高了 21%，对菱镁矿识别的准确率提高了 9%，对石灰岩识别的准确率提高了 5%(图 7b)，证明了本研究对模型所做出改进的有效性与优越性，更加直观地表明了 RSHFNet 模型在岩性识别任务上的性能提升。

本研究构建的岩性识别数据集由于各类别的岩矿石图像数量不平衡，所以模型对不同类别的岩石图像识别准确率也不同。数据集中在菱镁矿与菱锰矿的图像数量较少，数据增强后分别为 360，466 张，而其余 6 类的岩石数据增强后均超过了 1 200 张，模型对菱镁矿的识别准确率为 73%，而对菱锰矿的识别准确率却达到了 100%(图 7b)。由于菱锰矿在外观上主要表现为粉红色与深红色，与其他类岩石具有较大的颜色差异，虽然其图像数量较少，但模型依旧能够提取出其独有的特征信息，故基础网络与 RSHFNet 模型对其均有很高的识别准确率。菱镁矿与白云岩、石灰岩、方解石均属于碳酸盐岩，四者的颜色、纹理具有较高的相似性，但白云岩、石灰岩与方解石图像数量较多，模型在训练阶段能够较充分地提取其独有特征，因此具有较高的识别准确率。对于砂岩、泥岩与页岩 3 类，用于模型训练的砂岩图像数量最多，为 1 485 张，泥岩与页岩均为 1 250 张左右，三者也均有较高的识别准确率，其中泥岩的识别准确率最高。综上，由于数据集中在菱镁矿的图像数量较少且与石灰岩等碳酸盐岩类岩石的特征较为相似，故对菱镁矿的识别能力较差，本研究改进的模型对菱镁矿的识别能力虽然有提升，但仍存在一定的局限性，可能会导致模型对菱镁矿的泛化识别能力较差，对于数据集中图像数量较多的岩石类别，模型均具有较好的识别能力。

(4) 对比实验



a. ShuffleNetV2 混淆矩阵；b. RSHFNet 混淆矩阵；1. 白云岩；2. 方解石；3. 菱镁矿；4. 菱锰矿；5. 石灰岩；6. 砂岩；7. 泥岩；8. 页岩

图 7　改进前后混淆矩阵

Fig. 7　Confusion matrix before and after improvement

为了进一步评估 RSHFNet 模型对岩石图像的识别性能，在相同的实验条件下，将 RSHFNet 模型与经典卷积神经网络 VGG16、ResNet18、ResNet50、DenseNet169、GhostNet、MobileNetV3-small、MobileNetV3-large 以及袁硕等[31]提出的改进 ShuffleNet-V2 岩石图像识别模型（以下简称为 ShuffleNetV2-Y）进行对比。由结果可知，相较于对比网络中性能最优的 DenseNet169，RSHFNet 在准确率上与之相同，精确率高出了 0.74%，F1 分数高出了 0.48%，但 RSHFNet 的模型参数量与浮点运算量仅为 DenseNet169 的 7%、2.7%，模型复杂度明显优于 DenseNet169。较之于其他模型，MobileNetV3-small 模型的浮点运算量虽然小于 RSHFNet，但 RSHFNet 模型的岩性识别性能明显优于 MobileNetV3-small；ShuffleNetV2-Y 的模型参数量与浮点运算量都小于本研究提出的 RSHFNet 岩性识别模型，但针对本研究的数据集，该模型的识别性能也远低于 RSHFNet（表 5）。由此可知，本研究提出的 RSHFNet 模型具有较强的综合能力。

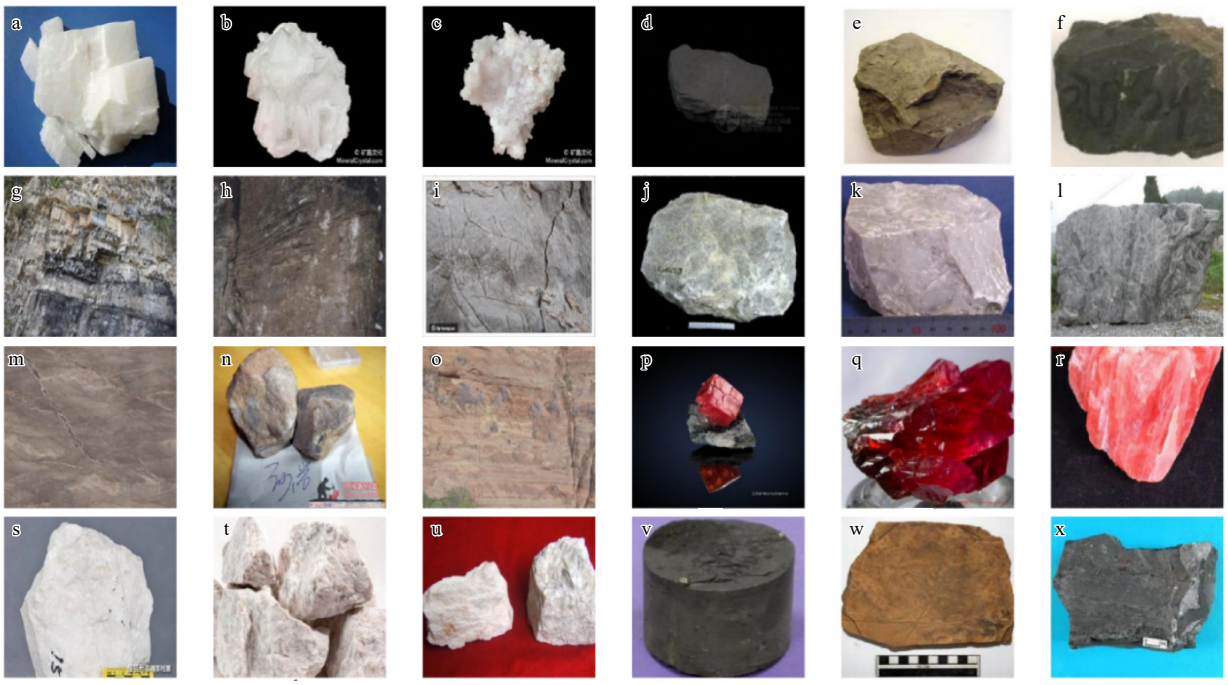
（5）泛化能力验证实验

为了进一步验证 RSHFNet 模型的泛化能力，通过互联网搜索途径采集数据集外 8 种岩石图像进行预测，所收集的图像包含岩石标本图像、岩心图像与野外露头图像等，每类岩石图像各 3 张，共计 24 张（图 8）。可以看出 24 张图像中预测正确 22 张，预测错误 2 张，其中将 1 张菱镁矿图像预测成石灰岩（图 8t）、1 张页岩图像预测成砂岩（图 8w）。对菱镁矿预测错误的原因可能是用于模型训练的图像数量较少，模型未能提取到其独有特征。对页岩预测错误的原因可能是该图像中的页岩含有较多铁的氧化物或氢氧化物，使其在颜色上呈现出与多数砂岩相似的黄褐色，导致模型误将其预测为砂岩。总体而言，虽然模型在对数据集外图像的预测结果中出现

表 5　不同模型的对比实验结果

Table 5　Comparative experimental results of different models

模型	准确率/%	精确率/%	召回率/%	*F*1分数/%	模型参数量/个	浮点运算量/10^8
VGG16	67.77	70.68	67.66	66.96	134 301 768	155.20
ResNet18	85.71	84.86	86.29	85.27	11 180 616	18.24
ResNet50	85.05	85.09	83.73	84.11	23 524 424	41.32
DenseNet169	87.21	85.63	86.83	85.91	12 497 800	34.34
GhostNet	82.39	84.18	82.80	83.09	4 212 120	1.97
MobileNetV3-small	82.72	78.18	78.72	78.71	1 238 996	0.65
MobileNetV3-large	85.88	83.51	82.56	82.83	2 685 164	2.71
ShufleNetV2	82.23	82.21	81.42	81.54	1 295 037	1.47
ShuffleNetV2-Y	79.57	79.31	78.24	78.13	275 066	0.33
RSHFNet	87.21	86.37	86.68	86.39	869 702	0.93



a ~ c. 预测类别：方解石；d ~ f. 预测类别：泥岩；g ~ i. 预测类别：白云岩；j ~ l, t. 预测类别：石灰岩；m ~ o, w. 预测类别：砂岩；p ~ r. 预测类别：菱锰矿；s, u. 预测类别：菱镁矿；v, x. 预测类别：页岩

图 8　RSHFNet 模型预测结果

Fig. 8　Predicted results of the RSHFNet model

了错误，但总体的预测正确率达到了 91.7%，证明本研究提出的 RSHFNet 模型具有较好的泛化能力。

4　结　语

岩性识别是油气勘探和开发工作中的重要环节，相较于传统的人工识别方法，利用深度学习技术可以极大地提高岩性识别的效率。本研究基于 ShufleNetV2 模型，提出了一种适用于岩性识别的 RSHFNet 轻量化模型，并构建了一套适用范围相对较广的岩性识别数据集，以解决当前岩性识别研究中存在的模型参数量大、岩石图像适用范围有限的问题。RSHFNet 模型首先融入了多尺度特征融合模块（MSF），以加强模型在初始阶段的特征提取能力；其次又将混合注意力机制模块（CBAM）融入到 ShuffleNetV2 单元中，从空间与通道两方面增强了模型提取特征的能力；最后又对模型深度进行了优化，降低了模型的参数量与复杂度。实验结果表明：本研究提出 RSHFNet 模型的岩性识别准确率达到了 87.21%，相较于基础 ShuffleNetV2 模型提高了 4.98%，精确率、召回率与 *F*1 分数等模型评价指标均有提升，同时模型的参数量与浮点运算量分别降低了 33% 与 37%，复杂度明显降低，并且通过与 DenseNet169、MobileNetV3-small 等卷积神经网络的一系列对比实验证明了 RSHFNet 模型综合性能的优越性；对数据集外 24 张岩石图像预测的准确率为 91.7%，进一步验证了模型的泛化能力，故本研究所做的研究对实现野外岩性实时识别具有一定的指导意义。本研究提出的 RSHFNet 轻量化模型相较于基础 ShuffleNetV2 模型的综合性能有一定的提升，在后续的研究中将扩展岩性类别和数据来源，构建岩性类别更多、图像数量更加平衡的岩性识别数据集，同时考虑将该轻量化模型部署至移动设备，以实现岩石岩性的现场识别。

所有作者不存在利益冲突。

The authors declare that no competing interests exist.

参考文献：

[1] 陈丽，刘娟，王末，等. 中国油料作物能源利用效率与温室气体排放 [J]. 中国生态农业学报（中英文），2023，31（12）：1984-1996.
CHEN L，LIU J，WANG M，et al. Energy use efficiency and greenhouse gas emissions of oil crops in China[J]. Chinese Journal of Eco-Agriculture，2023，31（12）：1984-1996.（in Chinese with English abstract）

[2] ZOU C N，YANG Z，ZHU R K，et al. Progress in China's unconventional oil & gas exploration and development and theoretical technologies[J]. Acta Geologica Sinica（English Edition），2015，89（3）：938-971.

[3] 王婷婷，孙振轩，戴金龙，等. 松辽盆地中央坳陷区储层岩性智能识别方法 [J]. 吉林大学学报（地球科学版），2023，53（5）：1611-1622.
WANG T T，SUN Z X，DAI J L，et al. Intelligent identification method of reservoir lithology in Central Depression of Songliao Basin[J]. Journal of Jilin University（Earth Science Edition），2023，53（5）：1611-1622.（in Chinese with English abstract）

[4] 胡前泽，王玲利，崔娥. 三塘湖盆地致密油储层评价及成藏主控因素 [J]. 大庆石油地质与开发，2013，32（3）：164-169.
HU Q Z，WANG L L，CUI E. Tight reservoir evaluation of the unconventional oil and study on the main controlling factors of hydrocarbon accumulation in Santanghu Basin[J]. Petroleum Geology & Oilfield Development in Daqing，2013，32（3）：164-169.（in Chinese with English abstract）

[5] 陈国俊，吕成福，王琪，等. 珠江口盆地深水区白云凹陷储层孔隙特征及影响因素 [J]. 石油学报，2010，31（4）：566-572.
CHEN G J，LYU C F，WANG Q，et al. Characteristics of pore evolution and its controlling factors of Baiyun Sag in deepwater area of Pearl River Mouth Basin[J]. Acta Petrolei Sinica，2010，31（4）：566-572.（in Chinese with English abstract）

[6] MŁYNARCZUK M，GÓRSZCZYK A，ŚLIPEK B. The application of pattern recognition in the automatic classification of microscopic rock images[J]. Computers & Geosciences，2013，60：126-133.

[7] 付光明，严加永，张昆，等. 岩性识别技术现状与进展 [J]. 地球物理学进展，2017，32（1）：26-40.
FU G M，YAN J Y，ZHANG K，et al. Current status and progress of lithology identification technology[J]. Progress in Geophysics，2017，32（1）：26-40.（in Chinese with English abstract）

[8] MARTIN K G，CARR T R. Developing a quantitative mudrock calibration for a handheld energy dispersive X-ray fluorescence spectrometer[J]. Sedimentary Geology，2020，398：105584.

[9] SINGH N，SINGH T N，TIWARY A，et al. Textural identification of basaltic rock mass using image processing and neural network[J]. Computational Geosciences，2010，14（2）：301-310.

[10] 郭超，刘烨. 多色彩空间下的岩石图像识别研究 [J]. 科学技术与工程，2014，14（18）：247-251.
GUO C，LIU Y. Recognition of rock images based on multiple color spaces[J]. Science Technology and Engineering，2014，14（18）：247-251.（in Chinese with English abstract）

[11] 程国建，殷娟娟. 基于 SVM 的岩石薄片图像分类 [J]. 科技创新与应用，2015，5（1）：38.
CHENG G J，YIN J J. Classification of rock slice images based on SVM[J]. Technology Innovation and Application，2015，5（1）：38.（in Chinese with English abstract）

[12] DENG C X，PAN H P，FANG S N，et al. Support vector machine as an alternative method for lithology classification of crystalline rocks[J]. Journal of Geophysics and Engineering，2017，14（2）：341-349.

[13] 张驰，潘懋，胡水清，等. 融合储层纵向信息的机器学习岩性识别

方法 [J]. 地质科技通报, 2023, 42(3): 289-299.
ZHANG C, PAN M, HU S Q, et al. A machine learning lithologic identification method combined with vertical reservoir information[J]. Bulletin of Geological Science and Technology, 2023, 42(3): 289-299. (in Chinese with English abstract)

[14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014: 1409.1556.

[15] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Anon. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2016: 770-778.

[16] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design[M]// Computer vision-ECCV 2018. Cham: Springer International Publishing, 2018: 122-138.

[17] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Anon. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2023: 7464-7475.

[18] 程国建, 郭文惠, 范鹏召. 基于卷积神经网络的岩石图像分类 [J]. 西安石油大学学报(自然科学版), 2017, 32(4): 116-122.
CHENG G J, GUO W H, FAN P Z. Study on rock image classification based on convolution neural network[J]. Journal of Xi'an Shiyou University (Natural Science Edition), 2017, 32(4): 116-122. (in Chinese with English abstract)

[19] 李娜, 顾庆, 姜枫, 等. 一种基于卷积神经网络的砂岩显微图像特征表示方法 [J]. 软件学报, 2020, 31(11): 3621-3639.
LI N, GU Q, JIANG F, et al. Feature representation method of microscopic sandstone images based on convolutional neural network[J]. Journal of Software, 2020, 31(11): 3621-3639. (in Chinese with English abstract)

[20] 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 [J]. 地理与地理信息科学, 2022, 38(3): 17-22.
TAN Y J, TIAN M, XU D X, et al. Rock image classified identification based on Xception network[J]. Geography and Geo-Information Science, 2022, 38(3): 17-22. (in Chinese with English abstract)

[21] RAN X J, XUE L F, ZHANG Y Y, et al. Rock classification from field image patches analyzed using a deep convolutional neural network[J]. Mathematics, 2019, 7(8): 755.

[22] 马泽栋, 马雷, 李科, 等. 基于岩石图像深度学习的多尺度岩性识别 [J]. 地质科技通报, 2022, 41(6): 316-322.
MA Z D, MA L, LI K, et al. Multi-scale lithology recognition based on deep learning of rock images[J]. Bulletin of Geological Science and Technology, 2022, 41(6): 316-322. (in Chinese with English abstract)

[23] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Anon. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2015: 1-9.

[24] 高广尚. 深度学习推荐模型中的注意力机制研究综述 [J]. 计算机工程与应用, 2022, 58(9): 9-18.
GAO G S. Survey on attention mechanisms in deep learning recommendation models[J]. Computer Engineering and Applications, 2022, 58(9): 9-18. (in Chinese with English abstract)

[25] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[M]//Computer vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.

[26] HUANG G, SUN Y, LIU Z, et al. Deep networks with stochastic depth[M]//Computer vision-ECCV 2016. Cham: Springer International Publishing, 2016: 646-661.

[27] 马永生, 蔡勋育, 赵培荣. 深层、超深层碳酸盐岩油气储层形成机理研究综述 [J]. 地学前缘, 2011, 18(4): 181-192.
MA Y S, CAI X Y, ZHAO P R. The research status and advances in porosity evolution and diagenesis of deep carbonate reservoir[J]. Earth Science Frontiers, 2011, 18(4): 181-192. (in Chinese with English abstract)

[28] 盛双占, 潘海峰, 石湘, 等. 吐哈盆地胜北凹陷三间房组致密砂岩气藏沉积特征及有利区优选 [J]. 地质科技通报, 2024, 43(4): 181-190.
SHENG S Z, PAN H F, SHI X, et al. Sedimentary characteristics and favourable area selection of tight sandstone gas reservoirs in the Sanjianfang Formation of the Shengbei Depression, Tuha Basin[J]. Bulletin of Geological Science and Technology, 2024, 43(4): 181-190. (in Chinese with English abstract)

[29] 付小平, 刘苗苗. 涪陵地区凉高山组富有机质泥岩微相特征及油气富集规律 [J]. 断块油气田, 2023, 30(2): 230-237.
FU X P, LIU M M. Microfacies characteristics of organic-rich mudstone and oil and gas enrichment law of Lianggaoshan Formation in Fuling area[J]. Fault-Block Oil & Gas Field, 2023, 30(2): 230-237. (in Chinese with English abstract)

[30] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 60.

[31] 袁硕, 刘玉敏, 安志伟, 等. 基于改进 ShuffleNetV2 网络的岩石图像识别 [J]. 吉林大学学报(信息科学版), 2023, 41(3): 450-458.
YUAN S, LIU Y M, AN Z W, et al. Rock image recognition based on improved ShuffleNetV2 network[J]. Journal of Jilin University (Information Science Edition), 2023, 41(3): 450-458. (in Chinese with English abstract)

Summary for “基于轻量级卷积神经网络的岩石图像岩性识别方法”

Rock image lithology recognition method based on lightweight convolutional neural network

LIU Shanwei[1]，MA Zhiwei[1]，WEI Shiqing[1]，WEI Zhongyong[2*]

1. College of Ocean and Space Information, China University of Petroleum (East China), Qingdao Shandong 266580, China；
2. Geophysical Exploration and Surveying Team, Shandong Bureau of Coalfield Geology, Jinan 250104, China

First author，E-mail：shanweiliu@163.com

* Corresponding author， E-mail：187304503@qq.com

Abstract：[**Objective**] Lithology identification is a crucial step in the process of oil and gas detection and exploration， providing important guidance for exploration positioning， reservoir evaluation， and the establishment of reservoir models. However， traditional manual lithology identification methods are time-consuming and labor-intensive. Although classical deep learning models achieve high identification accuracy，they often have a large number of parameters. To enhance model accuracy while reducing the number of parameters，the aim of this research is to make the model suitable for real-time lithology identification. [**Methods**] This paper first collected a dataset of 3 016 rock images consisting of eight types of rocks，including dolomite and sandstone. Based on the lightweight convolutional neural network ShuffleNetV2，the paper proposes a Rock-ShuffleNetV2 lithology identification model (RSHFNet model). The model incorporates the convolutional block attention module (CBAM) and multi-scale feature fusion module (MSF) into the basic network to enhance feature extraction capabilities and improve identification performance. Additionally，the number of stacked ShuffleNetV2 units is optimized to reduce the model's parameters. [**Results**] The experimental results show that the RSHFNet model achieved an accuracy of 87.21%， which is a 4.98% improvement over the baseline model. Furthermore， the model's parameters and floating-point operations were reduced to 869 702, 0.93×10^{8}， respectively， representing 67% of the model's parameters and 63% of the floating-point operations of the baseline model. This reduction significantly decreases the model's size. Additionally，the RSHFNet model demonstrates superior overall performance compared to existing convolutional neural networks. [**Conclusion**] The proposed RSHFNet lithology identification model offers high recognition accuracy and strong generalization capabilities while being more lightweight， providing a new approach for real-time lithology identification in the field.

Key words: lithological recognition；ShuffleNetV2 network；convolutional block attention module；multi-scale feature fusion module；convolutional neural network

Received：2024-06-24；Revised：2024-09-25；Accepted：2024-09-29

doi：10.19509/j.cnki.dzkq.tb20240348