

doi: [10.12097/gbc.2022.08.005](https://doi.org/10.12097/gbc.2022.08.005)

基于改进 ConvNeXt 网络的矿物图像智能识别

李雷^{1,2}, 卢才武^{1,2*}, 江松^{1,2}, 景文刚^{1,2}, 王洛锋³

LI Lei^{1,2}, LU Caiwu^{1,2*}, JIANG Song^{1,2}, JING Wengang^{1,2}, WANG Luofeng³

1. 西安建筑科技大学资源工程学院, 陕西 西安 710055;

2. 西安市智慧工业感知、计算与决策重点实验室, 陕西 西安 710055;

3. 洛阳栾川钼业集团股份有限公司, 河南 洛阳 471500

1. College of Resource Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China;

2. Xi'an Key Laboratory for Intelligent Industrial Perception, Calculation and Decision, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China;

3. Luoyang Luanchuan Molybdenum Industry Group Co., Ltd., Luoyang 471500, Henan, China

摘要: 矿物识别是地质研究的重要工作, 但是如何准确识别矿物仍然是一项重要的挑战。针对矿物形态特征, 提出了一种利用迁移学习策略并引入通道注意力的改进 ConvNeXt 网络矿物图像智能识别模型。首先, 利用 ImageNet 数据集上已预训练的 ConvNeXt 网络模型, 运用迁移学习的方式, 加载到矿物识别模型中; 其次, 在 ConvNeXt 网络的基础上, 以 ConvNeXt 块之后与注意力机制相结合的方式, 进一步提升其特征融合能力; 最后, 以 26 类矿物的矿石图像为研究对象, 总计 34576 张图像, 以 6:2:2 比例划分训练集、验证集与测试集, 模型在实验训练过程中与 VGG19、GoogLeNet、ResNet50、ResNeXt50 和 ConvNeXt 网络相比, 收敛速度明显加快。实验结果表明, 矿物智能识别模型在准确率、精确率和召回率上分别达到 98.58%、98.62% 和 98.73%, 而消融实验证明本文提出的优化方法有助于提升模型性能, 同时, 通过对不同模型矿物图像特征图的可视化对比分析, 验证了本文提出的矿物识别模型对于矿物特征的准确提取, 进一步证明了模型的有效性, 提高了矿物识别的准确率。

关键词: 矿物图像; ConvNeXt; 迁移学习; 注意力机制; 矿物识别

中图分类号: P57; P628 **文献标志码:** A **文章编号:** 1671-2552(2024)07-1266-10

Li L, Lu C W, Jiang S, Jing W G, Wang L F. Intelligent mineral image recognition based on improved ConvNeXt network. *Geological Bulletin of China*, 2024, 43(7): 1266-1275

Abstract: Mineral identification is a critical task in geological research, yet accurately identifying minerals remains a significant challenge. This study proposes an intelligent mineral image recognition model based on an improved ConvNeXt network, which utilizes transfer learning strategies and incorporates channel attention mechanisms to address the morphological characteristics of minerals. Firstly, the ConvNeXt network model pre-trained on the ImageNet dataset is employed and integrated into the mineral recognition model through transfer learning. Secondly, based on the ConvNeXt network, the model enhances feature fusion capabilities by combining the ConvNeXt blocks with attention mechanisms. Finally, a dataset comprising 34576 ore images of 26 mineral categories is used, divided into training, validation, and test sets in a 6:2:2 ratio. During experimental training, the proposed model demonstrates a significantly faster convergence compared to VGG19, GoogLeNet, ResNet50, ResNeXt50, and the ConvNeXt networks. Experimental results indicate that the intelligent mineral recognition model achieves an accuracy, precision, and recall of 98.58%, 98.62%, and 98.73%, respectively. Ablation experiments confirm that the optimization methods proposed in this study

收稿日期: 2022-08-08; 修订日期: 2022-10-08

资助项目: 国家自然科学基金面上项目《地下金属矿山岩体破坏多源异质流数据智能融合与态势评估研究》(批准号: 51974223)和国家自然科学基金青年项目《基于数据-知识混合驱动的露天矿复杂边坡灾害识别与预警》(批准号: 52104146)

作者简介: 李雷(1998-), 男, 硕士, 从事矿物智能识别方面的研究。E-mail: lilei_xauat@126.com

* 通信作者: 卢才武(1965-), 男, 博士, 教授, 从事矿山系统工程研究。E-mail: lucaiwu@126.com

enhance model performance. Additionally, comparative visual analysis of feature maps from different models substantiates that the proposed mineral recognition model accurately extracts mineral features, further validating the model's effectiveness and improving mineral identification accuracy.

Key words: mineral image; ConvNeXt; transfer learning; attention mechanism; mineral recognition

矿物识别是采矿工程、地质学等相关学科研究的基础。国内目前主要的矿物识别方法是依据人工经验,通常是依据其颜色、硬度和透明度,同时结合结晶习性、解理、断口等物理特征判断矿物种类(赵明, 2010)。这对于从事矿物识别及相关工作人员的专业知识与经验要求较高,而且对矿石识别的准确度较低,难以大规模应用于实际场景。近年来,为了提高矿石识别效率及解决识别准确度低的问题,以机器学习为代表的矿物智能识别成为了研究的热点(周永章等, 2018; 郭艳军等, 2020; 郝慧珍等, 2021)。

常规的矿物识别方法主要是依据矿物的属性,如外表形态、物理性质、化学成分等属性特征及相应的组合,使用能量散射光谱(EDS)(El Haddad et al., 2019)、电子探针(Tsuji et al., 2010)、激光诱导击穿光谱(LIBS)(Khajehzadeh et al., 2016)、显微光学观察(Maitre et al., 2019)、光谱分析(张旭等, 2020)等方法实现矿物与岩石的识别。这些方法虽然准确率较高,但是需要用到专业仪器,过程复杂,整体效率偏低。

近年来,深度学习在矿物识别领域取得了良好效果(Trejbal et al., 2020; 许振浩等, 2021; 周永章等, 2021)。Baykan et al.(2010)利用多层感知机(MLP, Multilayer Perceptron)识别石英、白云母、黑云母等矿物。Izadi et al.(2013)利用两层神经网络,成功识别 14 种矿物。徐述腾等(2018)设计了 Unet 模型,建立对矿石矿物镜下照片的智能识别。Liu et al.(2019)基于深度学习和迁移学习算法,以 12 种岩石矿物为研究对象,建立了岩石矿物识别模型。彭伟航等(2019)基于改进 InceptionV3 模型对常见的 16 类矿物进行图像识别,准确率达到 86%。李明超等(2020)基于迁移学习的方法与深度学习模型,根据 19 类矿物的颜色与纹理特征,建立了矿物图像识别耦合模型,并搭建矿物识别应用系统。刘艳鹏等(2020)利用卷积神经网络进行成矿预测。王李管等(2020)利用预训练的卷积神经网络对黑钨矿图像进行迁移学习实验,对黑钨原矿石的识别准确率达到 99.6%。Liu et al.(2021)结合模型深度、模型结构和数据集大小,建立了小型矿石图像分类的深度学习

模型。Zhou et al.(2022)在 MobileNet 网络的基础上,利用迁移学习与 SENet 机制相结合的方式,使矿石分类的准确率达到 96%。通过上述文献不难看出,基于深度学习的矿物识别可以较好地解决传统方法中效率偏低的问题,并且准确率有所提高,但仍存在矿物识别种类较少,同时矿物形态特征发生变化时,已有的深度学习模型难以准确提取特征,以致于泛化能力不强的问题。

针对现有问题,本文以 26 类矿物的矿石图像作为研究对象,提出了一种基于改进的 ConvNeXt 网络的矿物智能识别模型,通过利用迁移学习策略,引入超强通道注意力机制,在加强矿物特征提取的同时,进一步提高特征之间的融合,以及矿物识别准确率与模型的泛化性,从而实现对不同矿物的高效准确识别。

1 基本概念

在计算机视觉中,卷积神经网络凭借其其对图像特征的快速提取与识别物体准确率高一直占据着重要地位。ConvNeXt 网络是 2022 年由 Facebook 团队提出的纯卷积神经网络架构模型(Liu et al., 2022)。对于 ConvNeXt 网络,根据模型计算复杂度的不同,共有 4 个(T/S/B/L)版本。本文综合考虑矿物的类别与数据集的大小,选取了 ConvNeXt-T 版本。在 ConvNeXt-T 网络中,主要是由 ConvNeXt 块构成。对于 ConvNeXt 块,其输入特征图为 $h \times w \times dim$ (高度 $\times$ 宽度 $\times$ 特征图的通道深度),经过深度可分离卷积(DepthwiseConv2D)及层标准化(Layer Normalization),对多个特征通道进行融合,之后经过普通卷积(Conv2D)等操作,对特征图进行一系列升维与降维处理,使其输出特征图为 $h \times w \times dim$,具体变化过程如图 1 所示。对于整个 ConvNeXt-T 网络结构,首先输入大小为 $224 \times 224 \times 3$ (即高度与宽度大小为 224,通道数为 3)的图像,经过第一次卷积操作后特征图变为 $56 \times 56 \times 96$;其次,经过一系列 ConvNeXt 块同时结合下采样(Downsample)操作,特征图变为 $7 \times 7 \times 768$,最后,经全局平均池化(Global

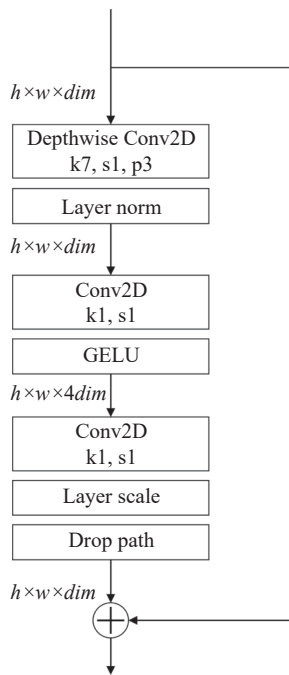


图 1 ConvNeXt 块

Fig. 1 ConvNeXt Block

Avg Pooling)、层标准化与全连接(Linear)等操作后,特征图输出大小为 1000。具体结构如表 1 所示。

2 模型构建

2.1 迁移学习

理论上,深度学习可以提取图像中不同层次的特征,表征物体之间的细微差别。然而,这需要以足够多且高质量的已标注数据集为前提,进而训练得出已建立网络模型中的巨量参数。在实际应用场景中,研究对象(如矿物图像领域)往往没有建立大量带标签的数据集,通常会采用迁移学习的方式来解决(Zamir et al., 2018)。目的主要是实现模型之间的参数共享与相应的特征迁移,提高其泛化能力,同时减少训练成本。

考虑到研究的矿物对象种类及数据集大小,本文通过迁移学习的方式,实现 ConvNeXt 原模型与-矿物识别模型之间参数共享,以降低其训练成本,提高矿物智能识别模型的泛化性。①利用原有的 ConvNeXt 网络对 ImageNet-1k 数据集进行预训练,得到 ConvNeXt 网络预训练模型与对应的模型权重文件;②通过微调的方式,删除上述预训练模型中原有的全连接层,并设计适合矿物研究对象的全连接层,从而适合本文研究的矿物识别领域;③基于迁移

表 1 ConvNeXt-T 结构

Table 1 Structure diagram of ConvNeXt-T

层名	输入	ConvNeXt-T	输出
conv1	224 × 224 × 3	4 × 4, 96, stride4 Layer Norm	56 × 56 × 96
conv2_x	56 × 56 × 96	$\begin{bmatrix} d7 \times 7, 96 \\ 1 \times 1, 384 \\ 1 \times 1, 96 \end{bmatrix} \times 3$	56 × 56 × 96
conv3_x	56 × 56 × 96	Downsample $\begin{bmatrix} d7 \times 7, 192 \\ 1 \times 1, 768 \\ 1 \times 1, 192 \end{bmatrix} \times 3$	28 × 28 × 192
conv4_x	28 × 28 × 192	Downsample $\begin{bmatrix} d7 \times 7, 384 \\ 1 \times 1, 1536 \\ 1 \times 1, 384 \end{bmatrix} \times 9$	14 × 14 × 384
conv5_x	14 × 14 × 384	Downsample $\begin{bmatrix} d7 \times 7, 768 \\ 1 \times 1, 3072 \\ 1 \times 1, 768 \end{bmatrix} \times 3$	7 × 7 × 768
	7 × 7 × 768	Global Avg PoolingLayer NormalizationLinear	1000

学习的方式,将以矿物种类的数量为标准而设计的全连接层与 ConvNeXt 模型相结合,以共享参数信息与特征权重,从而重新训练模型参数,加快训练速度。

2.2 超强通道注意力模块 (ECA-Net)

现有研究表明,将注意力机制添加到已有的卷积神经网络框架或模型中可以提高其性能。比如,SENet 通过对特征图中的通道维度进行信息权重分配,在 ImageNet 数据集上降低了 0.4%~1.8% 的错误率(Hu et al., 2018);SKNet 采用自适应的注意力编码机制,与 ResNeXt50 相比降低了 1.44% 的错误率(Li et al., 2019);CBAM 提出了通道与空间融合的注意力机制,较 ResNet50 降低了 1.9% 的错误率(Woo et al., 2018)。在复杂的环境变化下,矿物的特征发现多样化,本文在 ConvNeXt 网络基础上,在 ConvNeXt 块之后加入超强通道注意力机制 ECA(Efficient Channel Attention)(Wang et al., 2020),加强了对矿物特征的提取。其主要流程为:首先,输入矿物图像特征图大小为 $C \times H \times W$;在此基础上,进行 GAP(全局平均池化)操作以减少模型参数量;其次,使用一维卷积(大小为 k)获取矿物图像特征在每个空间通道的权重,其中, k 表示局部通道交互的覆盖率,同时利用 Sigmoid 进行对应权重的归

一化;最后,将已生成的各个通道权重与输入特征图相加权,进而得到矿物图像特征权重。超强通道注意力(ECA)模块如图 2 所示。

对于整个 ECA 模块来讲,主要是用 W_k 来表示学习到的通道注意力,其中, W_k 涉及 $K \times C$ 个参数,并且 W_k 避免了不同通道的完全独立:

$$\begin{bmatrix} w^{1,1} & \dots & w^{1,k} & 0 & 0 & \dots & \dots & 0 \\ 0 & w^{2,2} & \dots & w^{2,k+1} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & w^{C,C-k+1} & \dots & w^{C,C} \end{bmatrix} \quad (1)$$

其中,ECA 模块主要的部分是:利用卷积核大小为 k 的一维卷积实现每个通道之间的信息交互,从而生成每个通道对应的权重:

$$w = \sigma(C1D_k(y)) \quad (2)$$

其中,C1D 为一维卷积, σ 为 Sigmoid 函数。因此,ECA 在保证了模型性能的同时,提高了模型效率(降低了模型复杂度)。

2.3 矿物智能识别模型

针对矿物图像识别困难及准确率较低的缺陷,本文在原有 ConvNeXt 网络的基础上,利用迁移学习策略与引入 ECA 模块,构建矿物智能识别网络模型,名称为 ECA-ConvNeXt 网络模型。具体来讲,主要是利用在 ImageNet 已经预训练的 ConvNeXt 模型,经过一定处理后,通过迁移学习加载到矿物识别模型框架内,而在现有 ConvNeXt 网络的基础上,为了实现矿物图像特征之间的融合,在每一个 ConvNeXt 块后(即在 conv_2,3,4,5 后),都添加一个 ECA 模块,实现了矿物图像局部跨通道交互融合的目的,最后经全局平均池化等操作,实现了对于矿物

的识别。最终,本文构建的 ECA-ConvNeXt 网络模型,以不同种类矿物的矿石图像数据为研究对象,进行一系列实验后,矿物识别准确率及模型性能与现有主流网络相比都有较大提升。矿物智能识别模型整体结构如图 3 所示。

3 实验过程

3.1 矿物图像数据采集

本文选取常见的 26 类矿物作为研究对象,包括斑铜矿、辰砂、赤铁矿、磁黄铁矿、磁铁矿、毒砂、方解石、方铅矿、橄榄石、铬铁矿、黑钨矿、褐铁矿、黄铁矿、黄铜矿、辉钼矿、辉锑矿、辉铜矿、孔雀石、蓝铜矿、铝土矿、软锰矿、闪锌矿、石英、铁钼矿、雄黄、萤石。

由于矿物种类不同,而且同类矿物产地来源也有所不同。不同的矿物经采集、运输等过程,不可避免地会沾染灰尘等,影响矿物图像的清晰度,以致于影响模型的精度。因此,首先对 26 类矿物的矿石进行水洗、晾晒等操作,将已经处理过的矿物的矿石样本放置于实验平台之上,然后使用佳能单反相机从不同的方向拍摄,拍摄距离约 30 cm,其部分矿物的矿石样本图像如图版 I 所示。

3.2 矿物图像数据增强

在深度学习中,是否有充足的数据样本是保证训练模型效果优劣与泛化能力强弱的基础。但是在实际应用中,往往没有建立统一的大型数据集。为了解决因数据集的不足造成识别准确率低等问题,通常采用数据增强的方式提升模型的效果。针对本文中不同矿物矿石图像的特点,对其进行图像裁剪、旋转,图像亮度调整和随机添加噪声操作。对于这些数据增强方法操作后,每类矿物的矿石图像原始数量和增强后的数量如表 2 所示,按照序列号进行排序,共计 34576 张图片。

3.3 实验方法与环境配置

(1) 实验设计

在对矿物识别模型实验之前,需将矿物图片按 6 : 2 : 2 的比例划分为矿物训练集、矿物验证集与矿物测试集。其中,矿物训练集中的图片用于训练本文构建模型中的未知参数,验证集目的是为了验证模型训练的结果,而其测试集用于检验矿物识别模型的效果。根据本文矿物数据集的大小,设置 batch size 大小为 32,初始学习率为 0.001,迭代 200 次,采用交叉熵函数作为模型的损失函数。

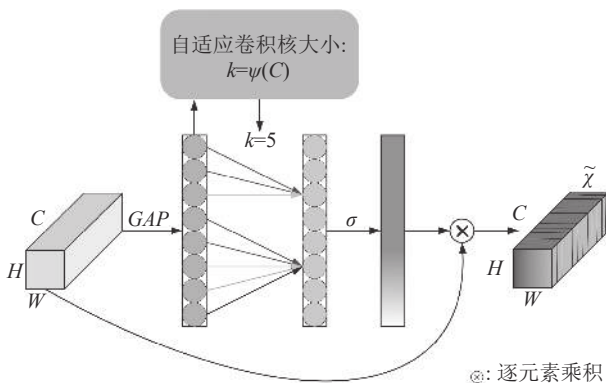


图 2 超强通道注意力(ECA)模块示意图

Fig. 2 Diagram of efficient channel attention (ECA) module

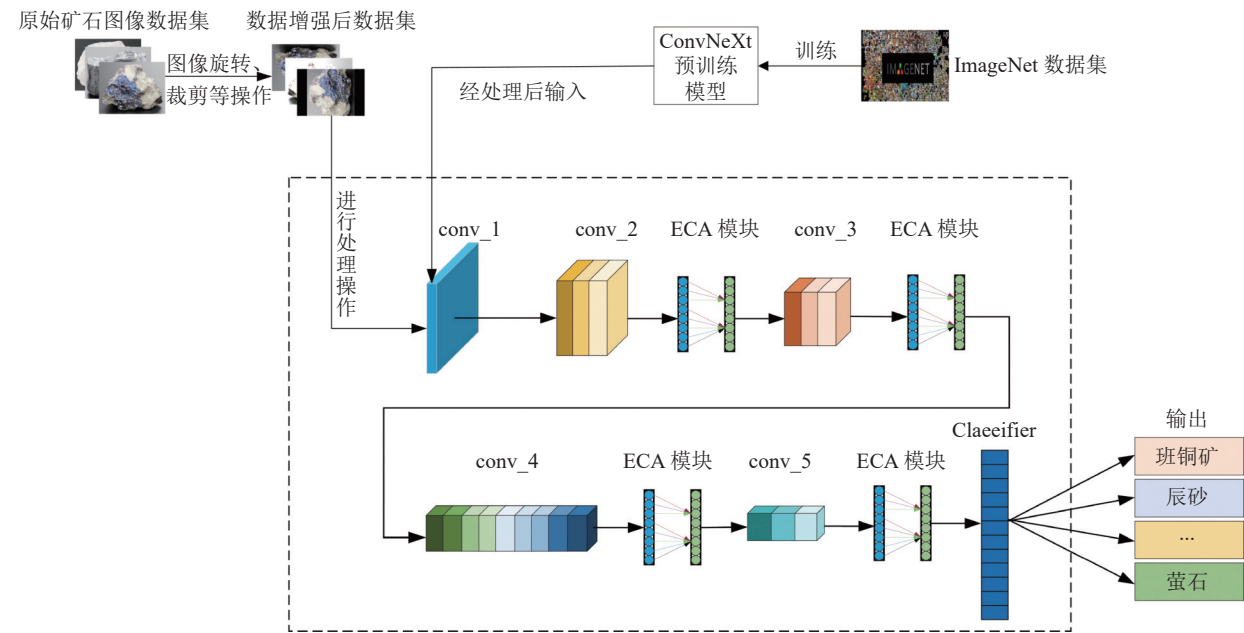
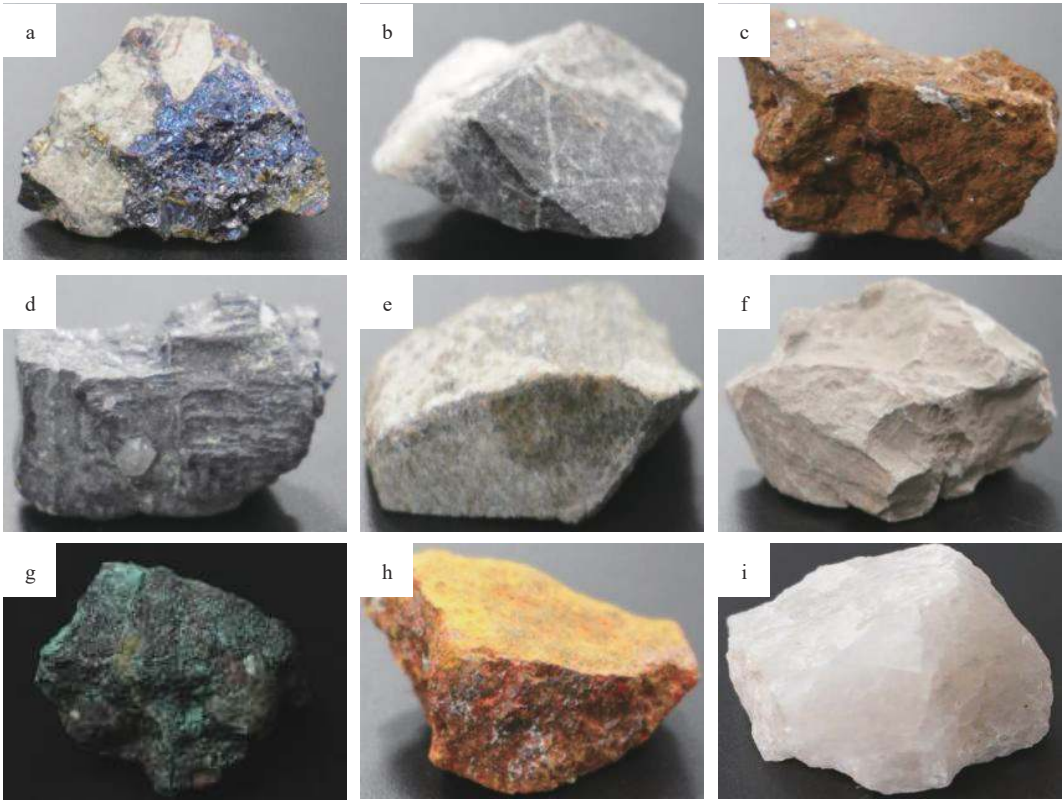


图 3 矿物智能识别模型

Fig. 3 Mineral intelligent identification model

图版 I Plate I



a. 斑铜矿; b. 辰砂; c. 褐铁矿; d. 方铅矿; e. 辉钼矿; f. 铝土矿; g. 孔雀石; h. 雄黄; i. 石英

在矿物图片实验过程中使用 Adam 作为优化器,同时利用余弦退火对矿物图片训练过程中学习率进行不断优化,使其接近全局最优值点,进而得到

矿物识别模型。

(2)环境配置

本实验基于 Pytorch 框架进行矿物识别模型的

表 2 矿物类别及图像数量统计
Table 2 Mineral category and image quantity statistics

序号号	矿物类别	原始图像数量/张	增强后的图像数量/张	序号号	矿物类别	原始图像数量/张	增强后的图像数量/张
1	斑铜矿	194	776	14	黄铜矿	360	1440
2	辰砂	214	856	15	辉钼矿	278	2224
3	赤铁矿	258	1032	16	辉锑矿	472	1888
4	磁黄铁矿	146	584	17	辉铜矿	300	1200
5	磁铁矿	286	1144	18	孔雀石	466	1864
6	毒砂	216	864	19	蓝铜矿	310	1240
7	方解石	374	1496	20	铝土矿	298	1192
8	方铅矿	516	2064	21	软锰矿	324	1296
9	橄榄石	142	568	22	闪锌矿	362	1448
10	铬铁矿	246	984	23	石英	257	1285
11	黑钨矿	324	1296	24	铁钼矿	121	968
12	褐铁矿	440	1760	25	雄黄	352	1408
13	黄铁矿	399	1995	26	萤石	426	1704

训练、验证与测试。其中, 硬件环境: 处理器为英特尔 i9-10900K(64GB), 显卡为 NVIDIA GeForce RTX 3080。软件环境: CUDA11.4, cuDNN8.2.4, Pytorch1.11, PyCharm2021 专业版。

4 结果分析

4.1 训练结果

本文利用训练准确率与验证准确率评估 ECA-ConvNeXt 网络矿物识别模型的训练效果。其中, 训练准确率与验证准确率相差太大表示模型过拟合, 导致网络泛化能力差。

为了验证本文模型的有效性, 在一定条件下(数据集与处理方法相同, 训练 200 次), 分别在训练集与验证集上与 5 种典型图像分类网络进行横向的对

比, 其结果如图 4 所示。图 4-a 为训练集上不同模型准确率的对比, 而图 4-b 为验证集上不同模型的准确率对比。在训练集上, 6 个网络的训练准确率均达到 90% 以上; 在验证集上, VGG19 准确率为 92.22%, GoogLeNet 准确率为 92.43%, ResNet50 准确率为 95.04%, ResNeXt50 准确率为 95.43%, ConvNeXt 准确率为 97.53%, 本文模型准确率为 99.42%。从以上结果可知, 本文构建的 ECA-ConvNeXt 网络模型与其他 5 种网络相比, 取得了较优成绩, 总体上性能最佳。

4.2 实验结果

4.2.1 模型性能

混淆矩阵是评判模型优劣的一种指标, 常用于评判图像分类模型的优劣。本文基于改进的 ConvNeXt

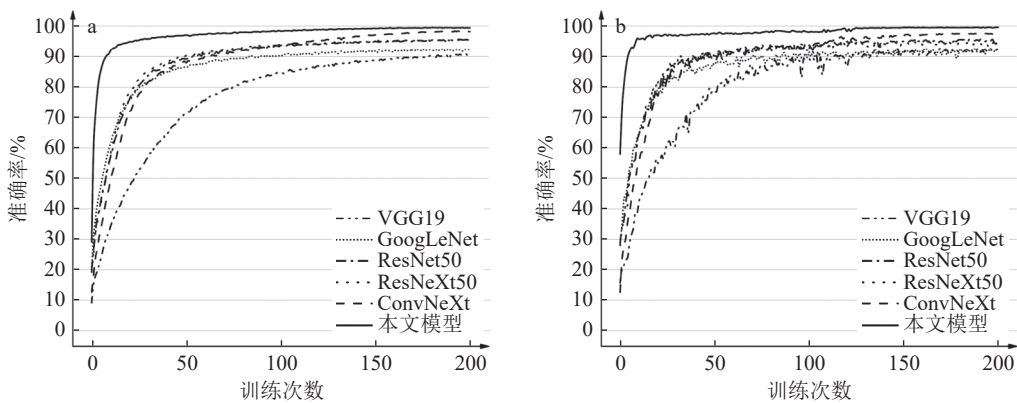


图 4 训练集(a)与验证集(b)不同网络的准确率对比

Fig. 4 Comparison of the accuracy of different networks in training sets (a) and validation sets (b)

矿物智能识别模型对于 26 类矿物的混淆矩阵如图 5 所示。其中,大多数准确率达 97% 及以上,对黄铁矿进行识别分类时只有 94%,识别成毒砂与方解石有 2% 的概率,而识别成黄铜矿与闪锌矿有 1% 的概率。

对于图像分类任务,如何对图像所属类别进行正确分类是其主要的目的。在实际情况下,预测值与真实值的情况如表 3 所示。其中,TP 为真实值为正,预测值为正;FP 为真实值为正,预测值为负;FN 为真实值为负,预测值为正;TN 为真实值为负,预测值为负。

为了验证本文构建的矿物智能识别模型的有效性,选用准确率(Accuracy)、精确率(Precision)与召回率(Recall)指标对 6 种矿物识别模型进行效果评估。计算公式如下所示:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

在相同训练集上进行训练的 6 种矿物识别模型,在同一测试集上进行对应的测试,结果如表 4 所示。从表 4 可以看出,分类准确率与模型的复杂度呈正相关。其中,GoogLeNet 优于 VGG19,而 ResNet50 引入残差结构,ResNeXt50 则是在 ResNet 的基础上结合分组卷积的优点,使其效果得到提升,而 ConvNeXt 是借鉴 Swin-Transformer 的结构,改进卷积神经网络的结构提升其性能。本文在 ConvNeXt 的基础上,引入注意力机制,使其准确率、精确率、召回率分别达到 98.58%、98.62%、98.73%,与其他网络相比均是最佳。因此,对于 ConvNeXt 网络的改进较有效。

4.2.2 消融实验分析

为了验证所提出的优化方法对于原 ConvNeXt 模型性能的影响,在相同实验条件下,采用不同优化方法对 ConvNeXt 模型进行测试,其结果如表 5 所示。在原 ConvNeXt 网络模型的基础上,利用

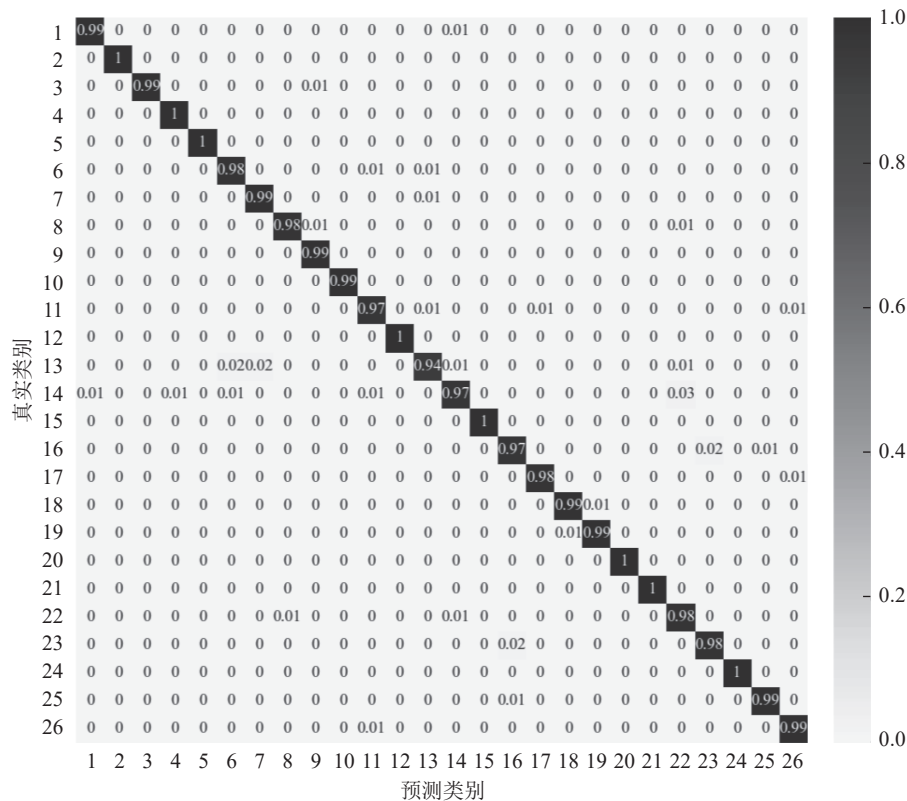


图 5 混淆矩阵

(横坐标与纵坐标的序列号与表 2 一致,分别表示真实标签与预测标签,颜色的深浅代表准确率的大小)

Fig. 5 Confusion matrix

表 3 分类指标
Table 3 Classification index

		预测值	
		正例(Positive)	反例(Negative)
真实值	正例(Positive)	TP	FP
	反例(Negative)	FN	TN

表 4 各个模型的测试评估结果
Table 4 Test evaluation results of each model

模型	准确率	精确率	召回率
VGG19	91.25%	91.62%	91.70%
GoogLeNet	92.66%	93.36%	92.33%
ResNet50	94.88%	95.04%	94.95%
ResNeXt50	95.07%	95.45%	95.03%
ConvNeXt	96.60%	96.62%	96.73%
本文模型	98.58%	98.62%	98.73%

ECA 或迁移学习的方法,使其在准确率、精确率与召回率 3 项指标上都有 1% 左右的性能提升,而本文提出的改进 ConvNeXt 模型,只增加了极少的参数量与模型大小,约有 2% 的性能提升,取得了不错的效果。因此,本文提出的改进 ConvNeXt 模型更适合于对矿物的智能识别。

4.3 矿物图像特征图可视化

卷积神经网络主要是通过提取图像中物体或其他本身所具有的特征,达到识别、检测、分割等目的,

表 5 不同优化方法对于 ConvNeXt 模型的影响
Table 5 The impact of different optimization methods on the ConvNeXt model

模型	参数量	模型大小	准确率	精确率	召回率
ConvNeXt	27.80M	106.25M	96.60%	96.62%	96.73%
ECA	27.82M	106.26M	97.62%	97.65%	97.61%
+ConvNeXt					
迁移学习+ConvNeXt	27.82M	106.26M	97.87%	97.90%	98.02%
本文模型	27.84M	106.27M	98.58%	98.62%	98.73%

但对于其所提取的特征是什么,哪些区域对于识别起作用,以及由此得出的最终结果都无从得知。为了确定本文提出的矿物识别模型对于矿石图像中哪些区域识别的影响较大,本文利用 Grad-CAM(Selvaraju et al., 2017)获得可视化矿石图像的特征图。主要是将本文提出的网络模型的最后一层 feature maps(特征图),利用 Grad-CAM 进行可视化输出,可以观察矿物图像中的有效区域对于模型识别的影响。

图 6 是在不同模型的条件下,所提取特征的可视化热力图。其中,图 6 中从上到下的 5 种矿物分别是:斑铜矿、赤铁矿、橄榄石、辉钼矿和石英。而图 6 中从左到右依次是原图与依据 VGG19、GoogLeNet、ResNet50、ResNeXt50、ConvNeXt 和本文所提出的模型生成与之对应的特征图。其中,通过对于原图与不同模型之间特征图的对比(图中颜

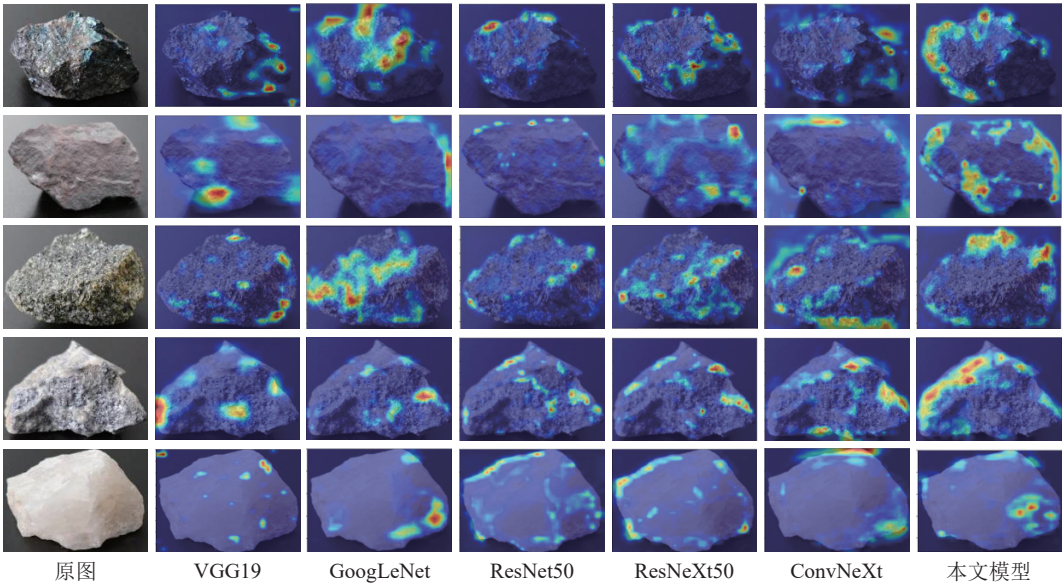


图 6 原图与可视化特征热力图

Fig. 6 Original image and visualized feature heat map

色越深,代表模型对于所在区域关注度越高,对识别的结果影响越大),本文提出的矿物识别模型优于传统模型,同时在对矿物识别的过程中,主要以矿物的边缘和颜色为感兴趣区域,也从侧面说明本文模型对于矿物特征的提取效果好,实现了对矿物的准确识别,证明了模型的有效性。因此,本文提出的基于改进的 ConvNeXt 网络矿物识别模型,提高了对于矿物的特征提取能力,较好地解决了矿物识别效率低的问题。

5 结 论

在原有 ConvNext 网络的基础上,利用 26 类矿物 34576 张图像,通过迁移学习与引入 ECA 模块,构建了基于改进的 ConvNeXt 网络的矿物智能识别模型,实现了对于矿物分类的高效准确识别。得到了以下结论。

(1)通过在 ImageNet 上已预训练的 ConvNeXt 模型,经一定处理后利用迁移学习的方法,加载到矿物识别模型框架中,同时针对不同矿物图像自身特点,通过引入超强注意力通道(ECA)机制,构建了 ECA-ConvNeXt 网络矿物识别模型,提高了矿物图像特征融合的能力与训练速度。

(2)采用 ECA-ConvNeXt 矿物图像识别模型,将 26 类矿物图像与 VGG19、GoogLeNet、ResNet50、ResNeXt50 和 ConvNeXt 网络相比,其在训练过程收敛更快,同时在准确率、精确度与召回率 3 项数据上均是最佳,分别达到了 98.58%、98.62% 与 98.73%;而消融实验,进一步证明了模型的有效性。

(3)利用 Grad-CAM 方法,在矿物原图的基础上,通过与不同模型之间矿物图像特征图的可视化结果对比,进一步证明了本文模型的优越性与在矿物识别领域的显著性。

本文研究成果对于矿物智能识别领域具有一定的借鉴意义,如何进一步提高矿物智能识别的速度与准确率,以及如何将其应用到实际场景等仍然是重要的研究方向。

参考文献

- Baykan N A, Yilmaz N. 2010. Mineral identification using color spaces and artificial neural networks[J]. *Computers & Geosciences*, 36(1): 91–97.
- El Haddad J, de Lima Filho E S, Vanier F, et al. 2019. Multiphase mineral identification and quantification by laser-induced breakdown spectroscopy[J]. *Minerals Engineering*, 134: 281–290.
- Hu J, Shen L, Sun G. 2018. Squeeze-and-excitation networks[C]// *Proceedings of the IEEE conference on computer vision and pattern recognition*: 7132–7141.
- Izadi H, Sadri J, Mehran N A. 2013. Intelligent mineral identification using clustering and artificial neural networks techniques[C]// *First Iranian Conference on Pattern Recognition and Image Analysis (PRIA)*. IEEE, 2013: 1–5.
- Khajehzadeh N, Haavisto O, Koresaar L. 2016. On-stream and quantitative mineral identification of tailing slurries using LIBS technique[J]. *Minerals Engineering*, 98: 101–109.
- Li X, Wang W, Hu X, et al. 2019. Selective kernel networks[C]// *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*: 510–519.
- Liu C, Li M, Zhang Y, et al. 2019. An enhanced rock mineral recognition method integrating a deep learning model and clustering algorithm[J]. *Minerals*, 9(9): 516.
- Liu Y, Zhang Z, Liu X, et al. 2021. Ore image classification based on small deep learning model: Evaluation and optimization of model depth, model structure and data size[J]. *Minerals Engineering*, 172: 107020.
- Liu Z, Mao H, Wu C Y, et al. 2022. A convnet for the 2020s[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 11976–11986.
- Maitre J, Bouchard K, Bédard L P. 2019. Mineral grains recognition using computer vision and machine learning[J]. *Computers & Geosciences*, 130: 84–93.
- Selvaraju R R, Cogswell M, Das A, et al. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]// *Proceedings of the IEEE international conference on computer vision*: 618–626.
- Tsuji T, Yamaguchi H, Ishii T, et al. 2010. Mineral classification from quantitative X-ray maps using neural network: Application to volcanic rocks[J]. *Island Arc*, 19(1): 105–119.
- Trejbal J, Valentová T, Neerka V, et al. 2020. Mechanical and image analysis of adhesion between mineral aggregate and bituminous binder[J]. *Acta Polytechnica CTU Proceedings*, 26: 112–116.
- Wang Q, Wu B, Zhu P, et al. 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 11534–11542.
- Woo S, Park J, Lee J Y, et al. 2018. Cbam: Convolutional block attention module[C]// *Proceedings of the European Conference on Computer Vision (ECCV)*: 3–19.
- Zamir A R, Sax A, Shen W, et al. 2018. Taskonomy: Disentangling task transfer learning[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 3712–3722.
- Zhou W, Wang H, Wan Z. 2022. Ore image classification based on improved CNN[J]. *Computers & Electrical Engineering*, 99: 107819.
- 郭艳军, 周哲, 林贺洵, 等. 2020. 基于深度学习的智能矿物识别方法研

- 究[J]. 地学前缘, 27(5): 39–47.
- 郝慧珍, 顾庆, 胡修棉. 2021. 基于机器学习的矿物智能识别方法研究进展与展望[J]. 地球科学, 46(9): 16.
- 李明超, 刘承照, 张野. 2020. 耦合颜色和纹理特征的矿物图像数据深度学习模型与智能识别方法[J]. 大地构造与成矿学, 44(2): 203–11.
- 刘艳鹏, 朱立新, 周永章. 2020. 大数据挖掘与智能预测找矿靶区实验研究——卷积神经网络模型的应用[J]. 大地构造与成矿学, 44(2): 192–202.
- 彭伟航, 白林, 商世为. 2019. 基于改进 InceptionV3 模型的常见矿物智能识别[J]. 地质通报, 38(12): 2059–66.
- 王李管, 陈斯佳, 贾明滔, 等. 2020. 基于深度学习的黑钨矿图像识别选矿方法[J]. 中国有色金属学报, 30(5): 1192–1201.
- 许振浩, 马文, 林鹏, 等. 2021. 基于岩石图像迁移学习的岩性智能识别[J]. 应用基础与工程科学学报, 29(5): 1075–1092.
- 徐述腾, 周永章. 2018. 基于深度学习的镜下矿石矿物的智能识别实验研究[J]. 岩石学报, 34(11): 3244–3252.
- 赵明. 2010. 矿物学导论[M]. 北京: 地质出版社.
- 张旭, 于明鑫, 祝连庆, 等. 2020. 基于全光衍射深度神经网络的矿物拉曼光谱识别方法[J]. 红外与激光工程, 49(10): 168–175.
- 周永章, 张良均, 张傲多, 等. 2018. 地球科学大数据挖掘与机器学习[M]. 广州: 中山大学出版社: 1–269.
- 周永章, 左仁广, 刘刚, 等. 2021. 数学地球科学跨越发展的十年: 大数据、人工智能算法正在改变地质学[J]. 矿物岩石地球化学通报, 40(3): 556–573, 777.