



OPEN ACCESS

EDITED BY
Abbas Maghsoudi,
Amirkabir University of Technology, IranREVIEWED BY
Kalpana Bhatt,
Purdue University, United States
Sadegh Karimpouli,
University of Zanjan, Iran*CORRESPONDENCE
Lumei Su,
sulumei@163.comSPECIALTY SECTION
This article was submitted to
Environmental Informatics
and Remote Sensing,
a section of the journal
Frontiers in Earth ScienceRECEIVED 25 October 2022
ACCEPTED 22 November 2022
PUBLISHED 16 January 2023CITATION
Chen W, Su L, Chen X and Huang Z
(2023). Rock image classification using
deep residual neural network with
transfer learning.
Front. Earth Sci. 10:1079447.
doi: 10.3389/feart.2022.1079447COPYRIGHT
© 2023 Chen, Su, Chen and Huang. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

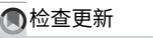
Rock image classification using deep residual neural network with transfer learning

Weihao Chen¹, Lumei Su^{1,2*}, Xinqiang Chen¹ and Zhihao Huang¹¹School of Electrical Engineering and Automation, Xiamen University of Technology, Xiamen, China,
²Xialong Institute of Engineering and Technology, Longyan, China

Rock image classification is a significant part of geological research. Compared with traditional image classification methods, rock image classification methods based on deep learning models have the great advantage in terms of automatic image features extraction. However, the rock classification accuracies of existing deep learning models are unsatisfied due to the weak feature extraction ability of the network model. In this study, a deep residual neural network (ResNet) model with the transfer learning method is proposed to establish the corresponding rock automatic classification model for seven kinds of rock images. ResNet34 introduces the residual structure to make it have an excellent effect in the field of image classification, which extracts high-quality rock image features and avoids information loss. The transfer learning method abstracts the deep features from the shallow features, and better express the rock texture features for classification in the case of fewer rock images. To improve the generalization of the model, a total of 3,82,536 rock images were generated for training via image slicing and data augmentation. The network parameters trained on the Texture Library dataset which contains 47 types of texture images and reflect the characteristics of rocks are used for transfer learning. This pre-trained weight is loaded when training the ResNet34 model with the rock dataset. Then the model parameters are fine-tuned to transfer the model to the rock classification problem. The experimental results show that the accuracy of the model without transfer learning reached 88.1%, while the model using transfer learning achieved an accuracy of 99.1%. Aiming at geological engineering field investigation, this paper studies the embedded deployment application of the rock classification network. The proposed rock classification network model is transplanted to an embedded platform. By designing a rock classification system, the off-line rock classification is realized, which provides a new solution for the rock classification problem in the geological survey. The deep residual neural network and transfer learning method used in this paper can automatically classify rock features without manually extracting. These methods reduce the influence of subjective factors and make the rock classification process more automatic and intelligent.

KEYWORDS

deep learning, image recognition, rock classification, transfer learning, convolutional neural network



开放获取

EDITED BY
阿巴斯·马格苏迪,
伊朗阿米尔卡比尔理工大学审阅人
卡尔帕纳·巴特,
美国普渡大学Sadegh
Karimpouli, 伊朗赞詹大学*通讯作者
苏璐梅,
sulumei@163.com专业章节
本文投稿至期刊《地球科学
前沿》的“环境信息与遥
感”专栏收稿日期：2022年10月25日
收稿日期2022年11月22日
刊发日期2023年1月16日CITATION
陈伟、苏力、陈晓与黄志(2023)撰写的
《基于深度残差神经网络迁移学习的
岩石图像分类》发表于《地球科学前
沿》第10卷, 文献编号1079447。doi:
10.3389/feart.2022.1079447COPYRIGHT
© 2023 陈、苏、陈与黄。本文依据知
识共享署名许可协议(CC BY)开放获
取, 允许在其他论坛使用、传播或复
制, 但须署名原始作者及版权所有
者, 并引用本刊原始出版物, 符合学
术规范。未遵守此条款的使用、传播
或复制均不被允许。

基于迁移学习的深度残差神经网络岩石图像分类方法

陈伟豪¹、苏露梅^{1,2*}、陈新强¹、黄志豪¹¹ 厦门理工学院电气工程与自动化学院, 中国厦门, ² 龙岩下陇工程技术研究院, 中国龙岩

岩石图像分类是地质研究的重要组成部分。与传统图像分类方法相比, 基于深度学习模型的岩石图像分类方法在自动特征提取方面具有显著优势。然而, 由于网络模型特征提取能力不足, 现有深度学习模型的岩石分类精度尚不理想。本研究提出采用深度残差神经网络(ResNet)结合迁移学习方法, 针对七类岩石图像建立自动化分类模型。ResNet34通过引入残差结构, 在图像分类领域表现优异, 既能提取高质量的岩石图像特征, 又可避免信息丢失。迁移学习方法通过从浅层特征中抽象深层特征, 在岩石图像样本较少的情况下更好地表征岩石纹理特征用于分类。为提升模型泛化能力, 通过图像切片和数据增强共生成382,536张岩石图像用于训练。网络参数先在包含47类纹理图像且能反映岩石特征的纹理库数据集上进行预训练, 随后将预训练权重加载至ResNet34模型进行岩石数据集训练, 并通过微调模型参数实现向岩石分类任务的迁移。实验结果表明: 未采用迁移学习的模型准确率为88.1%, 而采用迁移学习的模型准确率达99.1%。针对地质工程野外调查需求, 本文研究了岩石分类网络的嵌入式部署应用, 将所提模型移植至嵌入式平台, 通过设计岩石分类系统实现离线岩石分类, 为地质勘察中的岩石分类问题提供了新解决方案。本文采用的深度残差神经网络与迁移学习方法无需人工提取特征即可实现岩石自动分类, 有效降低了主观因素影响, 使岩石分类过程更具自动化与智能化。

KEYWORDS

deep learning, image recognition, rock classification, transfer learning, convolutional neural network

1 Introduction

Rock classification is the basis for studying geological reservoir characteristics and plays an essential role in vast fields, such as geotechnical engineering, mineralogy, petrology, rock mechanics, and mineral resource prospecting (Karimpouli and Tahmasebi, 2019; Guo et al., 2022; Houshmand et al., 2022). The efficiency of rock classification is closely associated with the efficiency of geological surveys and therefore needs urgent attention. Rocks classification can be accomplished *via* traditional methods, including remote sensing, electromagnetic field, geochemistry, hand specimen and thin section analysis (Ru and Jiong, 2019). These traditional methods are based on human observation, manual operation and empirical classification. Rock classification using traditional methods mainly extracts useful information features from rock images by professionals through specialized equipment, relying on people's experience and equipment sensitivity. These methods are often limited by the professionalism of experimental equipment and the theoretical level of researchers, resulting in much time spent, low efficiency and many other problems.

Rock classification using traditional machine learning methods usually need to manually design feature extraction methods and input rock features into the classifier for training, to realize rock classification. Singh et al. (2010) used a multilayer perceptron to extract 27 features from basalt rock slice images and achieved the classification of 140 rock sample slice images. Gonçalves and Leta (2010) proposed a neuro-fuzzy hierarchical classification method based on binary space division for macroscopic rock structure classification, and the final classification accuracy reached 73%. Mlynarczuk et al. (2013) used the nearest neighbor algorithm and k-nearest neighbor algorithm to realize the classification of 9 different types of rocks. Sharif et al. (2015) proposed an autonomous rock classification system based on Bayesian image analysis for planetary geological exploration. The rock sample surface was described by 13 Haralick texture parameters and the information was automatically catalogued into a 5-bin data structure, then the Bayesian probability was calculated and the recognition result was output. Patel and Chatterjee (2016) realized the classification of limestone by extracting color, shape and texture features from limestone images and inputting them into a probabilistic neural network. Wang and Sun (2021) proposed a rock classification method using geometric features of rock particles instead of local structural features, which effectively solved the problem of fuzzy boundaries.

With the development of artificial intelligence, machine learning and deep learning are widely used in various image classification problems. Since traditional machine learning need to manually extract rock features from a huge training dataset, the training work is difficult and rather laborious. Using deep learning methods to construct automatic rock classification models has become a new way for rock classification (Fan

et al., 2020; Falivene et al., 2022). Cheng et al. (2017) proposed an automatic rock grain size classification method based on the convolutional neural network. The convolutional neural network was trained with 4,800 samples from the Ordos Basin, which contains three categories, and the classification of rock slice samples under the microscope was realized. But its image data are thin sections of rock casts taken under a polarizing microscope, and the production of the data set is relatively complex and not easy to obtain. Based on the Inception-v3 network model, Zhang et al. (2018) used transfer learning to establish a classification model of rock images, which could identify and classify three types of rocks with obvious characteristics: granite, breccia and phyllite, and the accuracy of test data reached more than 85%. Bai et al. (2018) built a deep learning model for rock recognition based on the convolutional neural network and trained it on 1,000 rock pictures collected on the network or taken in real life, achieving a recognition accuracy of 63%. Bai et al. (2019) also used the VGG network model to establish a rock slice image recognition model to classify rock slice images of six common rocks such as granite and dolomite, and the recognition accuracy reached 82%. Imamverdiyev and Sukhostat, 2019 proposed a new 1D-CNN model trained on multiple optimization algorithms, which is suitable for the lithofacies classification of complex landforms. Shuteng and YongZhang (2018) designed a targeted U-net convolutional neural network model to automatically extract deep feature information of minerals under the mineral phase microscope and realize under-mirror ore mineral intelligent recognition and classification. Feng et al. (2019) established a rock recognition model based on the AlexNet twin convolutional neural network for fresh rock sections. Its advantage lies in the comprehensive consideration of global image information and local texture information of rocks, but its disadvantages are the large model and the lack of high classification accuracy. Hu et al. (2020) trained a lithology recognition model with an accuracy of 90% by applying image data in big geological data and based on deep learning. Zeng et al. (2021) used a two-layer fully connected neural network to increase the dimension of the scalar Mohs hardness, and used EfficientNet-b4 to extract the feature of the ore image, then fused the results of the two layers and finally sent them into the fully connected layer to complete the classification of 36 different types of ores. Liang et al. (2021) first used a ViT network structure that evolved from transformers to classify seven different types of ores. Koeshidayatullah et al. (2022) proposed a novel FaciesViT model based on the transformer framework for automatic core facies classification, which is much better than CNN and hybrid CNN-ViT models, and does not require preprocessing and feature extraction. In addition to rock images of natural scenes, many scholars also use microscopic rock images and spectral images for rock classification. Iglesias et al. (2019) used ResNet18 to classify the polarized light microscopic images of five ores, including amphibole, quartz, garnet, biotite, and olivine. The final model accuracy reached

等人, 2020; Falivene等人, 2022)。Cheng等人 (2017) 提出了一种基于卷积神经网络的岩石粒度自动分类方法。该卷积神经网络使用鄂尔多斯盆地的4800个样本进行训练, 包含三个类别, 实现了显微镜下岩石薄片样本的分类。但其图像数据为偏光显微镜下的岩石铸体薄片, 数据集制作较为复杂且不易获取。Zhang等人 (2018) 基于Inception-v3网络模型, 采用迁移学习建立了岩石图像分类模型, 可识别并分类花岗岩、角砾岩和千枚岩这三种特征明显的岩石类型, 测试数据准确率达85%以上。Bai等人 (2018) 构建了基于卷积神经网络的岩石识别深度学习模型, 在网络上收集或实地拍摄的1000张岩石图片上进行训练, 识别准确率达到63%。Bai等人 (2019) 还利用VGG网络模型建立了岩石薄片图像识别模型, 对花岗岩、白云岩等六种常见岩石的薄片图像进行分类, 识别准确率达82%。Imamverdiyev和Sukhostat (2019) 提出了一种基于多种优化算法训练的新型一维CNN模型, 适用于复杂地貌的岩相分类。Shuteng和YongZhang (2018) 设计了针对性的U-net卷积神经网络模型, 自动提取矿相显微镜下矿物的深层特征信息, 实现镜下矿石矿物的智能识别与分类。Feng等人 (2019) 针对新鲜岩石断面建立了基于AlexNet孪生卷积神经网络的岩石识别模型, 其优势在于综合考虑了岩石的全局图像信息和局部纹理信息, 但存在模型较大、分类精度不高的问题。Hu等人 (2020) 应用大地质数据中的图像数据, 基于深度学习训练了岩性识别模型, 准确率达90%。Zeng等人 (2021) 采用双层全连接神经网络对莫氏硬度标量进行升维, 利用EfficientNet-b4提取矿石图像特征, 将两者结果融合后送入全连接层, 完成36种不同类型矿石的分类。Liang等人 (2021) 首次使用由transformer演化的ViT网络结构对七种不同类型矿石进行分类。Koeshidayatullah等人 (2022) 提出基于transformer框架的新型

FaciesViT模型用于岩心相自动分类, 其性能显著优于CNN和混合CNN-ViT模型, 且无需预处理和特征提取。除自然场景的岩石图像外, 许多学者还利用显微岩石图像和光谱图像进行岩石分类。Iglesias等人 (2019) 使用ResNet18对闪石、石英、石榴石、黑云母和橄榄石等五种矿石的偏光显微图像进行分类, 最终模型准确率达到

(范

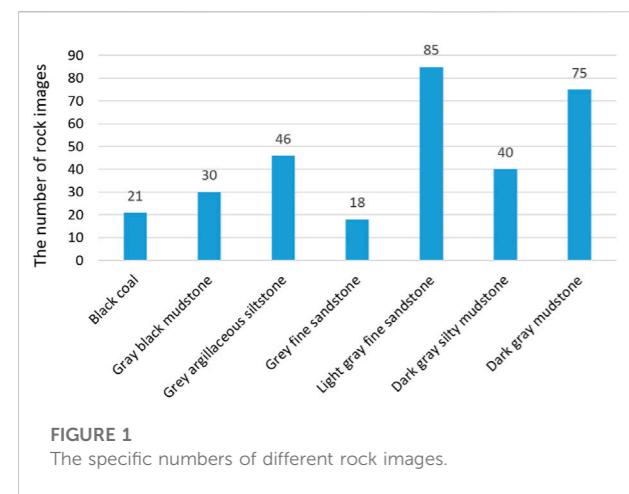


FIGURE 1
The specific numbers of different rock images.

89%. Xiao et al. (2021) first used the visible infrared reflectance spectrometer to obtain the spectral image of the ore, and then input it into the custom dilated convolutional neural network for training, and realized the classification of five kinds of ore such as hematite and magnetite.

Although the previous models have realized rock classification based on deep learning, the used models have redundancy and poor generalization. They can achieve low classification accuracy and do not consider the actual deployment and application of the model for geological exploration scenarios. To address these problems, a rock image classification method based on the pre-trained residual neural network (ResNet) by the way of transfer learning is proposed. ResNet can avoid feature loss of the convolution layer during information transmission, and can learn new features based on input features with better performance. In this study, ResNet is used to extract the deep feature information of rock images in order to classify all kinds of rocks. Transfer learning can reduce training time and consumption cost in the case of insufficient datasets, and achieve the goal of faster and better classification effect on small datasets. The texture feature is an important distinguishing point of all kinds of rocks. The Texture Library dataset is used to pre-train ResNet34 so that the model can extract texture features of rock images more quickly and effectively. The experimental results indicate that the model has high classification accuracy and good generalization ability. Finally, considering the application of geological surveys and construction sites, a rock classification system was developed. The rock classification model was deployed on the embedded device to achieve high accuracy of offline rock classification.

2 Materials

The rock dataset is provided by Guangdong TipDM Intelligent Technology Co., Ltd and includes the information

for 315 rock images. The rock samples were obtained by taking pictures of rock debris and drill core samples under the white light from an industrial camera at the mud logging site. The rock dataset consists of 7 categories of rock images: black coal, gray black mudstone, gray argillaceous siltstone, gray fine sandstone, light gray fine sandstone, dark gray silty mudstone and dark gray mudstone. The number of rock images varies by type and each image has dimensions of $4,096 \times 3,000$ pixels. Different types of rocks have slight differences in morphological characteristics. Sandstone is very small and contains a lot of sand grains. Mudstone is mostly lamellar and easily broken into fragments. The specific number is shown in Figure 1 and the corresponding characteristics of the seven rocks are shown in Table 1.

Datasets in deep learning are usually divided into the training set, validation set and test set, and different data subsets have different functions in model training. The training set is used to input data into the model to obtain results, then compare with the data labels to calculate the loss function, and finally update the parameters of the model through backpropagation to improve the performance of feature extraction and classification, so the training set accounts for the largest amount of data. The validation set is used to improve the training efficiency of the model. If the various hyperparameters are set or the model design is not reasonable when the model is under training, the model can respond to the accuracy of the validation set through the output, and then stop the training and make improvements in time. After the model is trained, the performance of the model can be evaluated using the test set. Similarly, the rock image dataset is randomly partitioned into the training set, validation set, and test set. If the ratio of the training set and validation set is too large, the model may overlearn and the model training time will grow, increasing the burden of model training, but a small ratio may also lead to model undertraining. The proportion of training, validation, and testing images in each label is set to 80%, 15%, and 5%, respectively. The dataset structure is shown in Table 2.

3 Methods

In order to fully extract the textural characteristics of different rocks, a rock image classification method based on the pre-trained residual neural network generated from transfer learning is proposed. Figure 2 presents the flowchart of the methods in this research.

3.1 The architecture of ResNet-34

Deep convolutional neural networks have made remarkable achievements in image classification, object detection, semantic segmentation and other fields. With the advancement of technology, more and more deep neural network models with better effects are constantly emerging

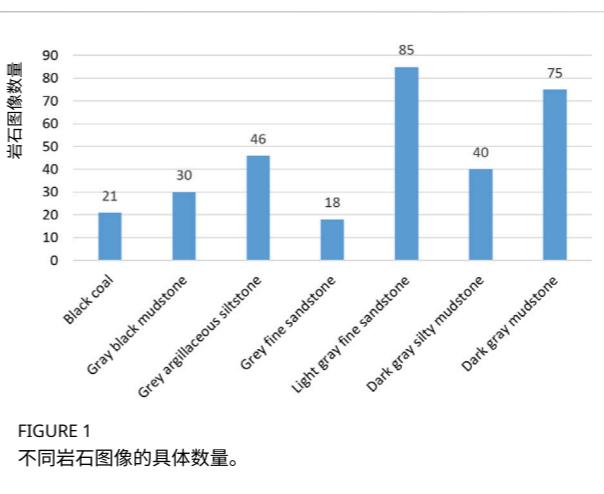


FIGURE 1
不同岩石图像的具体数量。

89%。肖等人（2021）首次利用可见光-红外反射光谱仪获取矿石光谱图像，将其输入定制扩张卷积神经网络进行训练，实现了赤铁矿、磁铁矿等五类矿石的分类识别。

尽管已有模型实现了基于深度学习的岩石分类，但所用模型存在冗余且泛化能力较差，分类精度较低，且未考虑地质勘探场景中模型的实际部署与应用。为解决这些问题，本文提出一种基于预训练残差神经网络(ResNet)的迁移学习岩石图像分类方法。ResNet能避免卷积层在信息传递过程中的特征丢失，并能基于输入特征学习新特征，性能更优。本研究利用ResNet提取岩石图像的深层特征信息，以实现各类岩石的精准分类。迁移学习可在数据集不足情况下减少训练时间与消耗成本，实现小数据集上更快、更好的分类效果。纹理特征是区分各类岩石的重要依据，采用纹理库数据集对ResNet34进行预训练，使模型能更快速有效地提取岩石图像的纹理特征。实验结果表明该模型具有较高分类精度和良好泛化能力。最后结合地质调查与施工现场应用需求，开发了岩石分类系统，将分类模型部署于嵌入式设备，实现了高精度的离线岩石分类。

2 材料

岩石数据集由广东泰迪智能科技股份有限公司提供，包含315幅岩石图像信息

。岩样通过工业相机在录井现场白光条件下拍摄岩屑和钻井岩芯获取。该数据集包含7类岩石图像：黑色煤、灰黑色泥岩、灰色泥质粉砂岩、灰色细砂岩、浅灰色细砂岩、深灰色粉砂质泥岩及深灰色泥岩。各类型岩石图像数量不等，单幅图像尺寸为 $4,096 \times 3,000$ 像素。不同岩石在形态特征上存在细微差异：砂岩颗粒极小且含大量砂粒，泥岩多呈层片状易碎成块。具体数量见图1，七类岩石对应特征详见表1。

深度学习中的数据集通常划分为训练集、验证集和测试集，不同数据子集在模型训练中具有不同功能。训练集用于向模型输入数据获取结果，通过与数据标签比对计算损失函数，最终通过反向传播更新模型参数以提升特征提取与分类性能，因此训练集数据占比最大。验证集用于提升模型训练效率，当模型训练时若各类超参数设置或模型设计不合理，可通过验证集准确率反馈输出，及时停止训练并改进。模型训练完成后，则通过测试集评估模型性能。同理，岩石图像数据集被随机划分为训练集、验证集和测试集。若训练集与验证集比例过大，模型可能过度学习且训练时长增加，加重训练负担；但比例过小又会导致模型欠学习。各标签中训练、验证与测试图像的占比分别设置为 80%，15% 和 5%，数据集结构如表2所示。

3 研究方法

为充分提取不同岩石的纹理特征，本研究提出基于迁移学习生成预训练残差神经网络的岩石图像分类方法。图2展示了本研究的流程图。

3.1 ResNet-34网络架构

深度卷积神经网络在图像分类、目标检测、语义分割等领域取得显著成就。随着技术进步，效果更优的深度神经网络模型不断涌现

TABLE 1 The characteristics of seven types of rocks.

Figure	Category	Color	Characteristic
	Black coal	Black	Lumpy, Granular
	Dark gray mudstone	Gray black	Argillaceous structure, Massive structure
	Dark gray silty mudstone	Dark gray	Silty and muddy structure, Bedding structure
	Gray black mudstone	Dark gray	Cryptocrystalline structure, Massive structure
	Grey argillaceous siltstone	Gray	Silty structure, Massive structure
	Grey fine sandstone	Gray	Fine grain structure, Massive structure
	Light gray fine sandstone	Light gray	Fine sand structure, Massive structure

表1 七类岩石的特征

图	类别	颜色	特征
	黑煤	黑色	块状、粒状
	深灰色泥岩	灰黑色	泥质结构，块状构造
	深灰色粉砂质泥岩	深灰色	粉砂质泥质结构，层理构造
	灰黑色泥岩	深灰色	隐晶质结构，块状构造
	灰色泥质粉砂岩	灰色	粉砂质结构，块状结构
	灰色细砂岩	灰色	细粒结构，块状结构
	浅灰色细砂岩	浅灰色	细砂结构，块状结构

TABLE 2 Details of the rock dataset.

Dataset	Number (Initial)	Number (After image slicing)	Number (After data augmentation)
Training	253	27,324	3,82,536
Validation	47	47	47
Test	15	15	15
Total	315	27,386	3,82,598

表2 岩石数据集详情

Dataset	编号 (初始)	编号 (图像切片后)	数量 (数据增强后)
训练集	253	27,324	3,82,536
验证集	47	47	47
测试集	15	15	15
总计	315	27,386	3,82,598

(Luo and Wang, 2021). However, it is found that not the deeper the number of network layers, the better the model effect. The increase in network depth not only does not make the accuracy achieved by the traditional network higher, but also produces problems such as gradient disappearance, gradient explosion, and degradation.

Residual neural networks (i.e., ResNet) enable feature information from the input or learned in the shallow layers of

(Luo与Wang, 2021)。但研究发现，并非网络层数越深模型效果越好。网络深度的增加不仅未能使传统网络达到更高准确率，反而还会导致梯度消失、梯度爆炸以及网络退化等问题。

残差神经网络（即ResNet）能够从输入或浅层学习到的特征信息

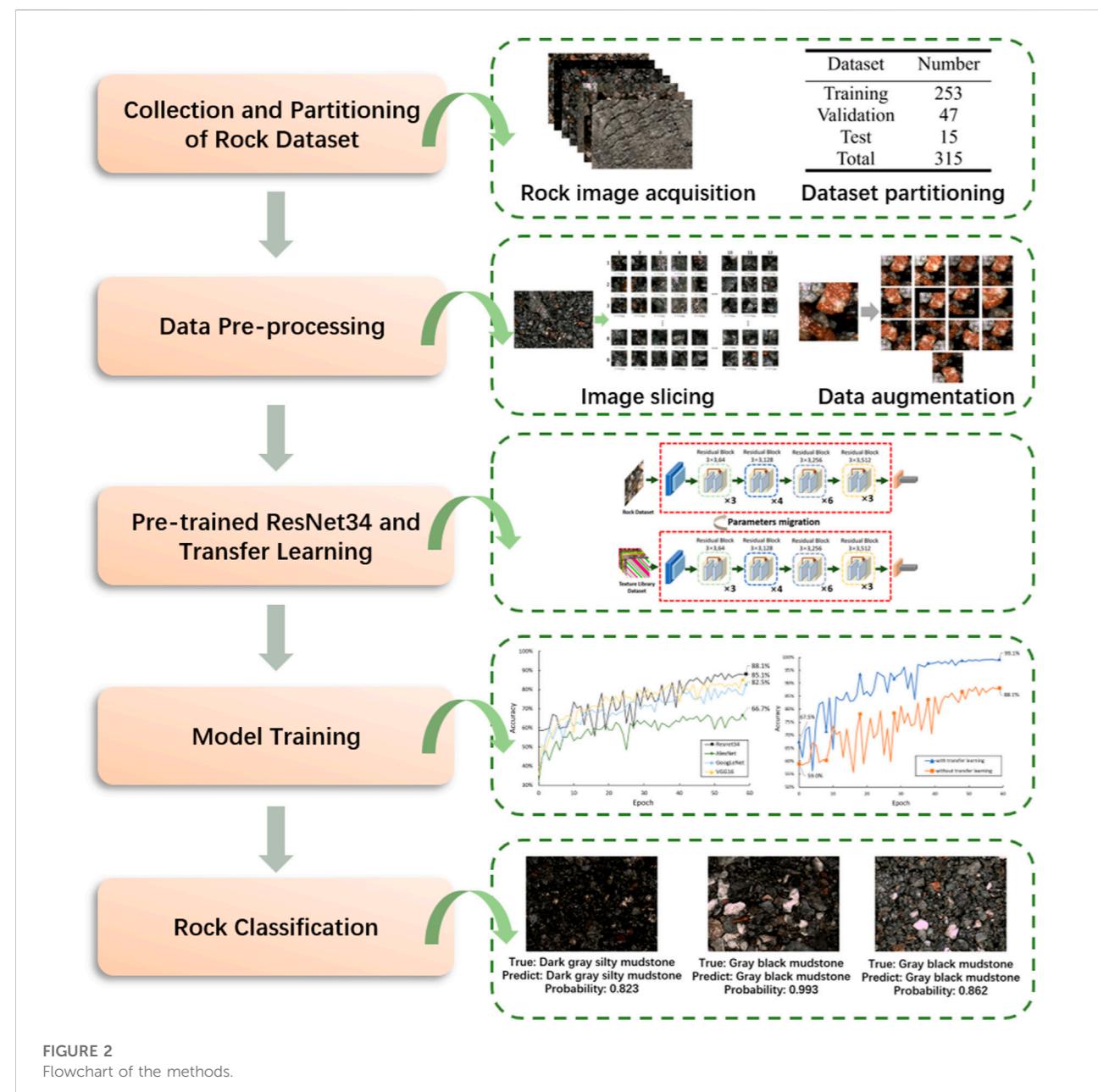


FIGURE 2
Flowchart of the methods.

the network to flow into the deeper layers by employing shortcut connections (He et al., 2016a). As the depth of the network increases, ResNet ensures the validity of gradient information by shortcut connections to prevent gradient disappearance and performance degradation caused by too deep layers of the network. Residual neural networks have achieved impressive results in image classification competitions such as ImageNet (He et al., 2016b) and MS COCO (Dai et al., 2016). In this study, ResNet is used to extract deep feature information from rock images to avoid the feature loss of the convolutional layer caused by gradient disappearance and gradient explosion in the process of information transmission.

$$H(x) = F(x) + x \quad (1)$$

Where x is the input data, $F(x)$ is the mapping function of the identity residuals and $H(x)$ is the mapped solution function.

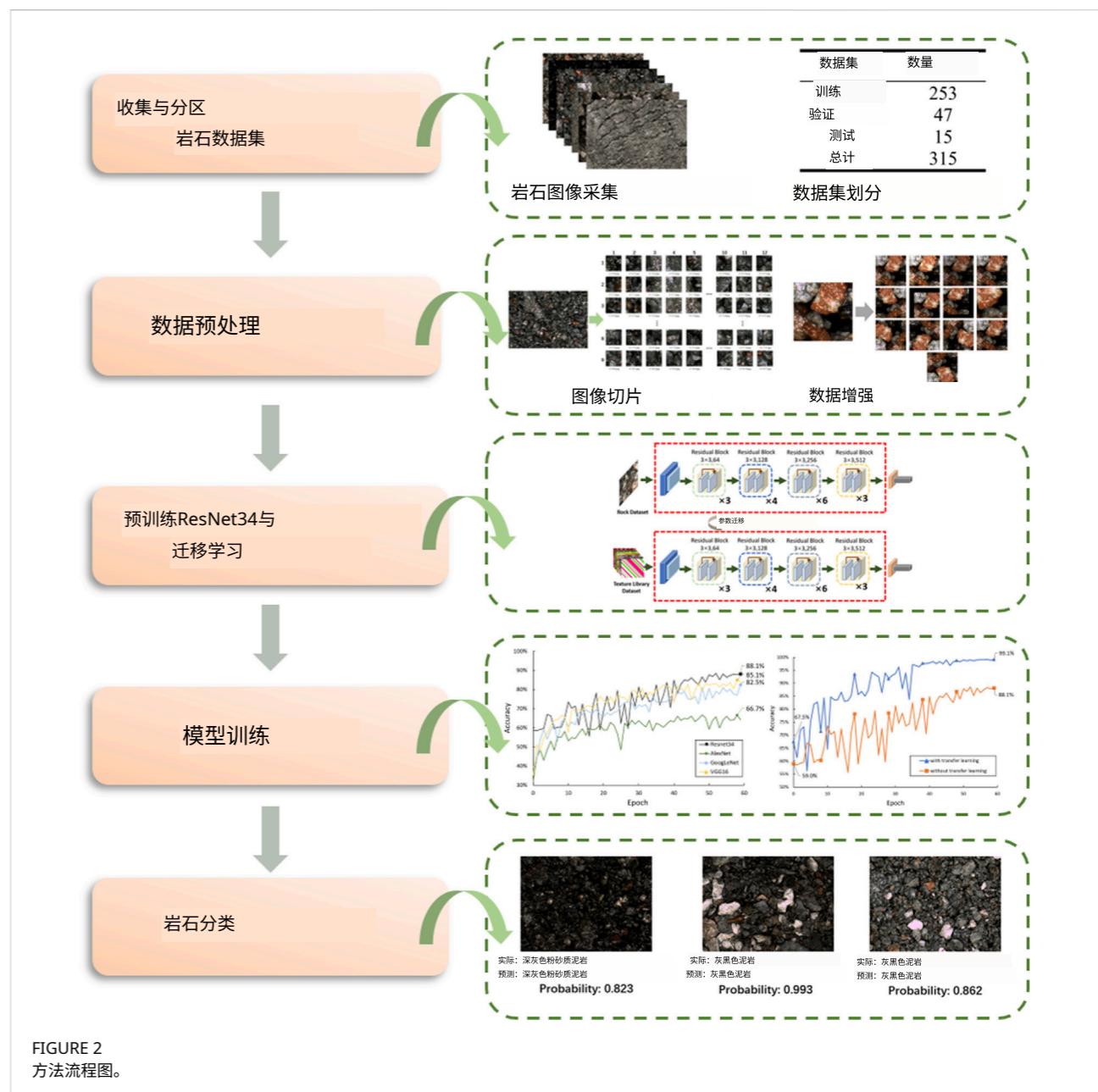


FIGURE 2
方法流程图。

ResNet consists of multiple residual blocks. The residual block not only has sequential convolutional layers, but also skips some convolutional layers through shortcut connections alongside the convolutional layers, and passes the data from the input residual block directly to the output, which is added with the result of the operation through the convolutional layer. Each residual structural unit can be defined as follows:

通过采用捷径连接(He等, 2016a), 网络得以向更深层流动。随着网络深度增加, ResNet 通过捷径连接确保梯度信息的有效性, 防止因网络层过深导致的梯度消失和性能下降。残差神经网络在ImageNet(He等, 2016b)和MS COCO(Dai等, 2016)等图像分类竞赛中取得了显著成果。本研究利用ResNet从岩石图像中提取深层特征信息, 避免信息传递过程中因梯度消失和梯度爆炸导致卷积层特征丢失。

ResNet由多个残差块组成。每个残差块不仅包含顺序排列的卷积层, 还通过并行于卷积层的捷径连接跳过若干卷积层, 将输入残差块的数据直接传递至输出端, 与卷积层运算结果相加。每个残差结构单元可定义为:

$$H(x) = F(x) + x \quad (1)$$

其中 x 为输入数据, $F(x)$ 是恒等残差的映射函数, $H(x)$ 为映射后的解函数。

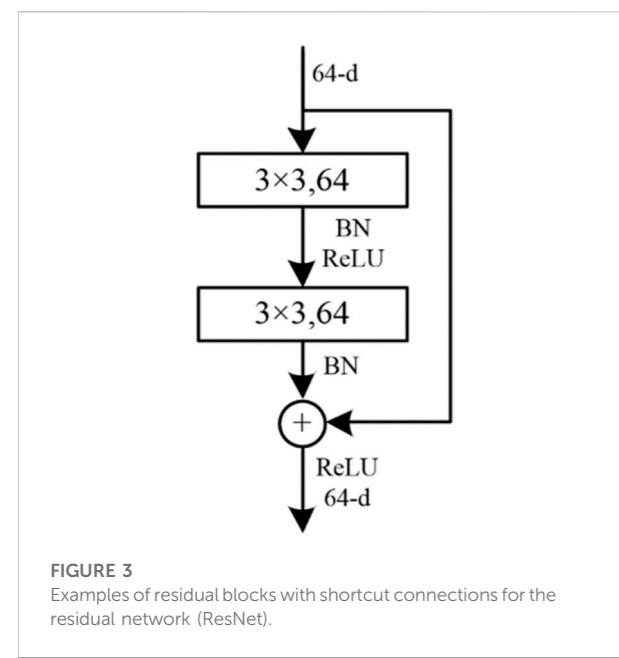


FIGURE 3
Examples of residual blocks with shortcut connections for the residual network (ResNet).

The residual block is shown in Figure 3. Shortcut connection skips two layers of 3×3 convolutional layer connected to the output. The output of the main line through the convolution operation is added to the input through the shortcut. Then the result is output through the ReLU activation function (Nair and Hinton, 2010).

Figure 4 shows the ResNet architecture with 34 layers (i.e. ResNet34). The rock image input to resnet34 is first passed through a 7×7 convolutional layer and a 3×3 max pooling layer (both with a stride of 2), and then fed into 16 residual blocks. All of these residual blocks have a total of 32 layers. Finally, the network ends with an average pooling layer, a fully connected layer, and a softmax layer.

3.2 Batch normalization

It is common for deep learning networks to consist of many layers. As the number of network layers increases, a significant deviation in data distribution across a layer will exacerbate, making it harder to optimize the model (Yan et al., 2020). Batch normalization (BN) can solve this problem well. Using batch normalization, data is divided into different groups and parameters are updated accordingly (Xiao et al., 2019). In the same group, the gradient direction is determined jointly, reducing randomness as the gradient declines. Furthermore, since the batch has fewer samples than the entire dataset, the amount of calculation has been significantly reduced. Batch normalization can avoid data offset because the batch normalization layer normalizes the input prior to the activation function.

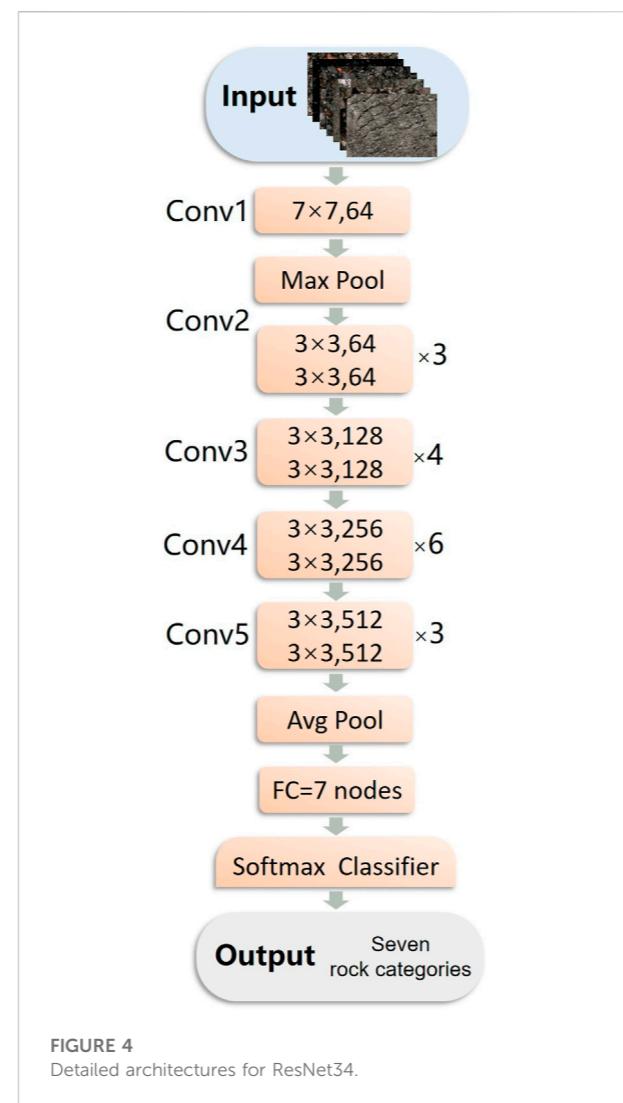


FIGURE 4
Detailed architectures for ResNet34.

In the ResNet34 rock image classification model we used, the BN is added before the ReLU activation function and after the convolutional layer. With the BN algorithm, parameter changes resulting from a different data distribution are minimized and the convergence speed during model training is accelerated. The formulas of batch normalization are as follows:

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{Var[x^{(k)}]}} \quad (2)$$

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)} \quad (3)$$

Where, $x^{(k)}$ is the characteristic distribution statistic of the current layer network. $E[x^{(k)}]$ is the mean value of data in current layer. $\sqrt{Var[x^{(k)}]}$ is the standard variance of data in current layer. $\gamma^{(k)}$ and $\beta^{(k)}$ are the learning parameters of the model.

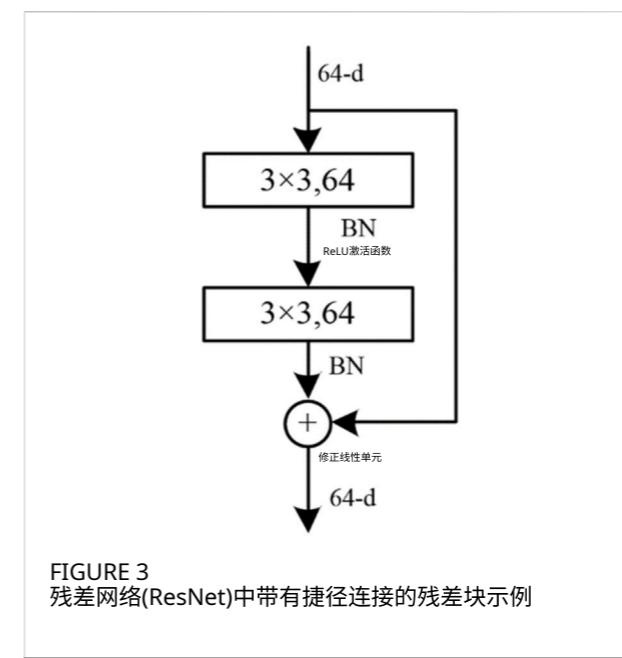


FIGURE 3
残差网络(ResNet)中带有捷径连接的残差块示例

图3展示了残差块结构。捷径连接跳过了与输出相连的两层 3×3 卷积层。通过卷积操作的主线输出与捷径输入的相加结果，最终经由ReLU激活函数输出(Nair和Hinton, 2010)。

图4展示了34层ResNet架构(即ResNet34)。输入resnet34的岩石图像首先经过 7×7 卷积层和 3×3 最大池化层(步长均为2)，随后输入16个残差块。这些残差块共计32层。网络末端由平均池化层、全连接层和softmax层构成。

3.2 批量归一化

深度学习网络通常由多层结构组成。随着网络层数增加，层间数据分布会出现显著偏差并加剧，导致模型优化难度增大(Yan等, 2020)。批量归一化(BN)能有效解决该问题。该方法将数据划分为不同组别并相应更新参数(Xiao等, 2019)。同组数据梯度方向由集体决定，随梯度下降可减少随机性。此外，由于批数据量远小于完整数据集，计算量得到显著降低。批量归一化层在激活函数前对输入进行标准化处理，可有效避免数据偏移。



FIGURE 4
ResNet34的详细架构。

在我们使用的ResNet34岩石图像分类模型中，BN层被添加在ReLU激活函数之前、卷积层之后。通过BN算法，不同数据分布导致的参数变化被最小化，同时加快了模型训练时的收敛速度。批量归一化的计算公式如下：

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{Var[x^{(k)}]}} \quad (2)$$

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)} \quad (3)$$

其中， $x^{(k)}$ 表示当前层网络的特征分布统计量， $E[x^{(k)}]$ 为当前层数据的均值， $\sqrt{Var[x^{(k)}]}$ 是当前层数据的标准差， $\gamma^{(k)}$ 和 $\beta^{(k)}$ 为模型的学习参数。

Through Eq. 3, the distribution of eigenvalues will be re-adjusted to a standard normal distribution and the eigenvalues are kept within the input-sensitive interval of the activation function, avoiding the disappearance of the gradient and speeding up the convergence.

3.3 ReLU activation function

The activation function is used to add nonlinear factors to the model because linear models are less expressive. In the absence of activation functions, the input of each layer node in the network is a linear function of the output of the upper layer, that is, inputs and outputs are linearly correlated (Liu et al., 2022). After adding the activation function, it is possible to apply neural networks to many nonlinear models arbitrarily because they can approach many nonlinear functions arbitrarily. As a result of the ReLU activation function, neurons are activated nonlinearly based on the feature map of the convolution layer output, enabling better learning by avoiding overfitting (Ran et al., 2019).

For each convolutional layer of ResNet34, the ReLU activation function is used:

$$f(x) = \max(0, x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

Where x is the input data. The ReLU activation function sets the output of some neurons in the network to zero, makes the network sparse and reduces the dependence between parameters, which solves the problem of overfitting. Another advantage of ReLU is that it is less computationally intensive and time consuming compared to other activation functions such as sigmoid, which involve exponential operations.

3.4 Softmax classifier

Softmax classifier is used in the establishment of the rock classification model. The input rock images can be converted into the corresponding category possibilities by the softmax classifier (Pham and Shin, 2020). At the end of ResNet34, the softmax classification function is added after the fully connected layer of the network, so that the output of the network is a one-dimensional vector of size 7, which represents the seven types of rocks to be classified in this study. The seven values in each one-dimensional vector reflect the rock class probability to which the input image belongs, so the sum of the seven values is 100%. The formula is as follows:

$$p(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}, \quad j = 1, 2, 3, \dots, n \quad (5)$$

Where, $p(z_i)$ is the probability of being identified as the i category, and n is the number of distinct categories. The

numerator maps input real value to zero to infinity and the denominator adds up all the results and normalizes them, as shown in Eq. 5.

3.5 Adaptive moment estimation

Adaptive moment estimation (Adam) is a stochastic optimization algorithm based on the adaptive estimation of low-order moments (Hang et al., 2019; Yang et al., 2019). The algorithm adaptively adjusted the learning rate update parameters through the first moment estimation and the second moment estimation of the gradient. In the past, many conventional deep neural networks use stochastic gradient descent algorithm (SGD), which iteratively updates the weights of the neural network until it reaches the global optimal solution. However, the model using SGD algorithm has a slow convergence speed in the early stage, and it is prone to decline in accuracy. The Adam algorithm is improved on the basis of SGD algorithm. The learning rate during network training is usually kept constant when using an optimization algorithm such as SGD, but Adam optimizes the network by iteratively updating the weights of the neural network and adaptively adjusting the learning rate as the network is trained, which makes the network converge faster and learn better.

In order to adjust the parameters of the rock classification model more efficiently and make it converge faster during training, Adam is chosen as the optimization algorithm. The updating formulas of Adam algorithm are as follows:

$$\theta_t = \theta_{t-1} - \alpha * \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (6)$$

$$g_t = \nabla \hat{L}(\theta_t) \quad (7)$$

$$m_t = \beta_1 * m_{t-1} + (1 - \beta_1) * g_t \quad (8)$$

$$v_t = \beta_2 * v_{t-1} + (1 - \beta_2) * g_t^2 \quad (9)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (10)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (11)$$

Where, t is the number of times, α is the learning rate, θ_t is the update parameter for the solution, ϵ is a very small constant which is set to prevent the denominator of Eq. 6 from being zero, $\hat{L}(\theta_t)$ is the loss function with parameter θ_t , g_t is the gradient of the partial derivative of the loss function $\hat{L}(\theta_t)$ with respect to θ_t . β_1 is the exponential decay rate of the first moment estimate. β_2 is the exponential decay rate of the second moment estimate. m_t is the estimate of the first moment of the gradient in momentum form. v_t is the second moment estimate of the gradient in momentum form. \hat{m}_t is the bias correction of m_t and \hat{v}_t is the bias correction of v_t . By default, $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$.

通过公式3，特征值分布将被重新调整为标准正态分布，并使其保持在激活函数的输入敏感区间内，从而避免梯度消失并加速收敛。

分子将输入实值映射到零至无穷大，分母汇总所有结果并进行归一化处理，如公式5所示。

3.5 自适应矩估计

自适应矩估计 (Adam) 是一种基于低阶矩自适应估计的随机优化算法 (Hang等, 2019; Yang等, 2019)。该算法通过梯度的一阶矩估计和二阶矩估计，自适应调整学习率更新参数。传统深度神经网络多采用随机梯度下降算法 (SGD)，通过迭代更新神经网络权重直至达到全局最优解。但SGD算法模型在初期收敛速度较慢，且易出现精度下降。Adam算法在SGD基础上改进，当使用SGD等优化算法时网络训练的学习率通常保持恒定，而Adam通过迭代更新神经网络权重并随训练过程自适应调整学习率，使网络收敛更快、学习效果更佳。

3.3 ReLU激活函数

激活函数用于为模型添加非线性因素，因为线性模型的表达能力较弱。若无激活函数，网络中每层节点的输入都是上层输出的线性函数，即输入与输出呈线性相关 (Liu等, 2022)。加入激活函数后，神经网络可任意应用于众多非线性模型，因其能任意逼近非线性函数。ReLU激活函数使神经元基于卷积层输出的特征图进行非线性激活，通过避免过拟合实现更优学习效果 (Ran等, 2019)。

ResNet34的每个卷积层均采用ReLU激活函数：

$$f(x) = \max(0, x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

其中 x 为输入数据。ReLU激活函数将网络中部分神经元的输出置零，使网络稀疏化并降低参数间依赖性，从而解决过拟合问题。相较于sigmoid等涉及指数运算的激活函数，ReLU的另一优势在于计算强度更低且耗时更少。

3.4 Softmax分类器

岩石分类模型构建采用Softmax分类器，该分类器可将输入的岩石图像转换为对应的类别概率(Pham and Shin, 2020)。在ResNet34末端网络的全连接层后添加Softmax分类函数，使网络输出为7维向量，对应本研究待分类的七类岩石。每个一维向量中的七个值反映输入图像所属岩类的概率，故七个数值之和为100%。计算公式如下：

$$p(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}, \quad j = 1, 2, 3, \dots, n \quad (5)$$

其中， $p(z_i)$ 表示被识别为 i 类别的概率， n 为不同类别的总数。

为更高效调整岩石分类模型参数并加速训练收敛，选用Adam作为优化算法。其参数更新公式如下：

$$\theta_t = \theta_{t-1} - \alpha * \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (6)$$

$$g_t = \nabla \hat{L}(\theta_t) \quad (7)$$

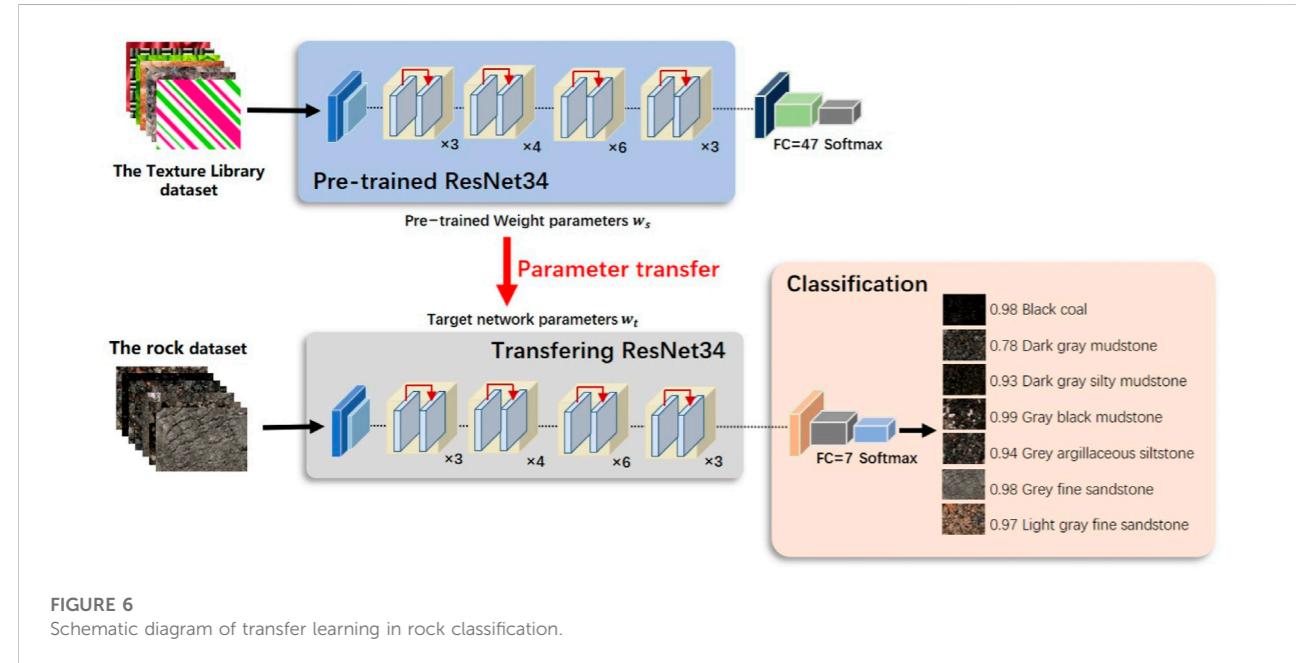
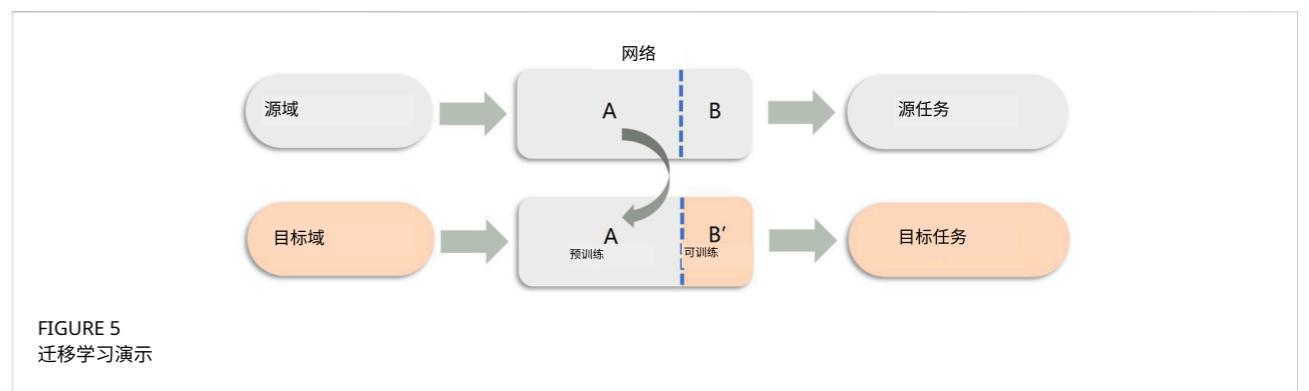
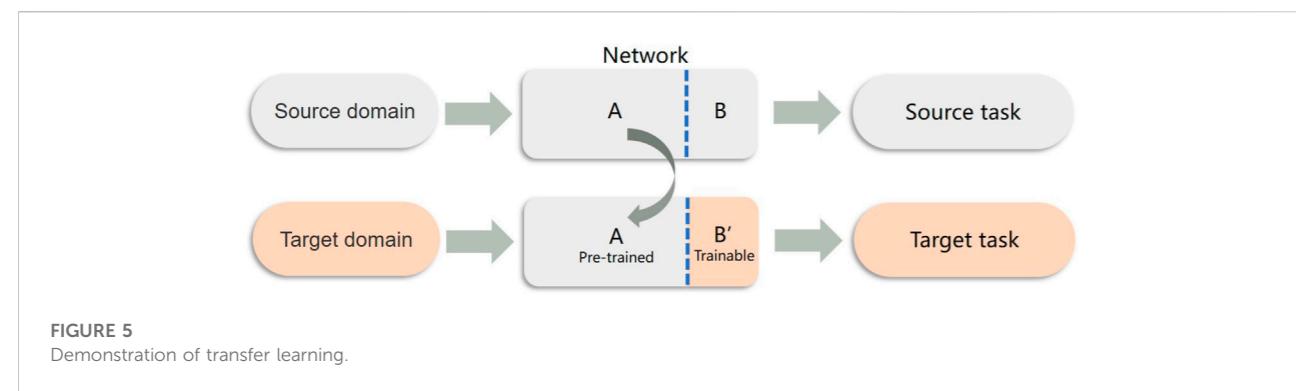
$$m_t = \beta_1 * m_{t-1} + (1 - \beta_1) * g_t \quad (8)$$

$$v_t = \beta_2 * v_{t-1} + (1 - \beta_2) * g_t^2 \quad (9)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (10)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (11)$$

其中， t 表示迭代次数， α 为学习率， θ_t 是解的更新参数， ϵ 为防止公式6分母为零而设置的极小常数， $\hat{L}(\theta_t)$ 是带参数 θ_t , g_t 的损失函数，表示损失函数 $\hat{L}(\theta_t)$ 对 θ_t 的偏导数梯度。 β_1 为一阶矩估计的指数衰减率， β_2 为二阶矩估计的指数衰减率。 m_t 是动量形式下梯度的一阶矩估计， v_t 为动量形式下梯度的二阶矩估计。 \hat{m}_t 和 \hat{v}_t 分别是对 m_t 和 v_t 的偏差校正。默认值设为 $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ 和 $\epsilon = 10^{-8}$ 。



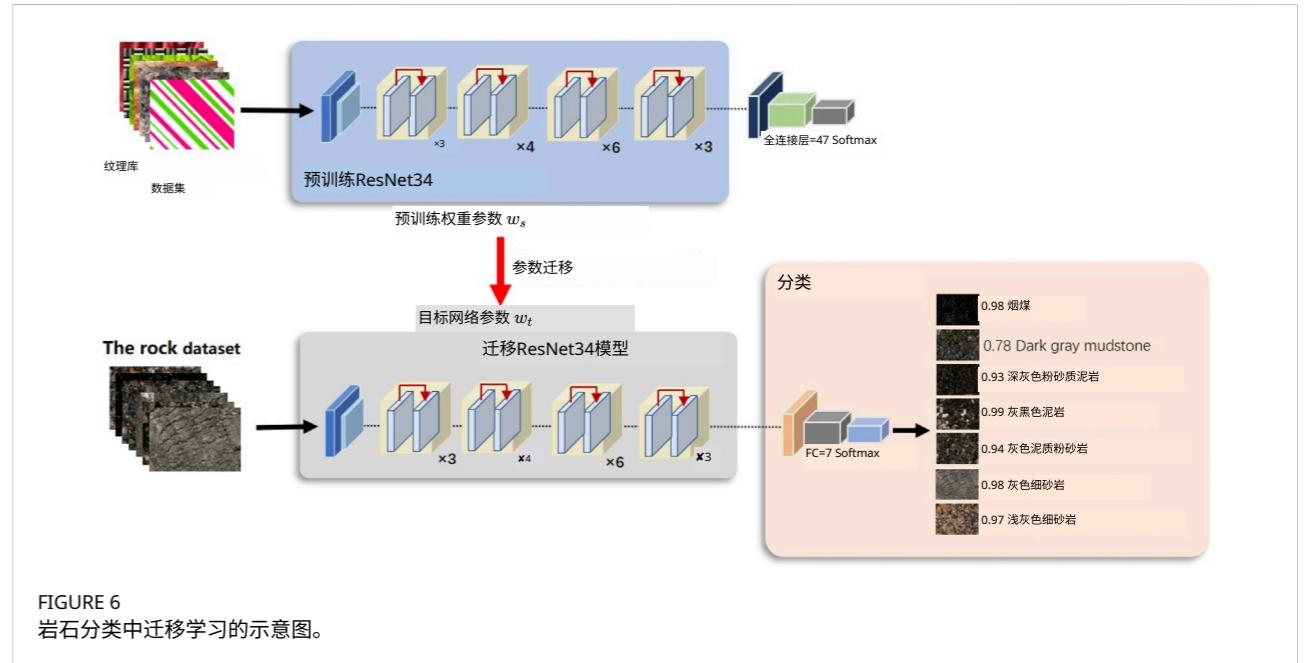
3.6 Transfer learning

Training convolutional neural networks usually require very large labeled datasets to achieve high accuracy. However, it is often difficult to obtain such data and it takes a lot of time to label the data. Due to the existence of these difficulties, the transfer learning method used in many studies to solve the cross-domain image classification problem has proven very effective. Transfer learning considers the correlation between different tasks, so that the knowledge obtained in the previous task can be directly applied to the new task through small transformation or even without any modification. Transfer learning is conducive to the construction of the mathematical model of the target task and reduces the dependence on the target task dataset (Gao et al., 2021). At present, the more complete convolutional neural

networks such as VGG, AlexNet, GoogLeNet and so on are pre-trained on the public image dataset of computer vision (Dabrowski and Michalik, 2017; Ali et al., 2020).

Since the model needs multiple rounds of iteration in the training process, and the number of rock pictures in this study is small, it will lead to the overfitting problem and low classification accuracy of the model. Consequently, transfer learning is a viable strategy (Figure 5). Given a labelled source domain D_S and learning task T_S , a target domain D_T and learning task T_T , transfer learning aims to help improve the learning of the target predictive function $f(\cdot)$ in D_T using the knowledge in D_S and T_S , where D_S is Texture Library dataset, D_T is rock dataset (Pan, 2017).

As shown in Figure 6, transfer learning is used to optimize the rock image classification. Transfer learning in the rock image



3.6 迁移学习

训练卷积神经网络通常需要大量标注数据集才能达到高精度。然而获取此类数据往往困难且耗时。针对这些难题，许多研究采用迁移学习方法解决跨领域图像分类问题，效果显著。迁移学习通过考量不同任务间的关联性，使得先前任务获得的知识只需微小调整甚至无需修改即可直接应用于新任务。该方法有助于构建目标任务的数学模型，并降低对目标数据集量的依赖 (Gao等, 2021)。目前较为完善的卷积神经

诸如VGG、AlexNet、GoogLeNet等网络已在计算机视觉领域的公共图像数据集上进行了预训练

(Dabrowski和Michalik, 2017; Ali等人, 2020)。由于模型训练过程需多轮迭代，而本研究中的岩石图片样本量较小，易导致模型过拟合及分类准确率低下。因此，迁移学习成为可行策略 (图5)。给定带标注的源域 D_S 和学习任务 T_S 、目标域 D_T 和学习任务 T_T ，迁移学习旨在利用 D_S 和 T_S 中的知识提升目标域 D_T 中预测函数 $f(\cdot)$ 的学习效果，其中 D_S 为纹理库数据集， D_T 为岩石数据集 (Pan, 2017)。

如图6所示，采用迁移学习优化岩石图像分类。岩石图像的迁移学习

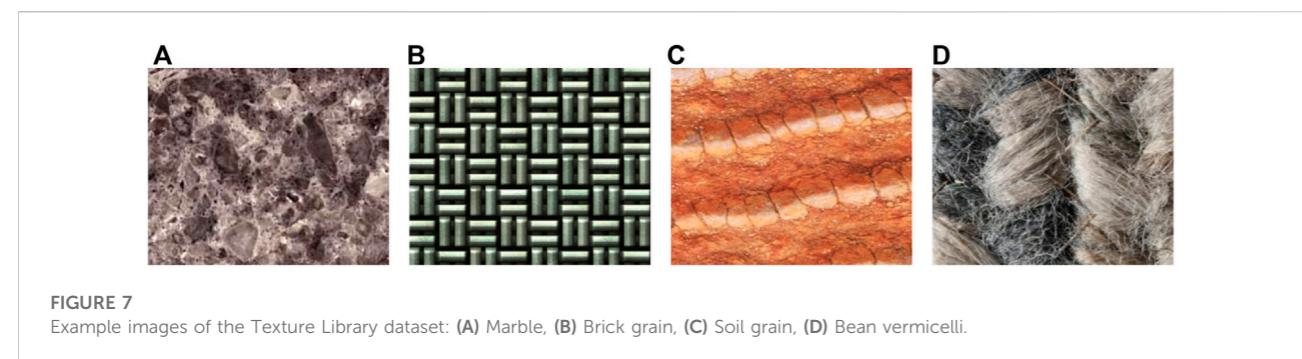


FIGURE 7
Example images of the Texture Library dataset: (A) Marble, (B) Brick grain, (C) Soil grain, (D) Bean vermicelli.

classification model includes pre-training and fine-tuning. Firstly, the ResNet34 model is pre-trained on the Texture Library dataset, the rock dataset is used to fine-tune the ResNet34 model afterwards.

The supervised learning architecture is used for pre-training. Pre-training usually requires a large enough dataset to help the model learn common features, and the learned features are parameterized to be ported to similar tasks for reuse (Zhu et al., 2021; Yi et al., 2022). The Texture Library dataset which contains 47 texture types total of 78,960 images is selected as the source domain for pre-training (Figure 7). The rock dataset and the Texture Library dataset are not identical, but the images of both have similar texture features, so the two dataset domains are related. Using the Texture Library dataset as the input of the pre-trained model for ResNet34, the characteristics of rocks can be well reflected. Therefore, it is reasonable to adopt the ResNet34 model pre-trained with the Texture Library dataset for rock image classification.

For fine-tuning, the parameters trained on the Texture Library dataset are used as initial values. The parameters of each layer of the network are frozen except for the last fully connected layer, and then input rock dataset and retrain the last fully connected layer to complete the fine-tuning.

In this study, the transfer learning method based on ResNet34 was applied to the rock image classification model. The ResNet34 pre-training weight parameters obtained by pre-training on the texture dataset are fine-tuned to speed up the convergence speed of the rock image classification network training, and spend less time training to obtain a model that can classify rock images. Transfer learning is used to simplify the original image training process, making the model learning more efficient and flexible.

4 Experiments and results

4.1 Data pre-processing

In the rock dataset used in the experiment, the number of rock images is too small and the pixel is too large. The number

of samples in each rock category is uneven, which will affect the recognition accuracy, so the rock training set is preprocessed.

4.1.1 Image slicing

Image information is composed of the spatial arrangement of pixels, so the features of an image are mainly represented by local adjacent pixels (Su et al., 2020). Large-scale images can represent more image detail information, so that the differences between images are more obvious. Image classification should make full use of image detail information. Therefore, we use the image slicing method to slice the 253 training sets at first. The original rock images acquired from the industrial camera contain $4,096 \times 3,000$ pixels and are sliced into 9 rows and 12 columns, meaning that each original image is divided into 108 sub-images. The size of each sub-image is 322×322 pixels. The original image and its cut part images are shown in Figure 8A.

4.1.2 Data augmentation

After image slicing, the training dataset is expanded to 27,324 images. The dataset used consisted of a relatively small number of images for training network. The data augmentation used in this study to expand the dataset were rotation, horizontal flip, vertical flip, blur, movement, brightness adjustment and Gaussian noise addition. The schematic of the data augmentation is shown in Figure 8B. The total number of training sets reached 382,536 by applying these transformations which fully expanded the original training set. The number of training set after pre-processing is also shown in Table 2.

4.1.2.1 Image resizing

Resizing changes the distance between different pixels in the image, typically along the x -axis and y -axis, and the matrix expression for image resizing is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (12)$$

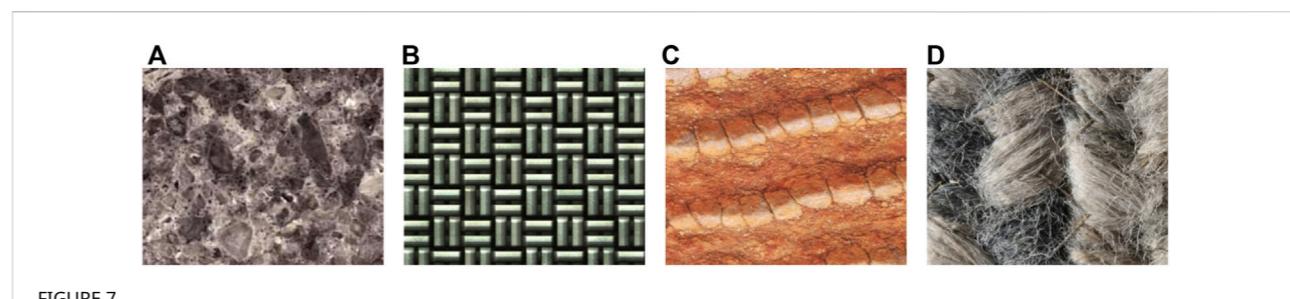


FIGURE 7
纹理库数据集示例图像：(A)大理石，(B)砖纹，(C)土壤颗粒，(D)粉丝纹理。

分类模型包含预训练与微调两阶段。首先在TextureLibrary数据集上预训练ResNet34模型，随后使用岩石数据集对该模型进行微调。

预训练采用监督学习架构，通常需要足够大的数据集帮助模型学习共性特征，并将学到的特征参数化迁移至相似任务复用（Zhu等，2021；Yi等，2022）。选择包含47种纹理类型、总计78,960张图像的纹理库数据集作为预训练源域（图7）。岩石数据集与纹理库数据集虽不完全相同，但两者图像具有相似纹理特征，属于关联域。以纹理库数据集作为ResNet34预训练模型输入，能充分反映岩石特征，因此采用该数据集预训练的ResNet34模型进行岩石图像分类具有合理性。

微调时，采用TextureLibrary数据集训练得到的参数作为初始值。冻结网络各层参数（除末层全连接层外），随后输入岩石数据集重新训练末层全连接层以完成微调。

本研究将基于ResNet34的迁移学习方法应用于岩石图像分类模型。通过在纹理数据集预训练获得的ResNet34权重参数进行微调，可加速岩石图像分类网络训练的收敛速度，用更少训练时间获得能实现岩石图像分类的模型。迁移学习简化了原始图像训练流程，使模型学习更高效灵活。

4 实验与结果

4.1 数据预处理

实验所用岩石数据集中，岩石图像数量过少且像素过大。

各类岩石样本数量不均衡会影响识别精度，因此对岩石训练集进行了预处理。

4.1.1 图像切片

图像信息由像素的空间排列构成，因此图像特征主要通过局部相邻像素来表征(Su等，2020)。大尺度图像能呈现更多细节信息，使图像间差异更显著。图像分类应充分利用细节信息，故我们首先对253个训练集采用切片处理。工业相机采集的原始岩石图像含 $4,096 \times 3,000$ 像素，按9行12列切分，每幅原始图像被划分为108个子图像，每个子图尺寸为 322×322 像素。原始图像及其切片示例如图8A所示。

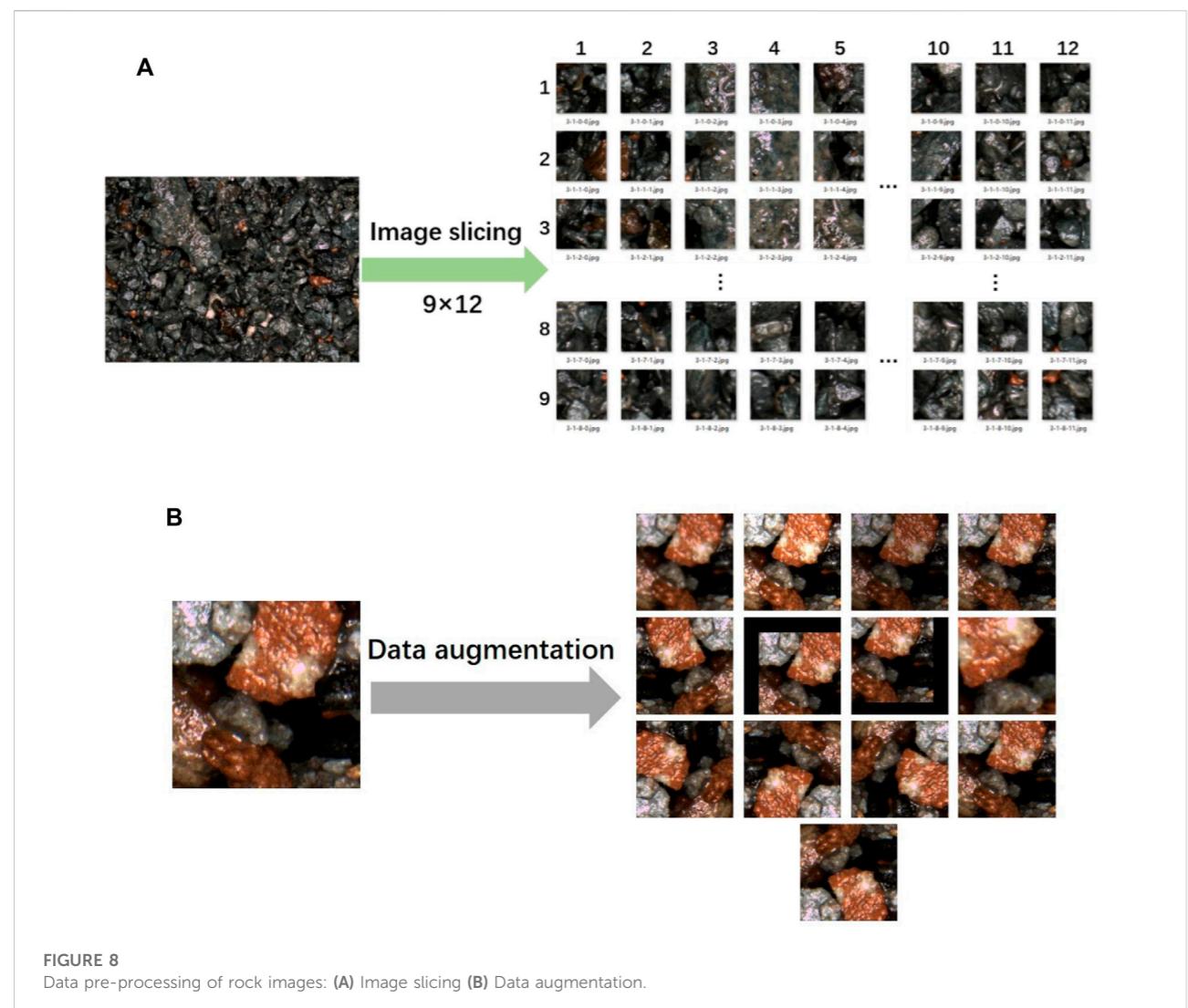
4.1.2 数据增强

图像切片后，训练数据集扩充至27,324张。本研究使用的原始训练集图像数量较少，通过旋转、水平翻转、垂直翻转、模糊处理、位移、亮度调节及高斯噪声添加等数据增强手段进行扩展。图8B展示了数据增强示意图。经上述变换后，训练集总量达到382,536张，实现了原始训练集的充分扩充。预处理后的训练集数量详见表2。

4.1.2.1 图像尺寸调整

尺寸调整会改变图像中像素点间的距离，通常沿 x 轴和 y 轴进行，其矩阵表达式如下：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (12)$$



Where, (x, y) is the original image and (x', y') is the resized image. S_x and S_y are the scaling factors along the x -axis and y -axis, respectively. The scaling factors are chosen randomly from 0.5 to 2.

4.1.2.2 Image rotation

Rotation is the process of rotating an image around a point to form a new image. The pixel values of the image before and after rotation remain unchanged. When the selected rotation point is the coordinate origin, the matrix expression for image rotation is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (13)$$

Where, θ is the rotation angle. The rotation angle is randomly selected from 0° to 360° .

4.1.2.3 Image movement

The matrix expression of image movement is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (14)$$

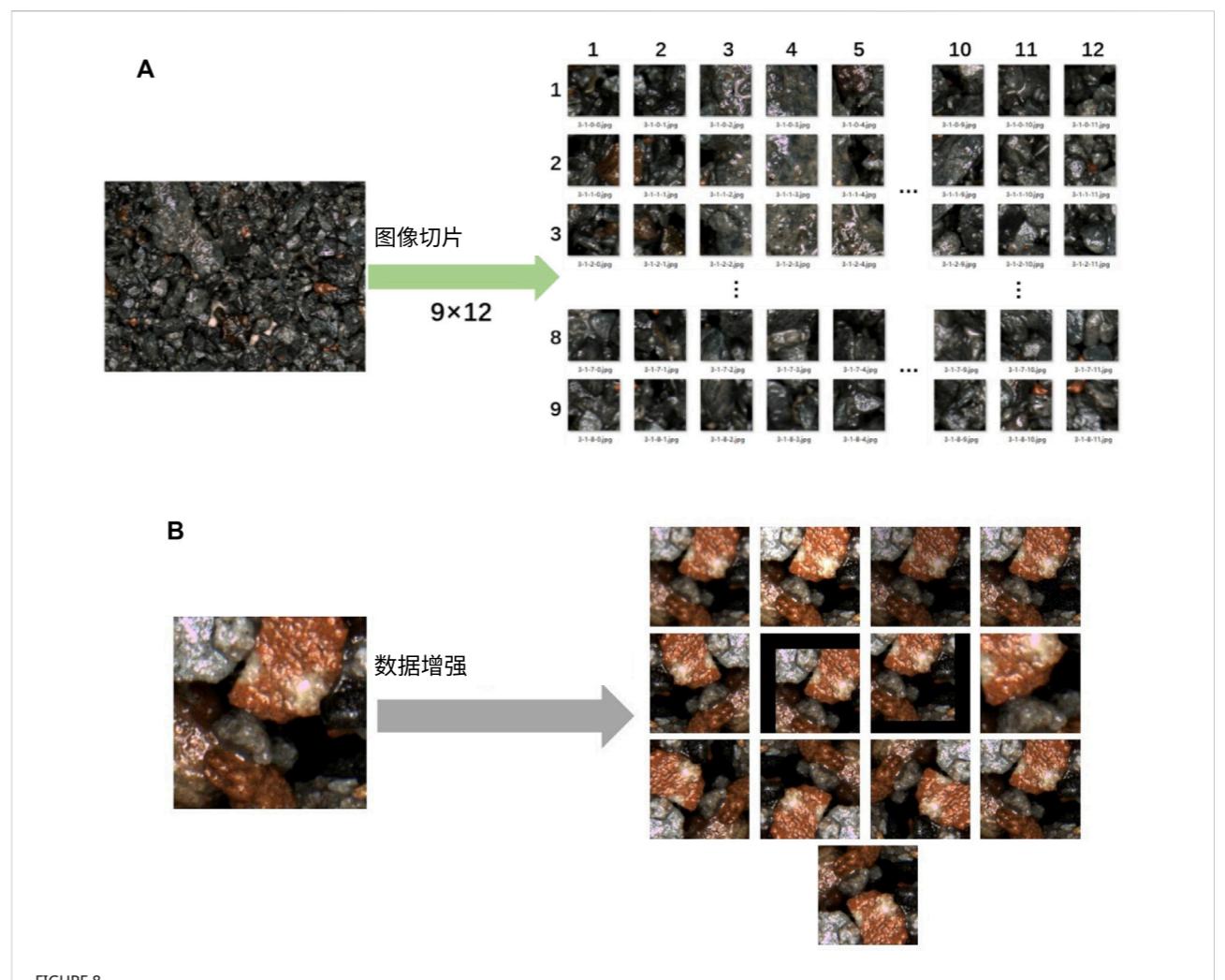
Where, t_x and t_y are the amount of translation to translate the image along the x -axis and y -axis, respectively. The amount of translation is chosen randomly from 30 to 80 pixels.

4.1.2.4 Image flip

The matrix expression for the horizontal flip is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & w \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (15)$$

The matrix expression for the vertical flip is as follows:



其中 (x, y) 为原始图像, (x', y') 为调整尺寸后的图像。 S_x 和 S_y 分别是沿 x 轴与 y 轴的缩放因子, 其取值范围为0.5至2的随机值。

4.1.2.2 图像旋转

旋转是指图像围绕某点转动形成新图像的过程, 旋转前后图像的像素值保持不变。当旋转点选为坐标原点时, 图像旋转的矩阵表达式如下:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (13)$$

其中, θ 为旋转角度, 该角度从 0° 到 360° 范围内随机选取。

4.1.2.3 图像平移

图像平移的矩阵表达式如下:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (14)$$

式中 t_x 和 t_y 分别表示图像沿 x 轴与 y 轴的平移量, 平移量在30至80像素间随机选取。

4.1.2.4 图像翻转

水平翻转的矩阵表达式如下:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & w \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (15)$$

垂直翻转的矩阵表达式如下:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (16)$$

Where w is the width of the image and h is the height of the image.

4.1.2.5 Brightness change

The change of image brightness belongs to the pixel transformation of the image, that is, the linear transformation is performed on each point of the two-dimensional matrix represented by the image. The transformation formula is as follows:

$$g(i, j) = \alpha \cdot f(i, j) + \beta \quad (17)$$

Where $f(i, j)$ is the pixel of the original image, $g(i, j)$ is the pixel of the output image, and i and j denote the pixel located in row i and column j . α is the gain parameter and β is the bias parameter. The brightness of the output image $g(i, j)$ is determined by β , and the contrast of the image is determined by α .

4.1.2.6 Noise addition

Due to the random interference of the external environment such as light and dust, the acquired rock image will contain noise. In order to simulate the real environment, Gaussian noise is added to the image. The probability density function of Gaussian noise is as follows:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right] \quad (18)$$

Where z is the gray value of the image pixel, μ is the mean value of the pixel value, and σ is the standard deviation of the pixel value. In the process of adding noise, a Gaussian noise with a mean value of 0 and a variance σ^2 of 0.01 is selected.

4.2 Evaluation metrics

The primary measures used to evaluate training effectiveness are classification accuracy and loss value. The classification accuracy is the percentage of the currently trained images that are accurately classified. It is formulated by Eq. 19:

$$Accuracy = \frac{t}{N} \quad (19)$$

Where, t is the number of samples whose predicted category is consistent with the actual category. N is the total number of samples. The effect of the model is measured by calculating the ratio of the number of samples correctly classified by the model to the total number of samples, and the goal is to measure the effect of the model.

Through the calculation of the loss function, the parameters of our model are updated. The goal is to reduce the optimization

error, that is, to reduce the empirical risk of the model under the joint effect of the loss function and the optimization algorithm (Chen et al., 2021). The cross-entropy is used as the loss function to evaluate the difference between the predicted value and the true value (Li et al., 2020). The loss value in this work is calculated by cross entropy, as follows:

$$Loss = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (20)$$

Where, M is the number of categories. y_{ic} is the indicator variable and takes the value of 0 or 1. If the category is the same as the category of sample i then it takes 1, and vice versa it takes 0. p_{ic} is the predicted probability that observation sample i belongs to category c . The function is convex and has good convergence properties when solved by gradient descent, and the global optimum can be obtained when solving the derivative. The loss value reflects the learning effect during the training of the model. The smaller the loss value is, the better the learning effect is.

4.3 Experiment details

The device information used in the experiment is as follows: the CPU model is Intel Xeon Silver 4,110 with 16 GB memory, and the GPU model is GeForce RTX 2080Ti with 11G memory. Windows10 was used as the operating system and Python 3.6 was used as the programming language. The deep learning framework is Pytorch, version 10.1 for CUDA, and version 7.6.5 for CuDNN.

The activation function selects the ReLU function. The optimizer selects is Adam. The learning rate is set to 0.001. The number of training epochs is 60 and the batch size is set to 16.

Different degrees of data preprocessing methods were used to conduct ablation experiments to explore the effectiveness of each preprocessing method. Resnet34 and three other different neural networks were trained to explore which worked best. The Texture Library dataset is selected as the source domain for transfer learning. The model parameter files are obtained after training. The other layers of the Resnet34 network are frozen except for the structural parameters of the fully connected layer. The pre-trained weights obtained from training on the texture dataset are loaded when the network is trained with the rock dataset. The prediction results are compared with the true label in each step so that the classification accuracy and loss value are both calculated to upload to the TensorBoard visual training tool.

4.4 Results analysis

4.4.1 The effectiveness of data pre-processing

The original data in this paper has been pre-processed by image slicing and data augmentation. In order to verify the

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (16)$$

其中 w 表示图像宽度, h 表示图像高度。

4.1.2.5 亮度变换

图像亮度变化属于像素级变换, 即对图像二维矩阵中的每个点进行线性变换。变换公式如下:

$$g(i, j) = \alpha \cdot f(i, j) + \beta \quad (17)$$

式中 $f(i, j)$ 为原始图像像素, $g(i, j)$ 为输出图像像素, i 和 j 表示位于第 i 行第 j 列的像素, α 为增益参数, β 为偏置参数。输出图像 $g(i, j)$ 的亮度由 β 决定, 图像对比度则由 α 决定。

4.1.2.6 噪声添加

由于光线、灰尘等外部环境的随机干扰, 获取的岩石图像会含有噪声。为模拟真实环境, 需在图像中添加高斯噪声, 其概率密度函数如下:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right] \quad (18)$$

其中 z 表示图像像素灰度值, μ 为像素值均值, σ 为像素值标准差。本实验添加噪声时选用均值为0、方差 σ^2 为0.01的高斯噪声。

4.2 评估指标

训练效果主要通过分类准确率和损失值评估。分类准确率指当前训练图像中被正确分类的百分比, 计算公式如式19所示:

$$Accuracy = \frac{t}{N} \quad (19)$$

式中 t 表示预测类别与实际类别一致的样本数, N 为样本总数。通过计算模型正确分类样本数与总样本数的比值来衡量模型效果, 旨在评估模型性能。

通过损失函数的计算, 模型参数得以更新, 其目的在于降低优化目标

误差, 即在损失函数与优化算法共同作用下减小模型的经验风险(Chen et al., 2021)。本研究采用交叉熵作为损失函数来评估预测值与真实值之间的差异(Li et al., 2020), 具体计算公式如下:

$$Loss = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (20)$$

式中 M 表示类别总数。 y_{ic} 为指示变量, 取值为0或1。当类别与样本 i 的类别相同时取1, 反之取0。 p_{ic} 。 c 表示观测样本 i 属于某类别的预测概率。该函数具有凸性, 采用梯度下降法求解时具有良好的收敛特性, 求导时可获得全局最优解。损失值反映了模型训练过程中的学习效果, 其数值越小表明学习效果越佳。

4.3 实验细节

实验所用设备信息如下: CPU型号为Intel至强银牌4110, 内存16GB; GPU型号为GeForce RTX 2080Ti, 显存11GB。操作系统采用Windows10, 编程语言使用Python 3.6。深度学习框架为Pytorch, CUDA版本10.1, CuDNN版本7.6.5。

激活函数选用ReLU函数, 优化器采用Adam算法。学习率设为0.001, 训练周期数为60, 批量大小设置为16。

采用不同程度的数据预处理方法进行消融实验, 以探究每种预处理方法的有效性。训练Resnet34及其他三种不同神经网络, 探究何种网络效果最佳。选取纹理库数据集作为迁移学习的源域, 训练后获得模型参数文件。除全连接层结构参数外, 冻结Resnet34网络其余层。在岩石数据集训练时加载纹理数据集预训练权重。每一步预测结果均与真实标签比对, 计算分类准确率与损失值并上传至TensorBoard可视化训练工具。

4.4 结果分析

4.4.1 数据预处理有效性

本文原始数据已通过图像切片与数据增强进行预处理。为验证

TABLE 3 Comparison of training results for different data preprocessing methods.

Method	Number of images in the training set	The highest accuracy achieved in 60 epochs (%)
ResNet34 + Original training set	253	73.8
ResNet34 + IS	27,324	76.2
ResNet34 + DA	3,542	84.4
ResNet34 + IS + DA	3,82,536	88.1

表3 不同数据预处理方法的训练结果对比

方法	训练集图像数量	60个训练周期内达到的最高准确率(%)
ResNet34+原始训练集	253	73.8
ResNet34+图像分割	27,324	76.2
ResNet34网络+数据增强	3,542	84.4
ResNet34网络+图像分割+数据增强	3,82,536	88.1

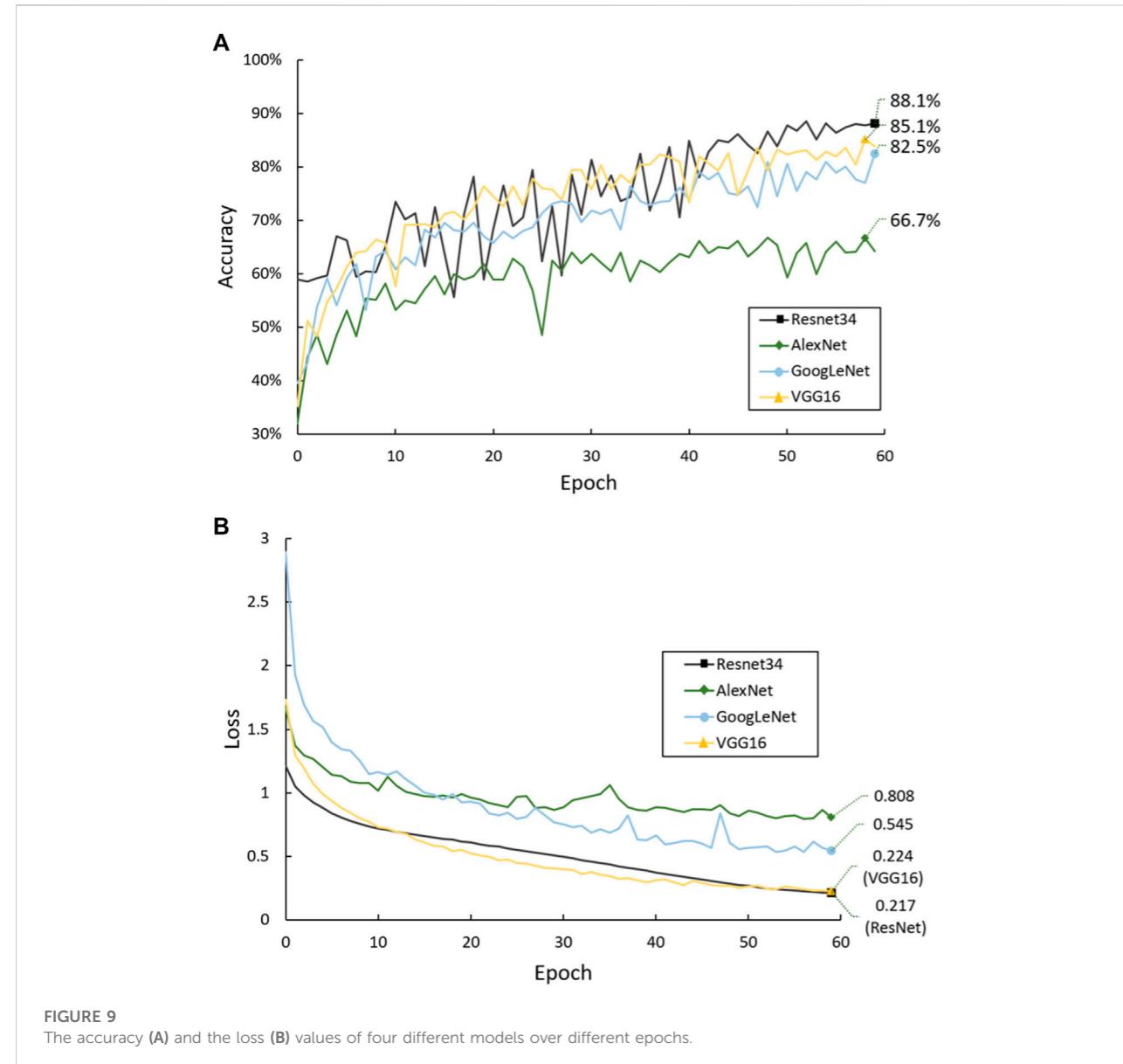


FIGURE 9
The accuracy (A) and the loss (B) values of four different models over different epochs.

effectiveness of data pre-processing, we conduct ablation experiments. The ResNet34 network was used to conduct four groups of experiments on different training sets: 1) no

data pre-processing is used; 2) using image slicing; 3) using data augmentation; 4) using image slicing and data augmentation both on the training set. The highest

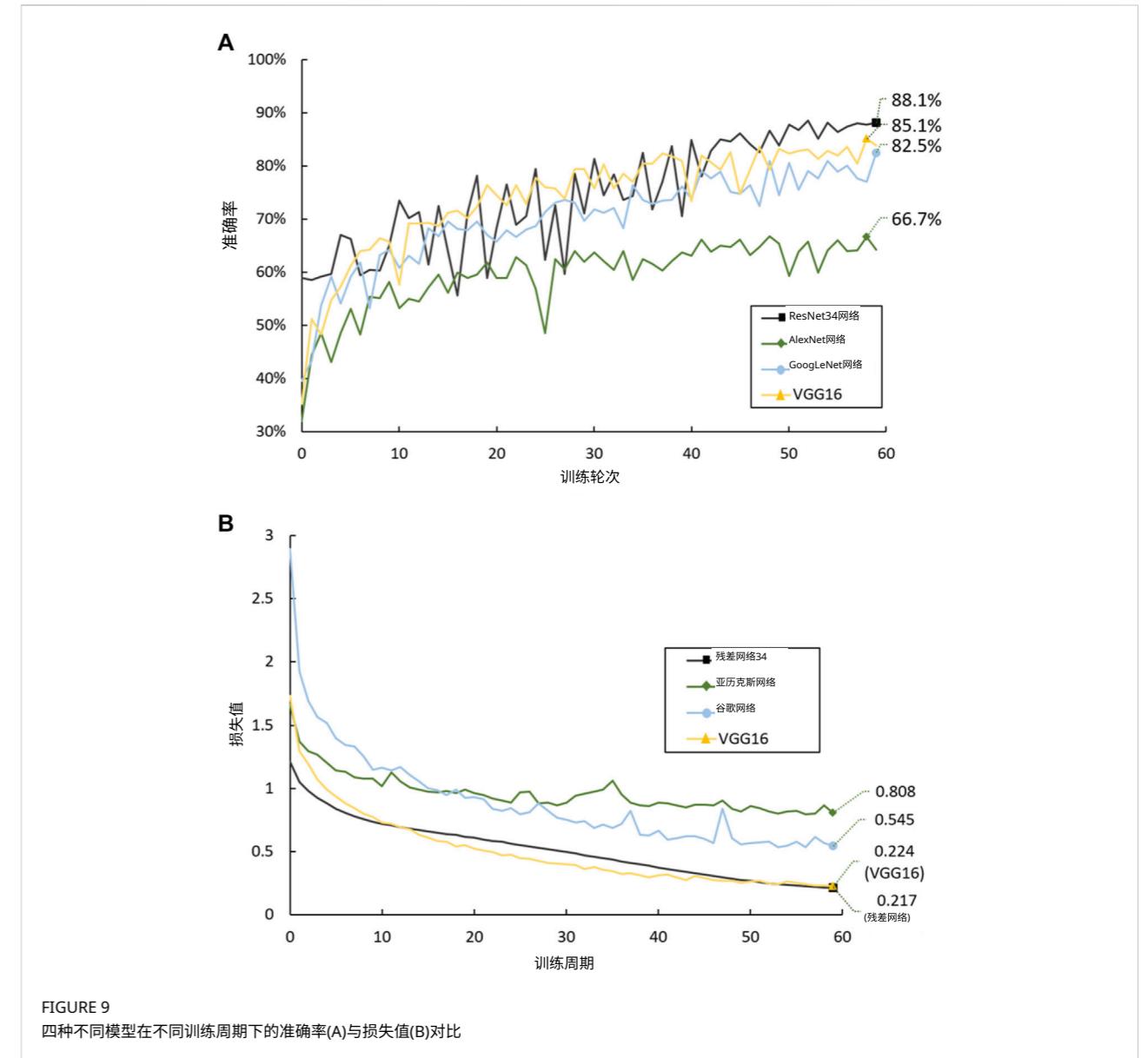
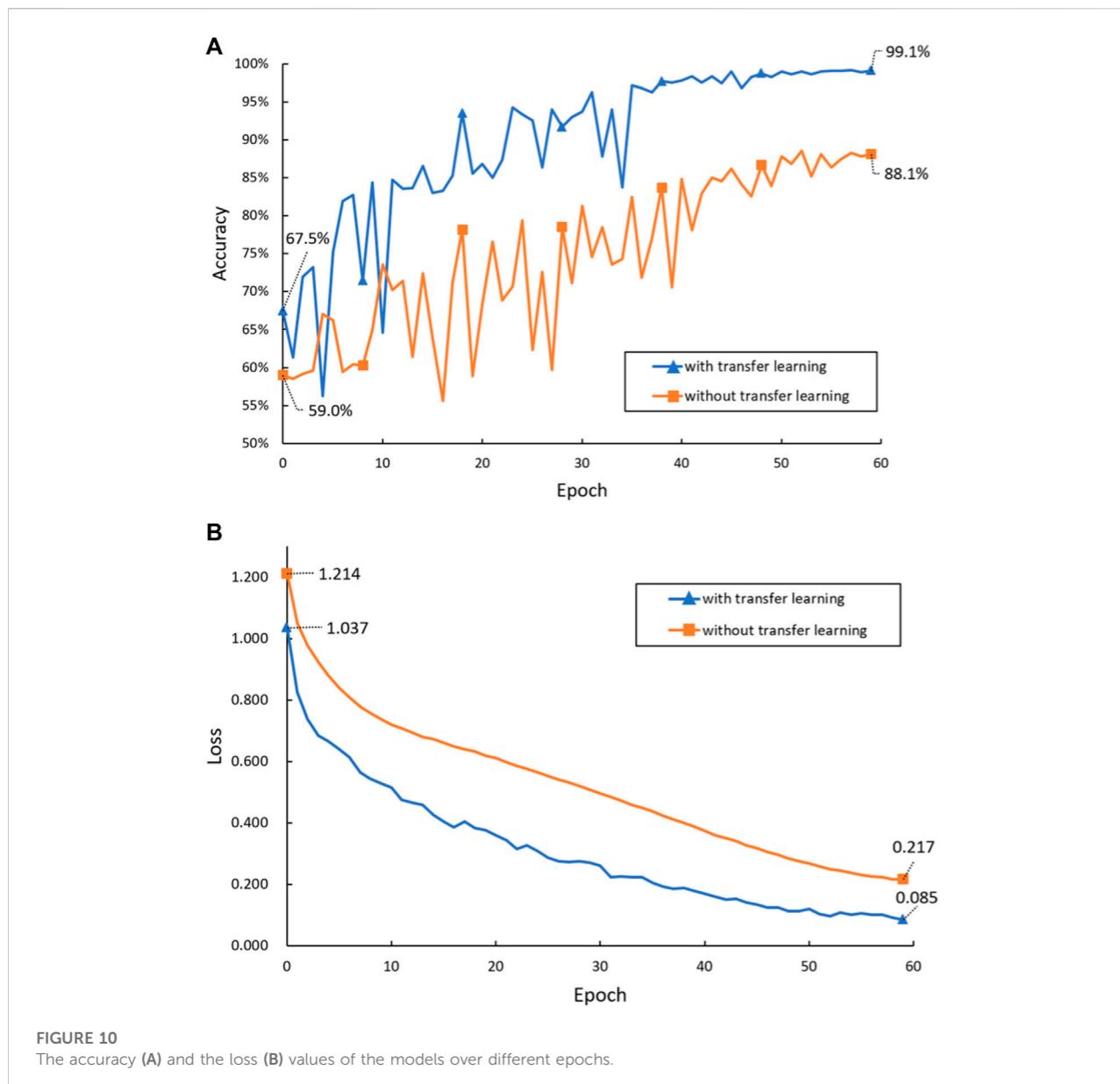


FIGURE 9
四种不同模型在不同训练周期下的准确率(A)与损失值(B)对比

为验证数据预处理的有效性，我们进行消融实验。采用ResNet34网络在四种训练集配置下测试：1)未使用

数据预处理；2)采用图像切片；3)采用数据增强；4)训练集同时应用图像切片与数据增强。最高



accuracy achieved by each method in 60 epochs is shown in Table 3.

After different degrees of image pre-processing, the classification accuracy of the network is improved in different degrees. Compared with the original training set, the accuracy of the training set after image slicing and data augmentation is improved by 14.3%. The result indicates that pre-processing of small sample data sets can make the network extract more comprehensive rock features and improve the generalization ability of the model. And it proves that the data pre-processing method in this paper can improve the overall accuracy of the classification network.

4.4.2 The effectiveness of residual networks

Four different network models to apply to rock classification in order to compare which network has the best effect are trained respectively. The training is visualized in the Pytorch framework using the TensorBoard tool.

Figure 9 illustrates the loss and accuracy changes for four deep learning methods (AlexNet, VGG16, GoogleLeNet, and ResNet34) as experiment steps increase. It shows that each of the four convolutional neural networks converges as the training process of rock image classification proceeds. In addition, it can be reflected from Figure 9A that the rock accuracy of the four networks from high to low is ResNet34, VGG16, GoogLeNet, and

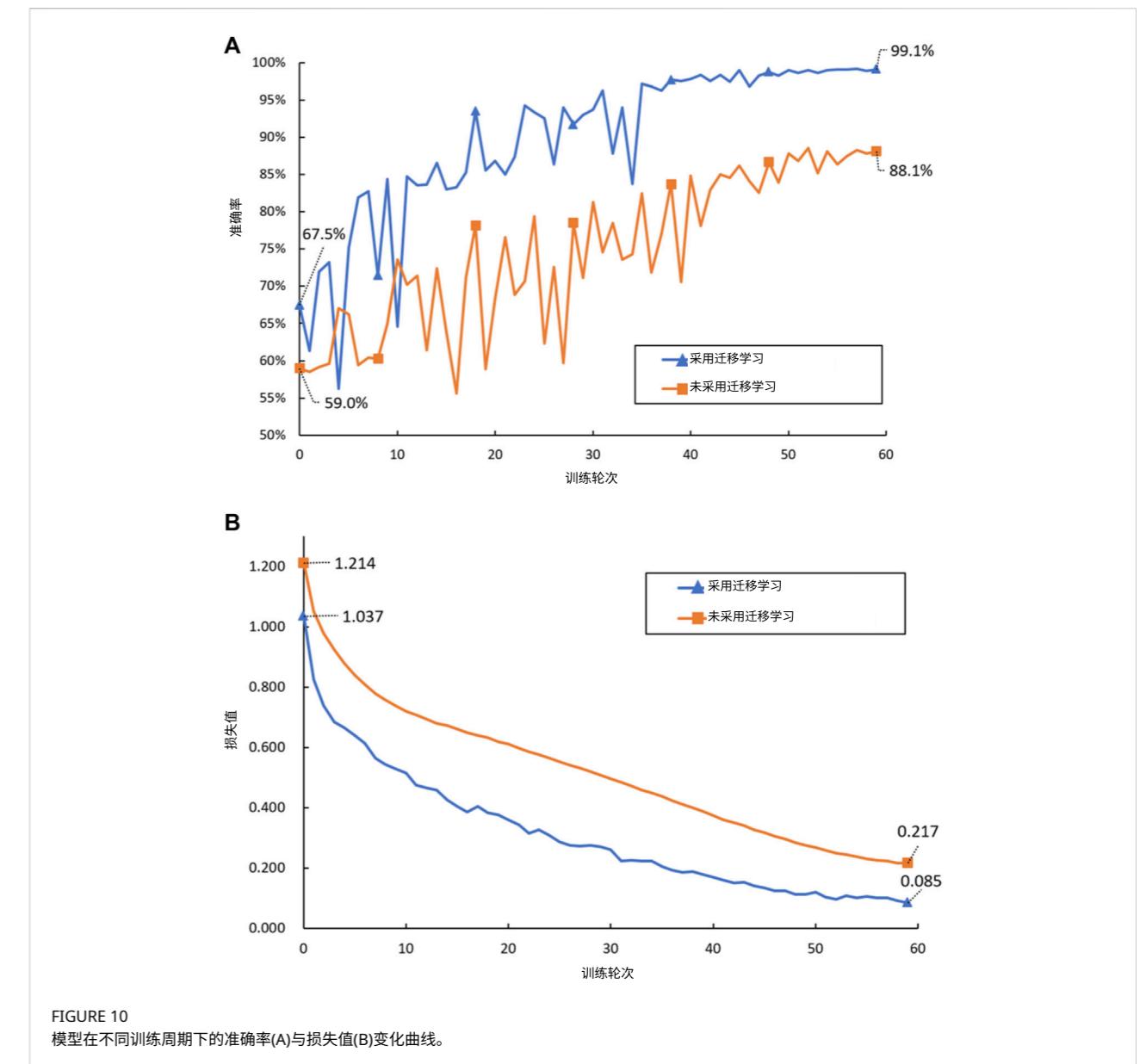


表3展示了各方法在60个训练周期内达到的准确率。

经过不同程度的图像预处理后，网络分类准确率均有提升。与原始训练集相比，图像切片和数据增强后的训练集准确率提高了14.3%。结果表明，小样本数据集的预处理能使网络提取更全面的岩石特征，提升模型泛化能力。这证实了本文提出的数据预处理方法可有效提高分类网络的整体准确率。

4.4.2 残差网络的有效性

为比较不同网络的岩石分类效果，分别训练了四种网络模型。训练过程通过TensorBoard工具在Pytorch框架中实现可视化。

图9展示了四种深度学习方法（AlexNet、VGG16、GoogleLeNet和ResNet34）随实验步骤增加的损失值与准确率变化。可见随着岩石图像分类训练的进行，四种卷积神经网络均呈现收敛趋势。另从图9A可知，四种网络的岩石分类准确率从高到低依次为ResNet34、VGG16、GoogLeNet和

TABLE 4 Comparison of training results with or without transfer learning.

Training results	Without transfer learning	With transfer learning
The accuracy in the first epoch	59.0%	67.5%
The loss value in the first epoch	1.214	1.037
The accuracy after 60 epochs	88.1%	99.1%
The loss value after 60 epochs	0.217	0.085

AlexNet. While the loss values (Figure 9B) are the opposite, from large to small are AlexNet, GoogLeNet, VGG16, and ResNet34.

The residual neural network has the highest accuracy and the lowest loss value, which is because the residual network uses residual structure to solve the model degradation problem of deep neural network. In conclusion, the ResNet34 network performs better than other networks in rock image classification.

4.4.3 The effectiveness of transfer learning

In order to explore whether the model using transfer learning performs better, the rock classification model using transfer learning method and the model without transfer learning method are trained respectively. The change of accuracy and loss value during the training process is shown in Figure 10. Compared with the training results with and without transfer learning, the first accuracy and the highest accuracy of the training epochs and the corresponding loss value are shown in Table 4. Combined with the graph, it can be observed that the model training without transfer learning has a low accuracy of 59.0% in the early stage. The highest accuracy reaches 88.1% after 60 epochs and the corresponding loss value is 0.217. In the model using transfer learning, the accuracy of the first epoch reaches 67.5% and the accuracy fluctuates slightly during the training process. The highest accuracy reached 99.1% which achieves an 11% improvement compared with the model without transfer learning, and the corresponding loss value is 0.085.

The accuracy and loss values reflect that the effect of the network model trained by transfer learning is obviously better than the original model. The model using transfer learning has high initial accuracy and high final accuracy. This is because the pre-training network based on the Texture Library dataset has learned rich texture spatial structure features and morphological correlation. The parameters of the pre-training model can be directly used in the model training, which can save training time and improve the precision of rock classification.

4.4.4 Reality testing

A total of 15 images covering 7 types of rocks from the testing dataset were classified by the model which had the weights with the highest accuracy using transfer learning method. The rock images to be recognized were fed into the rock prediction program, and the classification result was given in the form of names and probabilities. The test result of rock classification is

shown in Figure 11. All 15 images were correctly predicted with probabilities above 82%, and most of them were even above 95%.

Since the shooting angle and distance of rocks in the survey site are not fixed, the classification effect of rock images with different views is tested in this paper. Considering that the imaging resolution of each camera is not the same in practical applications, it is also necessary to test the effect of different resolutions of images on the rock classification results.

To simulate the camera changes at different resolutions, the image resolution was changed while keeping the view of the image unchanged. The original images of the test set are all $4,096 \times 3,000$ pixels. The original 15 test set images were downsampled multiple times to reduce the image resolution. The average accuracy of classifying 15 rock images is used as the evaluation criterion, and the experimental results are shown in Figure 12A.

It can be seen that the classification accuracy starts to suffer when the image is below 512×375 pixels. This is because the images of the original training set are processed to 322×322 pixels through image slicing in the previous data preprocessing, so the network can accurately identify the input images with a resolution higher than 322×322 . When the resolution of an image is lower than 322×322 pixels, the reduction of image features affects the rock classification results.

In order to simulate the change in the distance between the camera and the rock sample, the view of the image is changed while keeping the image resolution unchanged. The above experimental results show that the rock classification model can accurately classify rocks when the input image is in the pixels range of $4,096 \times 3,000$ to 512×375 . However, the accuracy starts to decrease after pixels are below 512×375 . Therefore 512×375 pixels are used as the minimum image resolution limit. After arbitrarily cropping an image with the same proportion as the original rock image and greater than 512×375 pixels, a random brightness change is added to simulate the field light change. And it is downsampled to 512×375 pixels to control the image resolution consistency, then input into the classification network for classification test. A total of 10 tests were performed, and the test set for each classification was 15 images. The experimental results are shown in Figure 12B.

The accuracy of rock classification does not change significantly due to the data pre-processing we have used. The pre-processing can improve the robustness and

表4 迁移学习与否的训练结果对比

训练结果	无迁移学习	采用迁移学习
首轮训练准确率	59.0%	67.5%
首轮训练损失值	1.214	1.037
60个训练周期后的准确率	88.1%	99.1%
60个训练周期后的损失值	0.217	0.085

AlexNet。而损失值（图9B）呈现相反趋势，从高到低依次为AlexNet、GoogLeNet、VGG16和ResNet34。

残差神经网络具有最高准确率和最低损失值，这是因为残差网络通过残差结构解决了深度神经网络模型退化问题。综上，ResNet34网络在岩石图像分类任务中表现优于其他网络。

4.4.3 迁移学习的有效性

为探究采用迁移学习的模型是否表现更优，本研究分别训练了使用迁移学习方法的岩石分类模型与未采用迁移学习的模型。图10展示了训练过程中准确率与损失值的变化曲线。对比两种训练结果，表4列出了训练周期中的初始准确率、最高准确率及对应损失值。结合图表可观察到：未采用迁移学习的模型初期准确率仅为59.0%，经过60个训练周期后最高准确率达到88.1%，对应损失值为0.217；而采用迁移学习的模型首轮训练准确率即达67.5%，训练过程中准确率波动较小，最终最高准确率达99.1%，较未采用迁移学习的模型提升11个百分点，对应损失值为0.085。

准确率和损失值表明，通过迁移学习训练的网络模型效果明显优于原始模型。采用迁移学习的模型具有较高的初始准确率和最终准确率，这是因为基于纹理库数据集的预训练网络已学习到丰富的纹理空间结构特征和形态关联性。预训练模型的参数可直接用于模型训练，既能节省训练时间，又能提高岩石分类精度。

4.4.4 实际测试

使用迁移学习方法中准确率最高的权重模型，对测试数据集中涵盖7类岩石的15张图像进行分类。将待识别岩石图像输入预测程序后，分类结果以名称和概率形式呈现。岩石分类测试结果

如图11所示。15张图像的预测正确率均超过82%，其中大部分甚至达到95%以上。

由于调查区域岩石的拍摄角度和距离不固定，本文测试了不同视角下岩石图像的分类效果。考虑到实际应用中各相机的成像分辨率存在差异，还需测试不同分辨率图像对岩石分类结果的影响。

为模拟不同分辨率下的相机变化，在保持图像视角不变的情况下调整分辨率。测试集原始图像均为 $4,096 \times 3,000$ 像素，通过对15张原始测试集图像进行多次降采样来降低分辨率。以15张岩石图像分类的平均准确率作为评价标准，实验结果如图12A所示。

可见当图像低于 512×375 像素时，分类准确率开始下降。这是因为原始训练集中的图像在前期数据预处理中通过切片处理为 322×322 像素，因此网络能准确识别分辨率高于 322×322 的输入图像。当图像分辨率低于 322×322 像素时，特征信息减少会影响岩石分类结果。

为模拟相机与岩石样本间距离变化，在保持图像分辨率不变的情况下改变成像视角。上述实验结果表明：当输入图像像素处于 $4,096 \times 3,000$ 至 512×375 范围内时，岩石分类模型能准确识别；但当像素低于 512×375 时准确率开始下降，故将 512×375 像素设为最小图像分辨率阈值。对原始岩石图像按相同比例任意裁剪出大于 512×375 像素的图片后，通过添加随机亮度变化模拟野外光照条件，并下采样至 512×375 像素以控制分辨率一致性，最后输入分类网络进行测试。共进行10次测试，每类测试集为15张图像，实验结果如图12B所示。

采用我们所使用的数据预处理方法后，岩石分类的准确性并未发生显著变化。该预处理过程能够提升模型的鲁棒性及

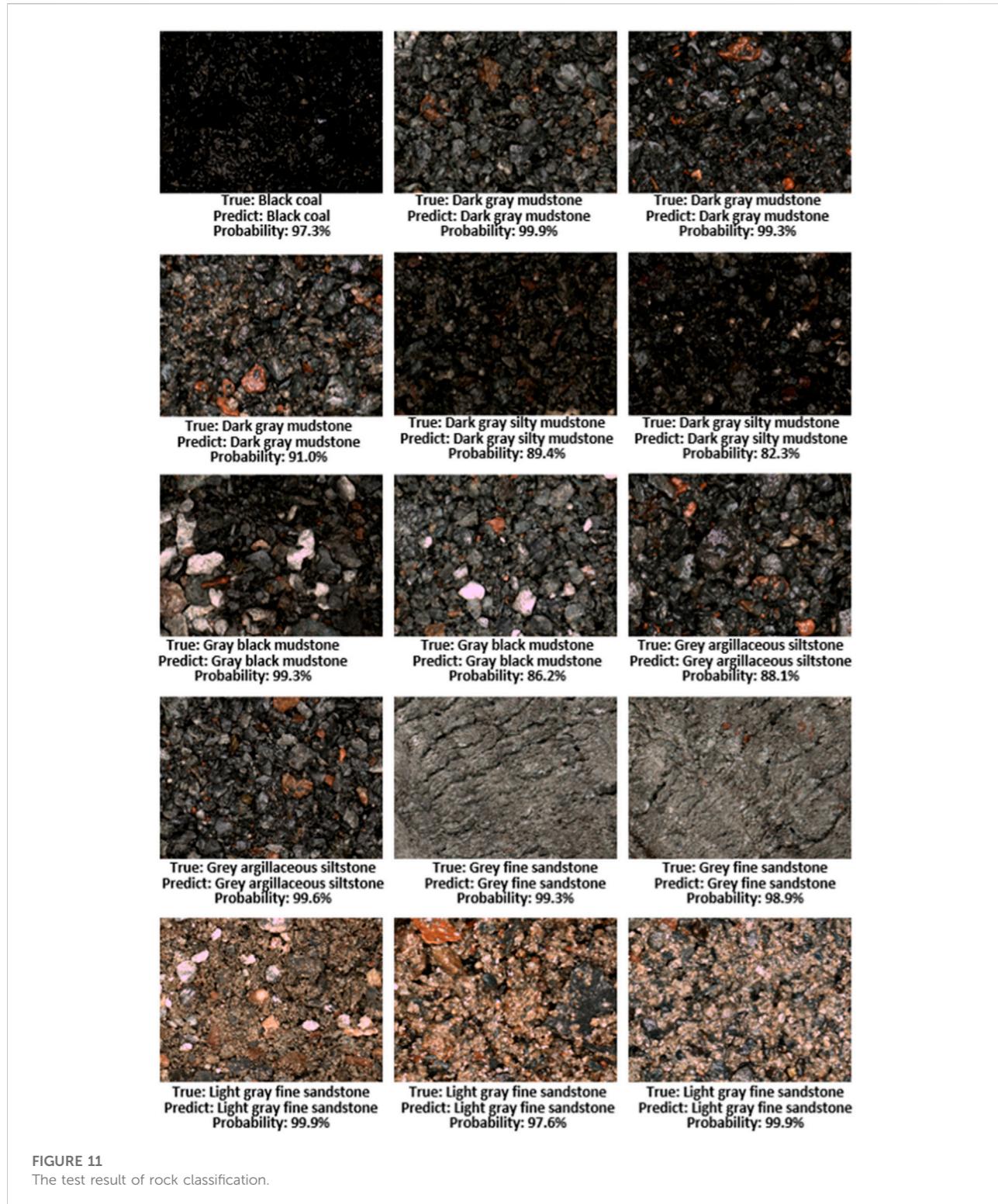


FIGURE 11
The test result of rock classification.

generalization of the model. Therefore, the model can adapt to the changes of different resolutions, shooting angles and shooting scenes. It indicates that the model learns more

about rock lithological features with the increase in data volume. This result also shows that the model has good robustness and generalization ability.

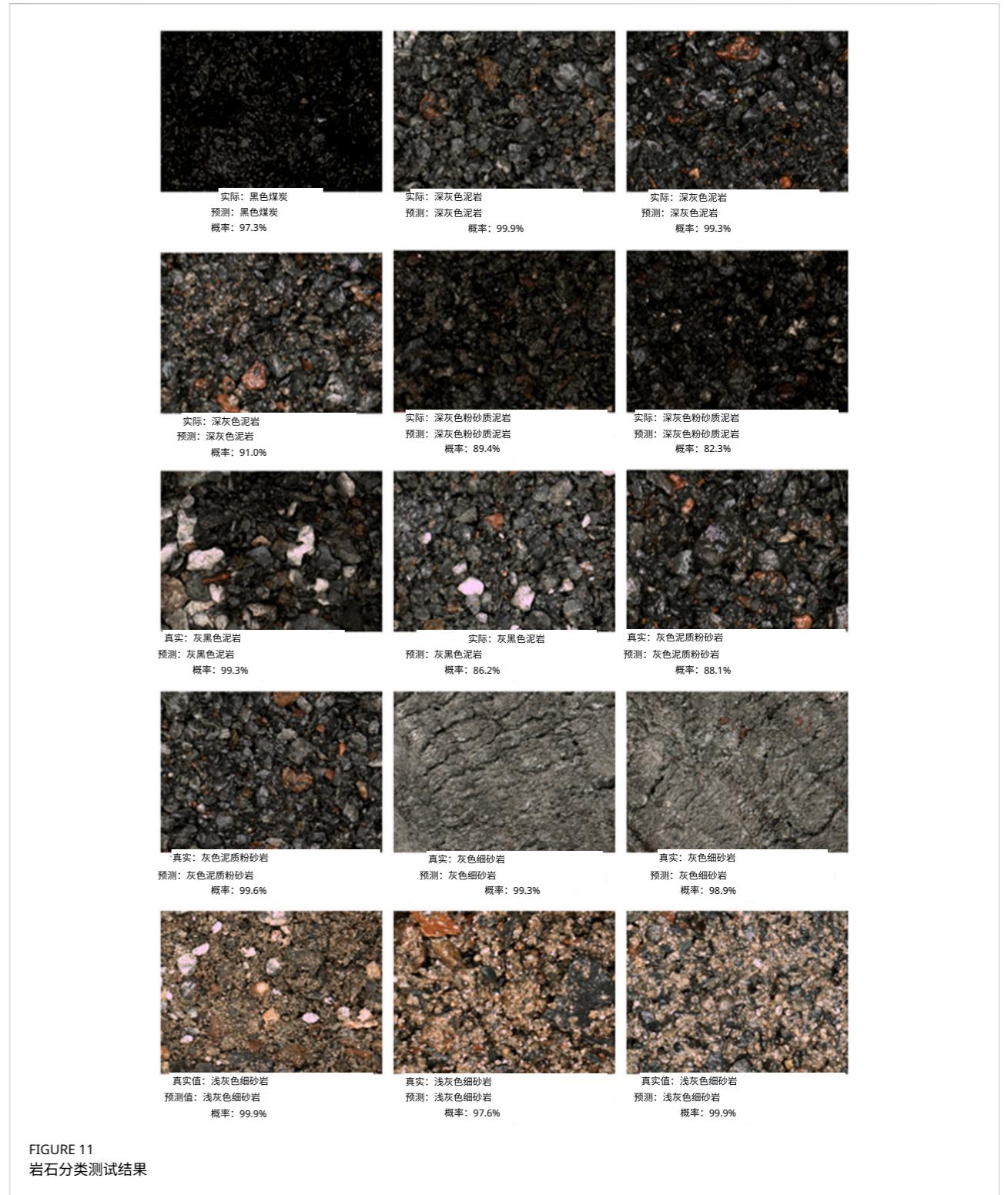


FIGURE 11
岩石分类测试结果

模型的泛化能力。因此该模型能适应不同分辨率、拍摄角度和拍摄场景的变化，表明随着数据量增加，模型能学习更多

岩石岩性特征。该结果也证明模型具有良好的鲁棒性和泛化能力。

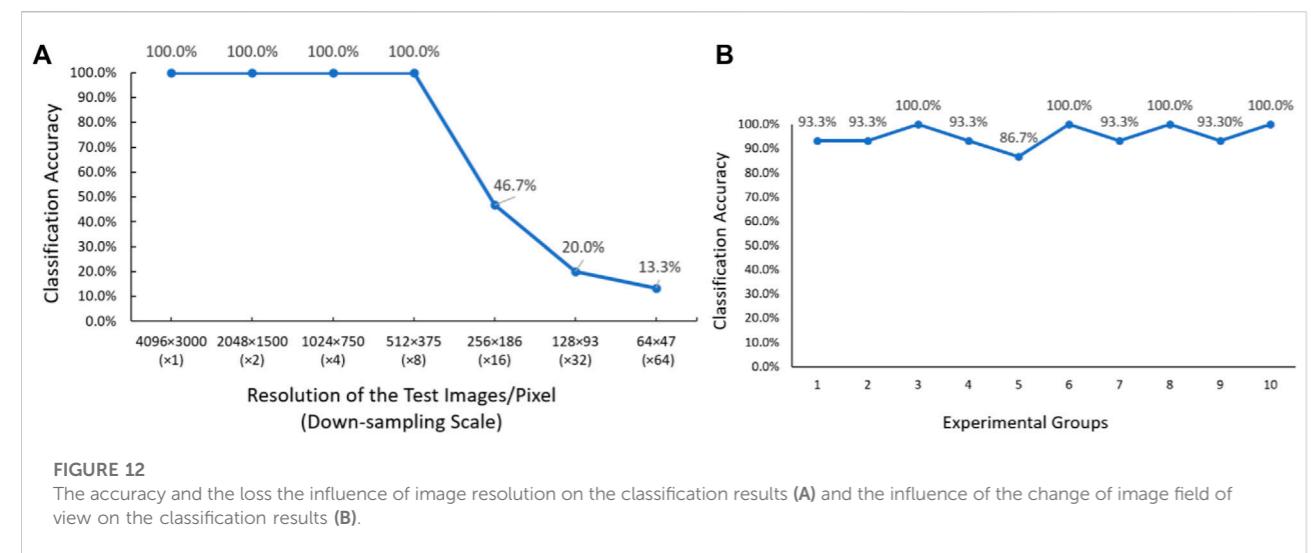


FIGURE 12

The accuracy and the loss the influence of image resolution on the classification results (A) and the influence of the change of image field of view on the classification results (B).

4.4.5 Comprehensive analysis

Ablation experiments were conducted to verify the effectiveness of the data preprocessing done in this paper. Since the original rock data set is too small, image segmentation and data augmentation can significantly improve the accuracy of rock classification. The effectiveness of residual neural networks is verified by the comparative experiment of AlexNet, VGG16, GoogLeNet, and ResNet34. The effectiveness of transfer learning is verified by the comparative experiment between transfer learning and non-transfer learning. The practical usability of the rock classification model was verified by testing 15 images containing all seven types of rocks. All 15 images were correctly predicted with probabilities above 82%, and most of them were even above 95%. By simulating and testing the actual situation of camera view changes and resolution changes, it is verified that the model has good robustness, slight scene changes will not affect the accuracy of rock classification, and the effectiveness of data preprocessing is also shown. These experimental results indicate that the model using transfer learning with the pre-trained residual neural network has higher classification accuracy and good generalization ability.

5 Deployment and application of rock classification network

Geological survey work often needs to be carried out on the construction site or in off-line conditions. Geological investigators need to carry all kinds of geological exploration equipment, such as GPS measurement, positioning instruments, measuring instruments and so on. It is inconvenient to take equipment with a certain weight and volume such as workstations, and it is impossible to obtain timely feedback

on rock types through the network to guide the following investigation. Deploying the rock image classification model proposed in this paper to the embedded end device can effectively solve this problem.

In this paper, rock image classification is shifted from theoretical research to practical applications. The trained rock classification network model is transplanted to Nvidia Jetson TX2 embedded platform, the TensorRT inference optimizer is used to accelerate the model, and the front-end interface that integrates all aspects of the system is developed, which makes the system both portable and easy to use, and meets the requirements of geological survey field deployment.

5.1 Design of rock classification system

A rock classification system is constructed based on the designed rock image classification model. The overall framework is shown in Figure 13A, and the specific functions of the system are as follows: 1) Get an image of the rock. Images are acquired in real-time from connected industrial cameras, or rock images are fetched from local data. Real-time detection and local data acquisition are introduced to meet the requirements of the geological survey sites. 2) Rock image preprocessing. The rock images that need to be input into the classification network are preprocessed first, the brightness of the rock images that are too bright or too dark is corrected, and the rock images are smoothed to remove the sharp noise, reduce the level of detail, and enhance the recognition effect of the image under different proportions. The preprocessed rock image is used as the input image of the subsequent classification network, and improving the image identifiability is beneficial to improve the accuracy of rock classification. 3) Rock image classification. The rock classification network is loaded, and the preprocessed rock

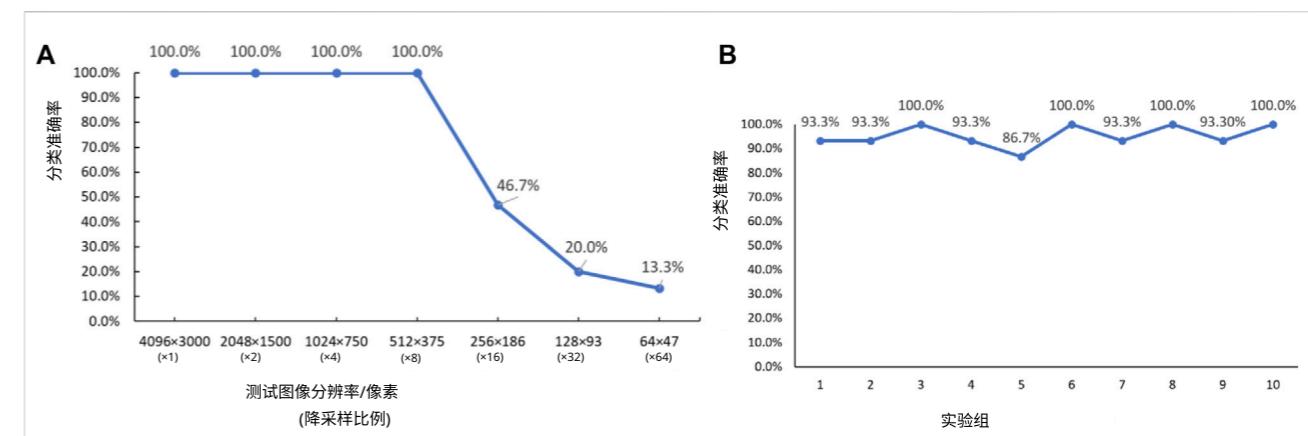


FIGURE 12

图像分辨率对分类结果的影响 (A) 与图像视场变化对分类结果的影响 (B) 所对应的准确率与损失值。

4.4.5 综合分析

通过消融实验证明了本文数据预处理的有效性。由于原始岩石数据集规模过小，图像分割与数据增强能显著提升岩石分类准确率。通过AlexNet、VGG16、GoogLeNet与ResNet34的对比实验证明了残差神经网络的有效性，迁移学习与非迁移学习的对比则证实了迁移学习的优越性。通过测试包含全部七类岩石的15张图像验证了模型的实用价值，所有图像均以高于82%的概率被准确预测，其中多置信度超过95%。模拟相机视角变化与分辨率变化的实际场景测试表明，该模型具有良好的鲁棒性，轻微场景变化不会影响分类精度，同时也印证了数据预处理的有效性。实验结果表明：采用预训练残差神经网络进行迁移学习的模型具有更高分类准确率和良好泛化能力。

5 岩石分类网络的部署与应用

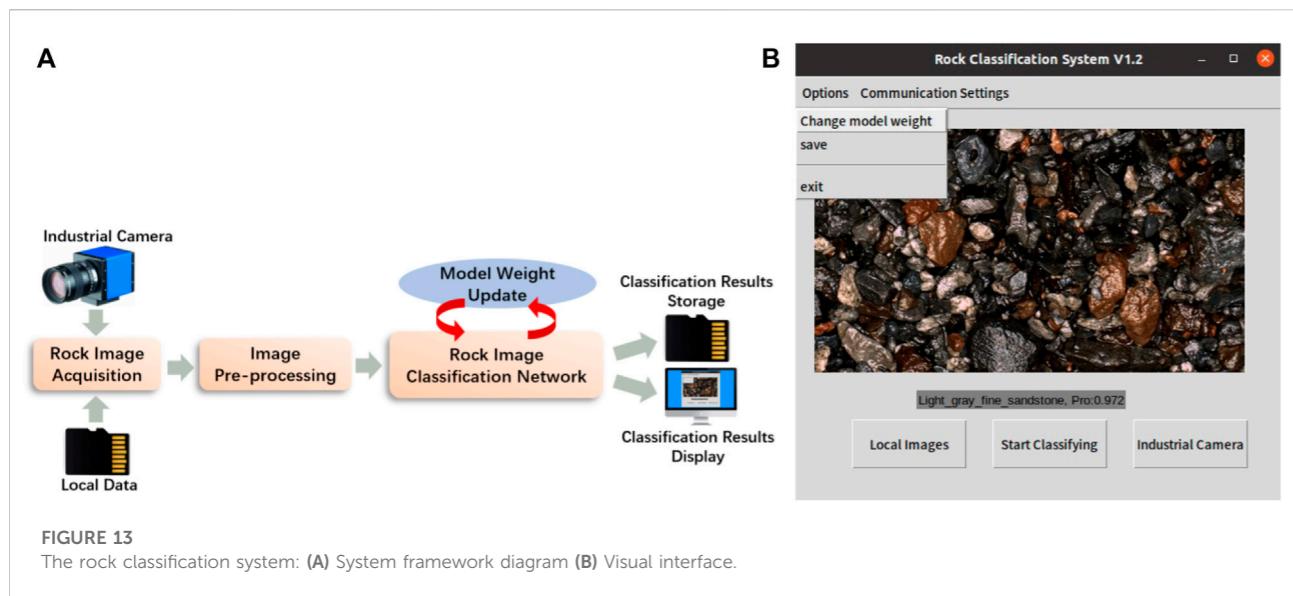
地质调查工作常需在施工现场或离线环境下开展。地质勘查人员需携带各类勘探设备，如GPS测量仪、定位仪、测绘仪器等。工作站等具有一定重量和体积的设备携带不便，且无法通过网络及时获取岩性反馈

来指导后续调查。将本文提出的岩石图像分类模型部署至嵌入式终端设备，可有效解决该问题。

本文将岩石图像分类从理论研究转向实际应用。将训练好的岩石分类网络模型移植至Nvidia Jetson TX2嵌入式平台，采用TensorRT推理优化器加速模型，并开发集成系统各环节的前端界面，使系统兼具便携性与易用性，满足地质调查现场部署需求。

5.1 岩石分类系统设计

基于设计的岩石图像分类模型构建了一套岩石分类系统。整体框架如图13A所示，系统具体功能如下：1) 获取岩石图像。通过连接工业相机实时采集图像，或从本地数据调取岩石图像。引入实时检测与本地数据获取功能以满足地质勘查现场需求。2) 岩石图像预处理。对需输入分类网络的岩石图像进行亮度校正(过亮/过暗)、平滑降噪处理，通过减少细节层次来增强图像在不同比例下的识别效果。预处理后的图像作为后续分类网络的输入，提升图像可辨识度有助于提高岩石分类准确率。3) 岩石图像分类。加载岩石分类网络对预处理后的



images are input for inference to obtain the rock classification results. If it is necessary to extend the rock category or use a better-trained rock model, the rock classification model can be updated by replacing the original model weights with the newly trained model weights. 4) The obtained rock classification results are stored or displayed on the visual interface.

5.2 System deployment

The preliminary development of the rock classification model proposed in this paper is carried out on the PC side, but the size of the PC side is huge, and it is not suitable for deployment in the industrial survey site. In contrast, embedded devices with deep learning computing capabilities are more in line with the needs of geological exploration. To consider the practical application, we port the algorithm from the PC to the embedded platform. Considering that there is usually no network support in the actual survey site, this paper adopts the offline deployment mode. After the model is trained in the PC server in advance, it is deployed on the embedded device.

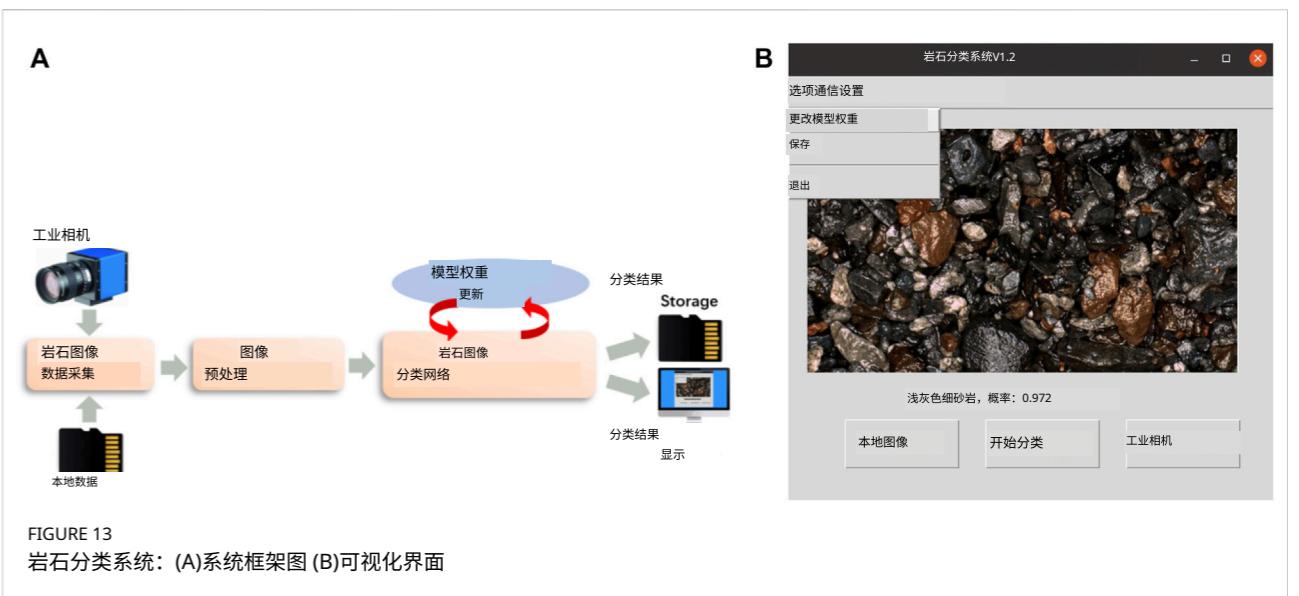
This paper implements the deployment process of the rock classification network on Nvidia Jetson TX2. Nvidia Jetson TX2 is an embedded AI computing device launched by Nvidia Corporation. Its GPU adopts Nvidia Pascal architecture, has 8 GB memory and 32 GB storage space, and is equipped with a variety of standard hardware interfaces. Jetson TX2 is compact and energy efficient, making it ideal for smart edge devices such as robots, drones, and smart cameras.

The deep learning network model trained on the PC usually has a large number of parameters, and it is easy to cause problems of slow inference speed and poor real-time

performance of the model when deployed on embedded devices with weak performance. In order to accelerate the reasoning of the model on embedded devices, the TensorRT framework developed by Nvidia is used to accelerate the reasoning. NVIDIA TensorRT is a special optimizer for neural network inference, which is mostly used in image classification, object detection and other fields. It uses a scheme to optimize the trained model, which can provide low latency and high throughput for deep learning model inference applications deployed in the production environment.

The steps for porting the rock classification algorithm are as follows: First, set up the software development environment on Jetson TX2 and install the libraries that the application depends on to run. Since the trained model is generated by the Pytorch framework and cannot be directly applied to the TensorRT framework, the Pytorch model is first converted to the ONNX (OpenNeural Network Exchange) format to make it suitable for the TensorRT framework. ONNX is a standard format for representing deep learning models that can be transferred between different frameworks (Chang et al., 2020). Many model formats can only be converted to ONNX to work with the TensorRT framework. Finally, the visual interface integrating each function was developed.

The user interface design of this paper takes into account that this system is mainly provided for geological exploration personnel. From the perspective of practical application, the code is encapsulated, and the PYQT module in Python is used for visual interface design. The rocks are classified through the visual interface, and the classification results are displayed and saved, which reduces the threshold of use and facilitates the use of engineers. The main interface is shown in Figure 13B.



输入图像进行推理获取岩石分类结果。如需扩展岩石类别或使用训练更优的岩石模型，可通过替换新训练模型权重来更新分类模型。4)最终分类结果将存储或显示在可视化界面上。

5.2 系统部署

本文提出的岩石分类模型初步开发在PC端完成，但PC端体积庞大，不适用于工业勘测现场部署。相比之下，具备深度学习计算能力的嵌入式设备更符合地质勘探需求。为考虑实际应用，我们将算法从PC端移植到嵌入式平台。鉴于实际勘测现场通常无网络支持，本文采用离线部署模式，模型在PC服务器端预先训练后部署至嵌入式设备。

本文在Nvidia Jetson TX2上实现了岩石分类网络的部署。该设备是英伟达公司推出的嵌入式AI计算设备，其GPU采用Pascal架构，配备8GB内存和32GB存储空间，具有多种标准硬件接口。Jetson TX2体积紧凑且能效优异，是机器人、无人机和智能摄像头等智能边缘设备的理想选择。

在PC端训练的深度学习网络模型通常参数量庞大，部署到性能较弱的嵌入式设备时容易导致推理速度缓慢、模型实时性差的问题。

为加速模型在嵌入式设备上的推理，采用Nvidia开发的TensorRT框架进行加速。NVIDIA TensorRT是专用于神经网络推理的优化器，主要应用于图像分类、目标检测等领域。该框架通过对训练好的模型进行优化，能为生产环境中部署的深度学习推理应用提供低延迟和高吞吐量。

岩性分类算法移植步骤如下：首先，在Jetson TX2上搭建软件开发环境并安装应用程序运行所需的依赖库。由于训练好的模型由Pytorch框架生成，无法直接应用于TensorRT框架，需先将Pytorch模型转换为ONNX（开放神经网络交换）格式以适应TensorRT框架。ONNX是深度学习模型的标准化表示格式，可实现不同框架间的模型转换（Chang等，2020）。许多模型格式需先转换为ONNX才能与TensorRT框架兼容。最终开发了集成各功能的可视化界面。

本文用户界面设计考虑到该系统主要面向地质勘探人员，从实际应用角度出发，对代码进行封装，采用Python中的PYQT模块进行可视化界面设计。通过可视化界面实现岩石分类，并展示保存分类结果，降低了使用门槛，便于工程师操作。主界面如图 13 B 所示。

6 Conclusion

In this study, a deep residual neural network model with transfer learning method is proposed to classify rock images quickly and accurately. The dataset is expanded by image slicing and data augment, and the Resnet34 is pre-trained by the Texture Library dataset for transfer learning. The comparative analysis shows that the model using transfer learning in ResNet34 structure for rock image classification has an excellent effect, and the classification accuracy is as high as 99.1%, which achieves an 11% improvement compared with the model without transfer learning. The excellent performance of the rock classification model is mainly due to the introduction of the residual module and the application of transfer learning. The pre-trained network based on the texture dataset learns rich texture spatial structure features and morphological correlation. Finally, a rock classification system is designed and deployed on embedded devices to meet geological survey tasks. The system extracts feature by the convolutional neural network without manual operation, which reduces the influence of subjective factors. This system has low requirements for rock image acquisition configuration and environment, which fully demonstrates its robustness and generalization ability.

Our future study will further increase the number of rock categories and ensure that the classification accuracy is further improved when more rock types are added, as the types and number of rock datasets in this paper are limited due to the limitations of shooting conditions.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <http://www.olegvolk.net/gallery/various/textures>.

Author contributions

WC was responsible for the implementation of the proposed method and the writing of the manuscript. LS was responsible for making important revisions to the manuscript to make it clearer

References

- Ali, S. B., Wate, R., Kujur, S., Singh, A., and Kumar, S. (2020). "Wall crack detection using transfer learning-based cnn models," in 2020 IEEE 17th India Council International Conference (INDICON) (IEEE), 1–7.
- Bai, L., Wei, X., Liu, Y., Wu, C., and Chen, L. (2019). Rock thin section image recognition and classification based on vgg model. *Geol. Bull. China* 38, 2053–2058.
- Bai, L., Yao, Y., Li, S., Xu, D., and Wei, X. (2018). Mineral composition analysis of rock image based on deep learning feature extraction. *China Min. Mag.* 27, 178–182.
- Chang, Y.-M., Liao, W.-C., Wang, S.-C., Yang, C.-C., and Hwang, Y.-S. (2020). A framework for scheduling dependent programs on gpu architectures. *J. Syst. Archit.* 106, 101712. doi:10.1016/j.jysarc.2020.101712
- Chen, J., Yang, T., Zhang, D., Huang, H., and Tian, Y. (2021). Deep learning based classification of rock structure of tunnel face. *Geosci. Front.* 12, 395–404. doi:10.1016/j.gsf.2020.04.003
- Cheng, G., Guo, W., and Fan, P. (2017). Study on rock image classification based on convolution neural network. *J. Xi'an Shiyou Univ. Nat. Sci. Ed.* 32, 116–122.
- Dabrowski, M., and Michalik, T. (2017). How effective is transfer learning method for image classification. Position papers of the Federated Conference on Computer Science and Information Systems, 3–9.
- Dai, J., He, K., and Sun, J. (2016). "Instance-aware semantic segmentation via multi-task network cascades," in Proceedings of the IEEE conference on computer vision and pattern recognition, 3150–3158.

6 结论

本研究提出采用迁移学习方法的深度残差神经网络模型，实现岩石图像的快速精准分类。通过图像切片与数据增强扩充数据集，并基于纹理库数据集对Resnet34进行迁移学习预训练。对比分析表明，在ResNet34结构中应用迁移学习的岩石图像分类模型效果优异，分类准确率高达99.1%，较未采用迁移学习的模型提升11%。该岩石分类模型的卓越性能主要得益于残差模块的引入和迁移学习的应用。基于纹理数据集预训练的网络学习了丰富的纹理空间结构特征与形态关联性。最终设计并部署于嵌入式设备的岩石分类系统可满足地质勘测任务需求，该系统通过卷积神经网络自动提取特征，避免了人工操作的主观因素影响，且对岩石图像采集配置与环境要求较低，充分展现了其鲁棒性与泛化能力。

由于拍摄条件限制，本文岩石数据集的类型和数量有限。我们后续研究将进一步增加岩石类别数量，确保在添加更多岩石类型时分类精度持续提升。

数据可用性声明

本研究分析了公开数据集，数据来源详见：
<http://www.olegvolk.net/gallery/various/textures>

作者贡献

WC负责方法实现与论文撰写；LS负责对论文进行重要修改以提高清晰度

参考文献

- Ali, S. B., Wate, R., Kujur, S., Singh, A., and Kumar, S. (2020). "Wall crack detection using transfer learning-based cnn models," in 2020 IEEE 17th India Council International Conference (INDICON) (IEEE), 1–7.
- Bai, L., Wei, X., Liu, Y., Wu, C., and Chen, L. (2019). Rock thin section image recognition and classification based on vgg model. *Geol. Bull. China* 38, 2053–2058.
- Bai, L., Yao, Y., Li, S., Xu, D., and Wei, X. (2018). Mineral composition analysis of rock image based on deep learning feature extraction. *China Min. Mag.* 27, 178–182.
- Chang, Y.-M., Liao, W.-C., Wang, S.-C., Yang, C.-C., and Hwang, Y.-S. (2020). A framework for scheduling dependent programs on gpu architectures. *J. Syst. Archit.* 106, 101712. doi:10.1016/j.jysarc.2020.101712
- 陈杰、杨涛、张栋、黄海、田阳 (2021)。基于深度学习的方法隧道掌子面岩体结构分类. 地学前缘 12, 395-404. doi:10.1016/j.gsf.2020.04.003
- 程光、郭伟、范鹏 (2017)。基于卷积神经网络的岩石图像分类研究。西安石油大学学报(自然科学版) 32, 116-122.
- Dabrowski, M. 与 Michalik, T. (2017). 迁移学习方法在图像分类中的有效性评估. 联邦计算机科学与信息系统会议立场报告, 3-9.
- 戴杰, 何恺明, 孙剑 (2016). "基于多任务网络级联的实例感知语义分割", 发表于IEEE计算机视觉与模式识别会议论文集, 3150-3158.

与合理性，并提出方法改进思路；XC负责构建初始研究构想并分析实验结果的合理性。ZH负责论文的语法检查与语言润色。

资助声明

本研究部分由国家自然科学基金（项目编号61903315）和龙岩市科技合作计划项目（项目编号2020LYF16004）资助；部分由福建省科技厅自然科学基金（项目编号2022J011255）资助。

致谢

作者感谢所有编辑和评审专家的意见与建议。特别鸣谢广东天鹏智能科技有限公司为本研究提供数据支持。

利益冲突

作者声明本研究不存在任何可能被视为潜在利益冲突的商业或财务关系。

出版者说明

本文表达的所有主张仅代表作者个人观点，未必反映其附属机构、出版商、编辑及审稿人的立场。对于文章中可能评估的任何产品或制造商作出的声明，出版商均不予以保证或认可。

- Falivene, O., Auchter, N. C., de Lima, R. P., Kleipool, L., Solum, J. G., Zarian, P., et al. (2022). Lithofacies identification in cores using deep learning segmentation: Turbidite deposits (gulf of Mexico and north sea) and the geoscientists role. *AAPG Bull.*
- Fan, G., Chen, F., Chen, D., Li, Y., and Dong, Y. (2020). A deep learning model for quick and accurate rock recognition with smartphones. *Mob. Inf. Syst.* 2020, 1–14. doi:10.1155/2020/7462524
- Feng, Y., Gong, X., Xu, Y., Xie, Z., Cai, H., and Lv, X. (2019). Lithology recognition based on fresh rock images and twins convolution neural network. *Geogr. Geo-Information Sci.* 35, 89–94.
- Gao, M., Qi, D., Mu, H., and Chen, J. (2021). A transfer residual neural network based on resnet-34 for detection of wood knot defects. *Forests* 12, 212. doi:10.3390/f12020212
- Gonçalves, L. B., and Leta, F. R. (2010). Macroscopic rock texture image classification using a hierarchical neuro-fuzzy class method. *Math. problems Eng.* 2010, 1–23. doi:10.1155/2010/163635
- Guo, W., Dong, C., Lin, C., Wu, Y., Zhang, X., and Liu, J. (2022). Rock physical modeling of tight sandstones based on digital rocks and reservoir porosity prediction from seismic data. *Front. Earth Sci. (Lausanne)*. 10, 932929. doi:10.3389/feart.2022.932929
- Hang, J., Zhang, D., Chen, P., Zhang, J., and Wang, B. (2019). Classification of plant leaf diseases based on improved convolutional neural network. *Sensors* 19, 4161. doi:10.3390/s19194161
- He, K., Zhang, X., Ren, S., and Sun, J. (2016a). “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016b). “Identity mappings in deep residual networks,” in European conference on computer vision (Springer), 630–645.
- Houshmand, N., GoodFellow, S., Esmaeili, K., and Calderón, J. C. O. (2022). Rock type classification based on petrophysical, geochemical, and core imaging data using machine and deep learning techniques. *Appl. Comput. Geosciences* 16, 100104. doi:10.1016/j.acags.2022.100104
- Hu, Q., Ye, W., Wang, Q., and Chen, Y. (2020). Recognition of lithology with big data of geological images. *J. Eng. Geol.* 28, 1433–1440.
- Iglesias, J. C. A., Santos, R. B. M., and Paciornik, S. (2019). Deep learning discrimination of quartz and resin in optical microscopy images of minerals. *Miner. Eng.* 138, 79–85. doi:10.1016/j.mineeng.2019.04.032
- Imamverdiyev, Y., and Sukhostat, L. (2019). Lithological facies classification using deep convolutional neural network. *J. Petroleum Sci. Eng.* 174, 216–228. doi:10.1016/j.petrol.2018.11.023
- Karimpouli, S., and Tahmasebi, P. (2019). Segmentation of digital rock images using deep convolutional autoencoder networks. *Comput. geosciences* 126, 142–150. doi:10.1016/j.cageo.2019.02.003
- Koeshidayatullah, A., Al-Azani, S., Baraboshkin, E. E., and Alfarraj, M. (2022). Faciesvit: Vision transformer for an improved core lithofacies prediction. *Front. Earth Sci. (Lausanne)*. 10, 992442. doi:10.3389/feart.2022.992442
- Li, J., Zhang, L., Wu, Z., Ling, Z., Cao, X., Guo, K., et al. (2020). Autonomous martian rock image classification based on transfer deep learning methods. *Earth Sci. Inf.* 13, 951–963. doi:10.1007/s12145-019-00433-9
- Liang, Y., Cui, Q., Luo, X., and Xie, Z. (2021). Research on classification of fine-grained rock images based on deep learning. *Comput. Intell. Neurosci.* 2021, 1–11. doi:10.1155/2021/5779740
- Liu, Y.-Z., Ren, S.-F., and Zhao, P.-F. (2022). Application of the deep neural network to predict dynamic responses of stiffened plates subjected to near-field underwater explosion. *Ocean. Eng.* 247, 110537. doi:10.1016/j.oceaneng.2022.110537
- Luo, Y., and Wang, Z. (2021). “An improved resnet algorithm based on cbam,” in 2021 International Conference on Computer Network Electronic and Automation (ICCNEA), Xi'an, China, 24–26 September 2021 (IEEE), 121–125. doi:10.1109/ICCNEA52021.9509902
- Mlynarczuk, M., Górszczyk, A., and Ślipek, B. (2013). The application of pattern recognition in the automatic classification of microscopic rock images. *Comput. Geosciences* 60, 126–133. doi:10.1016/j.cageo.2013.07.015
- Nair, V., and Hinton, G. E. (2010). “Rectified linear units improve restricted Boltzmann machines,” in Icml.
- Pan, J. (2017). “Review of metric learning with transfer learning,” in AIP Conference Proceedings (Melville, NY: AIP Publishing LLC), 020040.1864.
- Patel, A. K., and Chatterjee, S. (2016). Computer vision-based limestone rock-type classification using probabilistic neural network. *Geosci. Front.* 7, 53–60. doi:10.1016/j.gsf.2014.10.005
- Pham, C., and Shin, H.-S. (2020). A feasibility study on application of a deep convolutional neural network for automatic rock type classification. *Tunn. Undergr. space* 30, 462–472.
- Ran, X., Xue, L., Zhang, Y., Liu, Z., Sang, X., and He, J. (2019). Rock classification from field image patches analyzed using a deep convolutional neural network. *Mathematics* 7, 755. doi:10.3390/math7080755
- Ru, Z., and Jiong, M. (2019). Identification and evaluation of logging methods for quartz sandstone gas reservoir in yulin gas field. *J. Liaoning Univ. Petroleum Chem. Technol.* 39, 65.
- Sharif, H., Ralchenko, M., Samson, C., and Ellery, A. (2015). Autonomous rock classification using bayesian image analysis for rover-based planetary exploration. *Comput. Geosciences* 83, 153–167. doi:10.1016/j.cageo.2015.05.011
- ShuTeng, X., and YongZhang, Z. (2018). Artificial intelligence identification of ore minerals under microscope based on deep learning algorithm. *Acta Petrol. Sin.* 34, 3244–3252.
- Singh, N., Singh, T., Tiwary, A., and Sarkar, K. M. (2010). Textural identification of basaltic rock mass using image processing and neural network. *Comput. Geosci.* 14, 301–310. doi:10.1007/s10596-009-9154-x
- Su, C., Xu, S.-j., Zhu, K.-y., and Zhang, X.-c. (2020). Rock classification in petrographic thin section images based on concatenated convolutional neural networks. *Earth Sci. Inf.* 13, 1477–1484. doi:10.1007/s12145-020-00505-1
- Wang, Y., and Sun, S. (2021). Image-based rock typing using grain geometry features. *Comput. Geosciences* 149, 104703. doi:10.1016/j.cageo.2021.104703
- Xiao, D., Le, B. T., and Ha, T. T. L. (2021). Iron ore identification method using reflectance spectrometer and a deep neural network framework. *Spectrochimica Acta Part A Mol. Biomol. Spectrosc.* 248, 119168. doi:10.1016/j.saa.2020.119168
- Xiao, Q., Li, W., Kai, Y., Chen, P., Zhang, J., and Wang, B. (2019). Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinforma.* 20, 1–15. doi:10.1186/s12859-019-3262-y
- Yan, Q., Yang, B., Wang, W., Wang, B., Chen, P., and Zhang, J. (2020). Apple leaf diseases recognition based on an improved convolutional neural network. *Sensors* 20, 3535. doi:10.3390/s20123535
- Yang, B., Wang, M., Sha, Z., Wang, B., Chen, J., Yao, X., et al. (2019). Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors* 19, 4416. doi:10.3390/s19204416
- Yi, X., Li, H., Zhang, R., Gongqiu, Z., He, X., Liu, L., et al. (2022). Rock mass structural surface trace extraction based on transfer learning. *Open Geosci.* 14, 98–110. doi:10.1515/geo-2022-0337
- Zeng, X., Xiao, Y., Ji, X., and Wang, G. (2021). Mineral identification based on deep learning that combines image and mohs hardness. *Minerals* 11, 506. doi:10.3390/min11050506
- Zhang, Y., Li, M., and Han, S. (2018). Automatic identification and classification in lithology based on deep learning in rock images. *Yanshi Xuebao/Acta Petrol. Sin.* 34, 333–342.
- Zhu, W., Braun, B., Chiang, L. H., and Romagnoli, J. A. (2021). Investigation of transfer learning for image classification and impact on training sample size. *Chemom. Intelligent Laboratory Syst.* 211, 104269. doi:10.1016/j.chemolab.2021.104269
- Falivene, O., Auchter, N. C., de Lima, R. P., Kleipool, L., Solum, J. G., Zarian, P., et al. (2022). Lithofacies identification in cores using deep learning segmentation: Turbidite deposits (gulf of Mexico and north sea) and the geoscientists role. *AAPG Bull.*
- Nair, V., and Hinton, G. E. (2010). “Rectified linear units improve restricted Boltzmann machines,” in Icml.
- Pan, J. (2017). “Review of metric learning with transfer learning,” in AIP Conference Proceedings (Melville, NY: AIP Publishing LLC), 020040.1864.
- Patel, A. K., and Chatterjee, S. (2016). Computer vision-based limestone rock-type classification using probabilistic neural network. *Geosci. Front.* 7, 53–60. doi:10.1016/j.gsf.2014.10.005
- Pham, C., and Shin, H.-S. (2020). A feasibility study on application of a deep convolutional neural network for automatic rock type classification. *Tunn. Undergr. space* 30, 462–472.
- Ran, X., Xue, L., Zhang, Y., Liu, Z., Sang, X., and He, J. (2019). Rock classification from field image patches analyzed using a deep convolutional neural network. *Mathematics* 7, 755. doi:10.3390/math7080755
- Ru, Z., and Jiong, M. (2019). Identification and evaluation of logging methods for quartz sandstone gas reservoir in yulin gas field. *J. Liaoning Univ. Petroleum Chem. Technol.* 39, 65.
- Sharif, H., Ralchenko, M., Samson, C., and Ellery, A. (2015). Autonomous rock classification using bayesian image analysis for rover-based planetary exploration. *Comput. Geosciences* 83, 153–167. doi:10.1016/j.cageo.2015.05.011
- ShuTeng, X., and YongZhang, Z. (2018). Artificial intelligence identification of ore minerals under microscope based on deep learning algorithm. *Acta Petrol. Sin.* 34, 3244–3252.
- Singh, N., Singh, T., Tiwary, A., and Sarkar, K. M. (2010). Textural identification of basaltic rock mass using image processing and neural network. *Comput. Geosci.* 14, 301–310. doi:10.1007/s10596-009-9154-x
- Su, C., Xu, S.-j., Zhu, K.-y., and Zhang, X.-c. (2020). Rock classification in petrographic thin section images based on concatenated convolutional neural networks. *Earth Sci. Inf.* 13, 1477–1484. doi:10.1007/s12145-020-00505-1
- Wang, Y., and Sun, S. (2021). Image-based rock typing using grain geometry features. *Comput. Geosciences* 149, 104703. doi:10.1016/j.cageo.2021.104703
- Xiao, D., Le, B. T., and Ha, T. T. L. (2021). Iron ore identification method using reflectance spectrometer and a deep neural network framework. *Spectrochimica Acta Part A Mol. Biomol. Spectrosc.* 248, 119168. doi:10.1016/j.saa.2020.119168
- Xiao, Q., Li, W., Kai, Y., Chen, P., Zhang, J., and Wang, B. (2019). Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinforma.* 20, 1–15. doi:10.1186/s12859-019-3262-y
- Yan, Q., Yang, B., Wang, W., Wang, B., Chen, P., and Zhang, J. (2020). Apple leaf diseases recognition based on an improved convolutional neural network. *Sensors* 20, 3535. doi:10.3390/s20123535
- Yang, B., Wang, M., Sha, Z., Wang, B., Chen, J., Yao, X., et al. (2019). Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors* 19, 4416. doi:10.3390/s19204416
- Yi, X., Li, H., Zhang, R., Gongqiu, Z., He, X., Liu, L., et al. (2022). Rock mass structural surface trace extraction based on transfer learning. *Open Geosci.* 14, 98–110. doi:10.1515/geo-2022-0337
- Zeng, X., Xiao, Y., Ji, X., and Wang, G. (2021). Mineral identification based on deep learning that combines image and mohs hardness. *Minerals* 11, 506. doi:10.3390/min11050506
- Zhang, Y., Li, M., and Han, S. (2018). Automatic identification and classification in lithology based on deep learning in rock images. *Yanshi Xuebao/Acta Petrol. Sin.* 34, 333–342.
- Zhu, W., Braun, B., Chiang, L. H., and Romagnoli, J. A. (2021). Investigation of transfer learning for image classification and impact on training sample size. *Chemom. Intelligent Laboratory Syst.* 211, 104269. doi:10.1016/j.chemolab.2021.104269