

刘艳如, 吴晓红, 何小海, 等. 基于改进 ResNet50 的岩心图像分类研究 [J]. 智能计算机与应用, 2025, 15(2): 10-16. DOI: 10.20169/j.issn.2095-2163.24091102

基于改进 ResNet50 的岩心图像分类研究

刘艳如¹, 吴晓红¹, 何小海¹, 罗彬彬², 滕奇志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 610065)

摘要: 岩心岩性是反映地质条件的重要指标, 传统的岩性鉴定通常依赖于人工目视检查, 既费时又对专业水平要求高。近年来, 卷积神经网络技术的迅速进步, 为岩心图像的自动化预测开辟了一条新的途径。本文提出了一种基于改进的 ResNet50 网络结构的岩心图像分类算法, 通过引入 ECA (Efficient Channel Attention) 注意力机制和 PSA (Pyramid Scene Attention) 注意力机制, 改善了网络对岩心图像丰富地质信息的提取和理解能力, 对提高岩性分类的准确性和客观性起到了重要作用; 引入可变形卷积 (DCNv2), 使模型能够自动适应图像特征不规则性和形状变化, 显著提升了对岩心结构复杂性的识别能力; 使用迁移学习方法, 提高了模型的泛化能力和训练效率。实验结果表明, 改进的 ResNet50 网络模型在岩心图像分类任务上表现优异, 相较于其他主流卷积网络, 平均准确率明显提升, 较基线网络 ResNet50 提高了 2.33% 的准确率, 也有效地提高了对复杂岩心结构的识别精度与鲁棒性。

关键词: 注意力机制; 岩心图像; 图像分类; ResNet50; 可变形卷积

中图分类号: TP237

文献标志码: A

文章编号: 2095-2163(2025)02-0010-07

Research on core image classification based on improved ResNet50

LIU Yanru¹, WU Xiaohong¹, HE Xiaohai¹, LUO Binbin², TENG Qizhi¹

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

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

Abstract: The lithology of rock cores serves as a crucial indicator reflecting geological conditions. Traditionally, lithological identification relies heavily on manual visual inspection, which is both time-consuming and requires a high level of expertise. In recent years, the rapid development of CNNs (Convolutional Neural Networks) has provided an innovative approach for automated prediction of rock core images. This paper proposes an improved ResNet50 network-based algorithm for rock core image classification. By incorporating ECA (Efficient Channel Attention) and a specifically designed PSA (Pyramid Scene Attention) mechanism, the algorithm enhances the network's ability to extract and comprehend the rich geological information present in rock core images, thereby playing a pivotal role in improving the accuracy and objectivity of lithological classification. Furthermore, the introduction of DCNv2 (Deformable Convolution Networks version 2) enables the model to automatically adapt to irregularities and shape variations in image features, significantly enhancing its recognition capabilities for the complexity of rock core structures. Additionally, utilizing transfer learning methods improves the model's generalization ability and training efficiency. Experimental results demonstrate that the modified ResNet50 network model excels in the task of rock core image classification, achieving a notable increase in average accuracy compared to other mainstream CNNs. Specifically, it outperforms the baseline ResNet50 model by 2.33% in accuracy, effectively elevating the recognition precision and robustness for complex rock core structures.

Key words: attention mechanism; core images; image classification; ResNet50; deformable convolution

0 引言

岩心图像分类识别对于石油工程、地质勘探和

地质学研究十分重要。智能技术还未普及时, 岩心的岩性识别主要依赖于专业人员的经验和地质知识。

基金项目: 国家自然科学基金(62071315)。

作者简介: 刘艳如(1999—), 女, 硕士研究生, 主要研究方向: 图像处理与信息论, 人工智能与计算机视觉; 何小海(1964—), 男, 博士, 教授, 博士生导师, 主要研究方向: 图像处理与网络通信, 机器视觉与智能系统; 罗彬彬(1990—), 男, 学士, 主要研究方向: 软件工程; 滕奇志(1961—), 女, 博士, 教授, 博士生导师, 主要研究方向: 数字图像处理, 模式识别, 三维图像重建与分析。

通信作者: 吴晓红(1970—), 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: wxh@scu.edu.cn。

收稿日期: 2024-09-11

卷积神经网络(CNN)在计算机视觉领域发挥了重要作用^[1]。王浩等^[2]采用多尺度特征融合方法,利用向量机训练对400份岩心图像进行识别;Zhang等^[3]鉴定了砂岩、页岩和砾岩3个选定的岩性,获得了95%的预测准确率;Baraboshkin等^[4]收集了不同地区2000m的岩心,获得了6大类岩性,在GoogLeNet上取得72%的准确率;Alzubaidi等^[5]从钻孔中获得包含砂岩、石灰岩和页岩3种岩性,还添加了第四类来表征钻孔岩心的非岩心部分,实现了93.12%的预测准确率;许振浩等^[6]通过深度监督目标检测网络检测出图像中的岩石并进行自动裁剪分类;谭永健等^[7]使用Xception网络对岩石样本进行识别,引入残差机制提高识别准确率,能有效识别不同岩石,但对一些较为相似的岩石区分不够理想;Zhou等^[8]提出基于卷积神经网络的下一代CNN网络,对香港地区常见的7类岩石实现自动分类,取得了较好的效果;袁硕^[9]虽然对基于ShuffleNetV2的深度可分离卷积神经网络进行了深入研究,但对于背景较为复杂的图像分类问题尚未给予充分考虑。

岩心图像的岩性特征众多,包括颜色、结构与构造等^[10]。传统的岩性识别方法过于依赖地质研究人员的经验和地质专业知识积累,准确率受到影响,因此,本文提出了一种针对岩心图像分类的改进网络,该网络基于ResNet50网络,融合了DCNv2(Deformable Convolutional Networks version 2)的可变形卷积技术、ECA(Efficient Channel Attention)注意力机制以及PSA(Pyramid Scene Attention)注意力机制。DCNv2根据输入特征动态调整卷积核,增强了对复杂图像的处理能力;ECA注意力机制则通过高效计算通道权重,提升模型稳定性和可靠性;PSA机制通过并行处理图像的不同空间位置,实现对局部区域的选择性关注,加速了模型的训练和推理过程,利用迁移学习提高了训练效率。

1 ResNet50 网络

1.1 ResNet50 网络结构

ResNet50是残差网络系列架构的重要成员,是目前图像识别领域使用最广泛的模型之一,在分类、对象检测及语义分割等任务中表现尤为突出,有效克服了梯度消失难题。该模型以全局平均池化层作为输出前的关键处理步骤,并利用ImageNet等大规模数据集进行预训练,以获取初始权重。预训练的ResNet50架构划分为4个主要部分,每部分称为一

个阶段,各自承担不同的功能(如图1所示)。

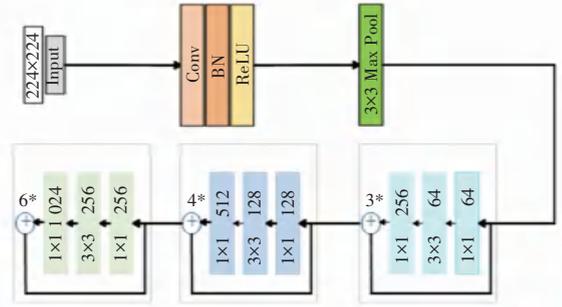


图1 ResNet50 结构示意图

Fig. 1 ResNet50 architecture

ResNet50在第一阶段使用了一个 7×7 的卷积核结合 3×3 的最大池化操作,有效缩小了输入图像尺寸,为后续处理提供了更紧凑的数据表示;在第二阶段,通过一系列创新的残差结构设计,即Conv2、Conv3、Conv4及Conv5等残差块,模型能够深入探索图像的高级特征;最终,这些高度抽象的特征被送入到第三阶段的全连接层中,完成最终的分类任务。

1.2 残差网络

ResNet残差网络通过跨层连接解决了梯度难题,这种结构允许输入直接跳过某些层,与经过处理的输出相加,形成“残差”概念。残差网络结构如图2所示,假设输入图像为 x ,输出为 $H(x)$,中间经过卷积之后的输出为 $F(x)$ 的非线性函数,最终的输出为 $H(x) = F(x) + x$,也就是 $F(x)$,而网络也就转化为求残差函数 $F(x) = H(x) - x$,这样残差函数要比 $F(x) = H(x)$ 更加容易优化。ResNet残差网络中引入残差结构,通过堆叠多个残差单元构建不同深度和复杂度的网络以适应任务需求,在各种图像相关任务中取得显著性能提升。

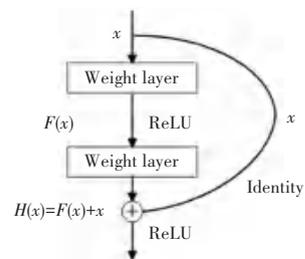


图2 残差网络结构图

Fig. 2 Residual structure diagram

2 改进的 ResNet50 网络

2.1 改进的 ResNet50 网络结构

ResNet50由多个残差块组成,每个块包含多个卷积层和恒等映射或者 1×1 卷积作为跳跃连接。因此,选择ResNet50网络作为主网络,其深度架构

能充分学习岩心图像中的复杂特征,包括纹理和细节等,能根据岩心图像的独特需求进行灵活调整和改进,更好地处理和理解岩心图像的地质信息。

本文在原 ResNet50 网络基础上进行了改进,改进的 ResNet50 网络结构如图 3 所示。以 224×224 像素的岩心图像作为网络的输入,使用 Batch Normalization 批量归一化层加快收敛速度,并使用 ReLU 激活函数增加非线性,通过最大池化层进行采样,然后依次通过 4 个 ResNet50 Bottleneck 残差网络,分别包含 3、4、6、3 个残差结构。针对每个 ResNet50 Bottleneck 模块内部,输入特征图先经过一个 1×1 的卷积层,调整特征图的通道数;经过一个 3×3 的 DCNv2 层进行多尺度特征提取和空间变换;再次经过一个 1×1 的卷积层,进一步调整通道数或融合特征;通过 ECA 注意力机制进行通道维度上的注意力机制增强;通过 PSA 注意力机制提高特征的

分层注意力;最后,所有处理后的特征通过一个加法操作进行融合,形成最终的输出特征图。同样的,经过第二组、第三组和第四组残差块后,进行全局平均池化,将特征图缩小为 1×1 ;最后通过一个全连接层 (FC) 进行分类,并通过 Softmax 输出最终的结果。

迁移学习作为一种高效的学习策略,允许将从源任务(如 ImageNet 图像分类)中获取的丰富知识迁移并应用于相关的目标任务(如岩心分类),显著加速了新任务的适应过程并优化了性能^[11]。利用预训练模型作为起点,不仅确保了模型在目标数据集上参数的有效性与合理性,还大大提升了训练效率,使得模型能够更快地收敛至较优状态。此外,迁移学习还有效减轻了过拟合风险,增强了模型的泛化能力。因此,本文融合了迁移学习策略,以加速并优化岩心分类模型的训练过程。

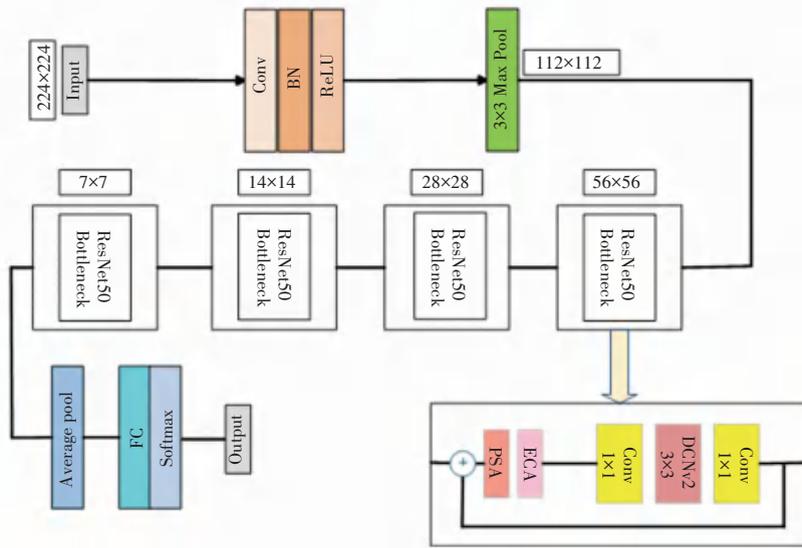


图 3 改进的 ResNet50 网络结构

Fig. 3 Improved ResNet50 Network Architecture

2.2 通道注意力机制

通道注意力机制在深度卷积神经网络 (CNN) 的性能提升中扮演着至关重要的角色。特别是 ECA 注意力机制,显著增强了 CNN 在特征提取方面的效能,为网络性能的优化提供了有力支持。

ECA 模块利用共享参数的 1 维卷积(具体为 $1 \times k$ 形式,而非常规的 1×1 卷积)来促进通道间的信息流通,减少了参数数量同时保持了高效的特征整合能力。ECA 通道注意力结构示意图如图 4 所示。

在 ResNet50 中引入 ECA 能提升特征提取能力,增强重要通道特征,弱化次要通道,以少量计算量提升性能,可轻松融入架构并优化模型。

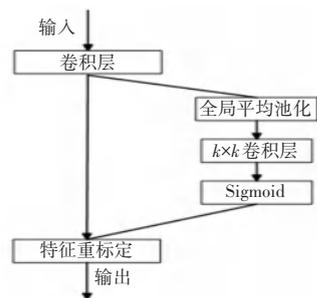


图 4 ECA 注意力机制结构

Fig. 4 ECA Attention Structure

2.3 空间注意力机制

PSA 注意力机制作为一种有效的空间注意力机

制,能够更好地捕捉到场景中的各种尺度信息,有效抑制背景噪声,克服传统空间注意力机制的局限性,

提高模型的鲁棒性和识别能力。PSA 注意力机制结构如图 5 所示。

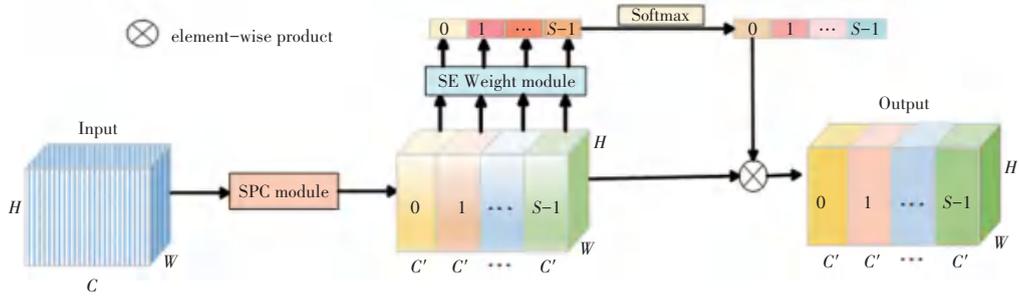


图 5 PSA 注意力机制结构

Fig. 5 PSA Attention Structure

在处理复杂背景和多尺度目标时,PSA 注意力机制表现出更强的灵活性,显著提升了模型的整体性能,尤其与 ResNet 等主流网络结合使用时效果更为明显。

2.4 可变形卷积模块

DCNv2 (Deformable Convolutional Networks version 2) 是基于可变形卷积网络的一种改进版本,在传统的卷积神经网络中引入了可变形卷积层,提高网络对图像中目标形状变化的适应能力。

DCNv2 网络结构如图 6 所示,由两大组件构成:偏移量生成模块与采样卷积模块。前者通过特定卷积操作,精准计算出卷积核在输入特征图上的水平(x轴)与垂直(y轴)方向上的采样点偏移量;后者则利用这些偏移信息及输入特征图,执行双线性插值操作,以精确定位卷积核的采样点,进而完成高效且灵活的卷积处理;最后,利用这些采样点进行卷积操作。

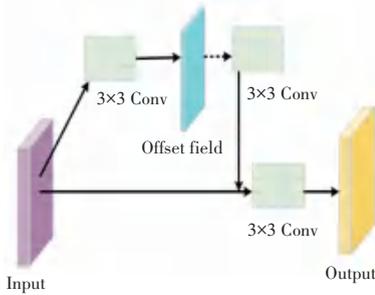


图 6 DCNv2 网络结构图

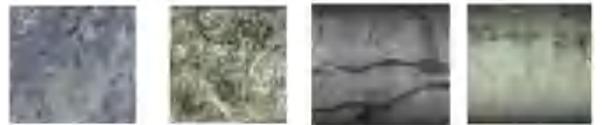
Fig. 6 DCNv2 network architecture diagram

3 实验结果与分析

3.1 数据集

本文实验所用的岩心薄片数据集来源于所在课题组自行构建的一个用于分类的岩心图像数据集。由于目前网络上缺少公开的岩心图像分类数据集,

该数据集是通过所在课题组与石油地质单位在若干油田勘探现场使用自行研制的数字化岩心图像采集系统,对采集到的图片进行分段裁剪、人工去除以及整理分类等操作得到一个岩心图像分类数据集,部分岩心图像数据样本示例如图 7 所示。



(a) 砂卡岩 (b) 灰白色粗粒 (c) 粉砂岩 (d) 白云岩正长岩

图 7 部分岩心图像数据集示例

Fig. 7 Partial core image dataset sample

该数据集共有 6 820 张图像,有 17 种不同种类的岩心,将每种岩心图像按 8 : 2 的比例随机划分为训练集和测试集,见表 1。

表 1 岩心图像样本分类及数量

Table 1 Classification and quantity of core image samples

类别名称	训练样本数/张	测试样本数/张
石英脉	45	12
白云岩	307	77
铅矿体	400	100
硅化岩	400	100
中细粒黑云辉石正长岩	320	81
灰白色粗粒正长岩	400	99
糜棱岩	100	25
粉砂岩	395	99
灰岩	399	100
灰白色正长斑岩	452	114
闪长岩	364	92
板岩	295	74
砂卡岩	374	94
细砂岩	400	100
辉绿岩	145	37
白云母花岗岩	400	100
二云母花岗岩	256	64

3.2 参数设置与评价指标

本文的实验环境为 Ubuntu 22.04.2 LTS 系统,处理器为 Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz, GPU 为 NVIDIA GeForce RTX2080Ti, 软件环境为 CUDA11.1, 深度学习框架为 Pytorch1.8.1。

先将图像调整为 256×256 像素,防止过多的计算量和显存占用,满足模型训练要求,并从中心截取 224×224 像素的区域,用随机旋转和水平反转的方式进行数据增强。

本文的 ResNet50 网络采用 SGD (Stochastic Gradient Descent) 优化算法,设定学习率为 0.01,权重衰减系数为 0.0001,动量系数为 0.9,模型数据批次大小为 32,训练轮次 (Epoch) 为 100。

鉴于岩心图像样本间存在数量不均衡现象,本文采用平均准确率 (AA) 作为评估模型性能的指标,综合考量数据集中各类别 (K 类) 的预测表现,定义如下式:

$$AA = \frac{1}{K} \sum_{i=1}^K \left(\frac{Y_i}{N_i} \right) \quad (1)$$

其中, N_i 代表第 i 类岩心图像的总样本数, Y_i 为该类中被正确分类的样本数。

在样本分布不均的情况下,平均准确率相较于总体准确率 (OA) 能提供更均衡、全面的评估视角,其值越高,表明模型分类效果越佳。

3.3 公共数据集对比实验

为了验证模型的泛化能力,本文选取了南京大学提供的教学薄片显微镜下的岩石图像数据集,此数据集详尽地分为火成岩、变质岩与沉积岩 3 个主要类别,具体包含火成岩样本 963 幅,变质岩样本 972 幅以及沉积岩样本 688 幅。这些岩石薄片样本均在单偏光和正交偏光两种独特的照明条件下拍摄,捕捉更丰富的岩石微观结构与特征。岩石薄片样本示例见图 8。



图 8 岩石薄片样本示例

Fig. 8 Rock thin section sample examples

针对沉积岩图像资源较少的情况,本文采取水平镜像、色彩调整、图像模糊化及噪声添加等多种技术手段进行数据扩增。平衡了数据集中各类岩石图像的数量分布,数据扩增前后的数量变化见表 2。

表 2 数据扩增前后对比

	火成岩	变质岩	沉积岩	总计
数据扩增前	963	972	688	2 623
数据扩增后	963	972	957	2 892

本文按 8:2 的比例划分数据集,即 2 313 张图像的训练集和 579 张图像的测试集。选择一系列主流的卷积网络,包括 VGG16、ResNet34、ResNet50、ResNet101、DenseNet169、EfficientNet、MobileViT、EfficientNetv2、DaViT 以及 MixMAE 进行迁移学习,并在相同的实验条件下进行训练并测试,所有网络在训练之前均加载了预训练权重。本文在南京大学岩石样本分类数据集上的对比实验结果见表 3。

表 3 不同网络在南京大学岩石样本数据集的对比试验指标

Table 3 Comparative evaluation metrics of different networks on the Nanjing University Rock Sample Dataset

网络模型	平均准确率/%
VGG16 ^[12]	96.53
ResNet34 ^[13]	96.70
ResNet50 ^[13]	97.88
ResNet101 ^[13]	96.67
DenseNet169 ^[14]	90.45
EfficientNet ^[15]	95.31
MobileViT ^[16]	98.13
EfficientNetv2 ^[17]	97.91
DaViT ^[18]	90.27
MixMAE ^[19]	97.92
改进的 ResNet50 ^[20]	98.25
本文改进的网络模型	98.56

从表 3 中可知,本文改进的网络平均准确率相比改进之前的 ResNet50 基准网络效果更优;与其他主流分类网络相比,本文网络平均准确率也更高。

3.4 岩心图像分类数据集对比试验

为了验证本文改进网络的有效性,将本文所改进的网络与上述网络在岩心图像分类数据集上进行训练,并将所有训练完成的网络使用同一测试集进行测试,实验结果见表 4。

由表 4 可知,本文改进的网络在岩心图像分类任务中表现出色,显示出较强的分类能力,比 ResNet50 网络提高了 2.33%,并且相比其他主流分类网络也具有更高的平均准确率。VGG16 网络由于网络深度相对较浅,难以捕捉岩心图像的更复杂的特征;ResNet 系列模型随着网络深度的增加,

ResNet50 和 ResNet101 识别的准确率有所上升; DenseNet169 网络以其密集连接特性提升了信息流通,但也带来了训练挑战; EfficientNet 则偏向轻量级,面对岩心数据集的细粒度分类时,特征提取可能受限; EfficientNetV2 网络通过优化结构与参数,克服了轻量级网络在细粒度分类任务中特征提取能力不足的问题; MobileViT 结合了卷积和 Transformer 的特性,能够在保持计算效率的同时,提高特征提取能力,达到了较高的准确率; DaViT 侧重于计算效率和特定任务适应性,准确率相对较低; MixMAE 模型在岩心图像岩性识别任务上学习到丰富的特征表示; 本文改进的网络在所有模型中表现最佳,能够更好地捕捉岩心图像中的复杂特征。

表 4 不同网络在岩心图像分类数据集上的指标对比

Table 4 Comparison of different networks on the core image classification dataset

网络模型	平均准确率/%
VGG16 ^[12]	89.23
ResNet34 ^[13]	88.76
ResNet50 ^[13]	90.81
ResNet101 ^[13]	90.77
DenseNet169 ^[14]	89.96
EfficientNet ^[15]	89.14
MobileViT ^[16]	90.85
EfficientNetv2 ^[17]	86.09
DaViT ^[18]	80.68
MixMAE ^[19]	89.81
改进的 ResNet50 ^[20]	92.89
本文改进的网络模型	93.14

3.5 消融实验

本文改进的网络主要包括 3 个模块,即 ECA 注意力机制、PSA 注意力机制、DCNv2 可变形卷积模块。消融实验共分为 5 组,第一组实验使用 ResNet50 网络,第二组实验在 ResNet50 网络的基础上添加 ECA 注意力机制,第三组实验在 ResNet50 网络的基础上添加 PSA 注意力机制,第四组实验是在 ResNet50 网络的基础上添加 DCNv2 卷积模块,第五组实验是在 ResNet50 网络的基础上添加 ECA 注意力机制、PSA 注意力机制和 DCNv2 可变形卷积模块,实验所使用的训练参数一致。在岩心图像分类数据集上的消融实验结果见表 5,可见各模块改进均有效提升模型性能,融合 3 个模块后网络精度最高,证明了本文改进的网络在岩心图像分类上的有效性。

表 5 不同模块在岩心图像分类数据集上的消融实验

Table 5 Ablation study of different modules on the core image classification dataset

网络模型	平均准确率/%
ResNet50	90.81
ResNet50 + ECA	90.92
ResNet50 + PSA	91.82
ResNet50 + DCNv2	90.97
本文提出的模型	93.14

4 结束语

本文聚焦于岩心图像分类识别,提出了一种优化 ResNet50 的网络。该算法融合了 ECA 与 PSA 注意力机制,显著增强了网络捕捉岩心图像中复杂地质特征的能力。同时,引入 DCNv2 可变形卷积,使模型能够灵活应对图像特征的多样性和形变,大幅提高了对复杂岩心结构的识别精度。通过迁移学习策略,利用 ImageNet 预训练的 ResNet50 网络参数作为起点,对岩心图像数据集进行微调,不仅加速了训练过程,还增强了模型的泛化能力。

实验结果表明,改进的 ResNet50 网络在岩心图像分类任务上表现优异,相较于其他主流卷积网络,平均准确率明显提升。消融实验验证了各改进模块的有效性,特别是同时引入 ECA、PSA、DCNv2 模块时,能够提升模型性能。

未来的工作中,考虑到地质样本的多样性和复杂性,还计划拓展数据集,包括不同地区和不同类型的岩心样本。

参考文献

- [1] LONG J, SHELHAMBER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2015: 3431-3440.
- [2] 王浩,宋文广,徐浩,等.基于支持向量机的岩心识别算法研究[J].湖北工程学院学报,2021,41(6):61-66.
- [3] ZHANG P Y, SUN J M, JIANG Y J, et al. Deep learning method for lithology identification from borehole images [C]//Proceedings of the European Association of Geoscientists & Engineers. 2017: 1-5.
- [4] BARABOSHKIN E, ISMAILOVA L S, ORLOV D M, et al. Deep convolutions for in-depth automated rock typing [J]. Computers & Geosciences, 2020(135): 104330.
- [5] ALZUBAIDI F, MOSTAGHIMI P, SWIETOJANSKI P, et al. Automated lithology classification from drill core images using convolutional neural networks[J]. Journal of Petroleum Science and Engineering, 2021(197): 107933.
- [6] 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[J]. 应用基础与工程科学学报, 2021, 29(5): 1075-1092.
- [7] 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.
- [8] ZHOU Y, WONG L N Y, TSE K C. Novel rock image classification: the proposal and implementation of HKUDES_Net [J]. Rock Mechanics and Rock Engineering, 2023, 56(5): 3825-3841.
- [9] 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.
- [10] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018 (77): 354-377.
- [11] PENG S, JIANG H, WANG H, et al. Modulation classification based on signal constellation diagrams and deep learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 30(3): 718-727.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large - scale image recognition [J]. Computer Science, 2014, 34(2): 1409-1422.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 770-778.
- [14] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 4700-4708.
- [15] TAN M. Efficientnet: Rethinking model scaling for convolutional neural networks[J]. arXiv preprint arXiv, 1905.11946, 2019.
- [16] MEHTA S, RASTEGARI M. Mobilevit: Light-weight, general-purpose, and mobile - friendly vision transformer [J]. arXiv preprint arXiv, 2110.02178, 2021.
- [17] TAN M, LE Q. Efficientnetv2: Smaller models and faster training [C]//Proceedings of the 38th International Conference on Machine Learning. IMLS, 2021: 10096-10106.
- [18] DING M, XIAO B, CODELLO N, et al. Davit: Dual attention vision transformers [C]//Proceedings of the European Conference on Computer Vision. Cham: Springer, 2022: 74-92.
- [19] LIU J, HUANG X, ZHENG J, et al. MixMAE: Mixed and masked autoencoder for efficient pretraining of hierarchical vision transformers [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2023: 6252-6261.
- [20] 曹迁义. 岩心图像处理与内容分析关键技术的研究与实现 [D]. 成都: 四川大学, 2022.