

doi: 10.12097/gbc.2022.10.002

基于 MobileViT 的岩石薄片图像岩性识别方法研究

王琼¹, 杨杰^{1,2,3,4}, 霍凤财^{1,2,3,4*}, 董宏丽^{1,2,3,4}, 任伟建^{1,2,3,4}, 于涛⁵

WANG Qiong¹, YANG Jie^{1,2,3,4}, HUO Fengcai^{1,2,3,4*}, DONG Hongli^{1,2,3,4}, REN Weijian^{1,2,3,4}, YU Tao⁵

1. 东北石油大学电气信息工程学院, 黑龙江 大庆 163318;
2. 东北石油大学人工智能能源研究院, 黑龙江 大庆 163318;
3. 黑龙江省网络化与智能控制重点实验室, 黑龙江 大庆 163318;
4. 东北石油大学三亚海洋油气研究院, 海南 三亚 572000;
5. 大庆油田有限责任公司第四采油厂工艺研究所, 黑龙江 大庆 163511

1. School of Electrical Information Engineering, Northeast Petroleum University, Daqing 163318, Heilongjiang, China;
2. Artificial Intelligence Energy Research Institute of NEPU, Daqing 163318, Heilongjiang, China;
3. Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Daqing 163318, Heilongjiang, China;
4. Sanya Offshore Oil & Gas Research Institute, Sanya 572000, Hainan, China;
5. Technology Research Institute of the Fourth Oil Production Plant of Daqing Oilfield Company Limited, Daqing 163511, Heilongjiang, China

摘要: 岩石薄片图像中包含了大量肉眼无法观察到的地质特征信息, 对岩石薄片图像的岩性识别结果为后续的石油勘探和开发奠定了基础。针对岩性识别数据集不均衡、识别模型参数多等问题, 提出一种改进的轻量化 MobileViT 模型, 该模型针对涵盖了 90% 以上常见岩性的岩石薄片图像进行建模分析。首先, 为使模型更好地学习到每类岩石薄片图像中所包含的独特特征, 对数据集进行数字增加。其次, 使用 GELU 替换 MobileViT 中 MV2 模块中常规 ReLU6, 从而作为该模块的激活函数, 有效解决神经元死亡的问题, 提升模型的收敛速度。最后, 划分训练集和测试集, 使用余弦退火算法自动更新学习率, 以迁移学习加速训练过程, 实现岩石薄片图像中针对岩性的自动识别。实验结果表明, 改进后的 MobileViT 对岩性识别的准确率达 82.9%, 模型的参数仅为 7.66M, 通过实例验证该算法具有较好的鲁棒性。

关键词: 岩石薄片; 岩性识别; MobileViT; 余弦退火; 轻量化

中图分类号: P588 **文献标志码:** A **文章编号:** 1671-2552(2024)06-0938-09

Wang Q, Yang J, Huo F C, Dong H L, Ren W J, Yu T. Lithology identification method of rock thin section images based on MobileViT. *Geological Bulletin of China*, 2024, 43(6): 938-946

Abstract: The rock thin-section images contain a large amount of geological feature information that cannot be observed with the naked eye. The lithology identification of rock thin-section images lays the foundation for subsequent oil exploration and production. Aiming at the problems of unbalanced lithology identification data set and many identification model parameters, an improved lightweight MobileViT model is proposed to model and analyze the rock slice images covering more than 90% of common lithology. First, to enable the model to better learn the unique features contained in each type of rock slice image, adding numbers of the dataset is performed on the image. Secondly, use GELU to replace the ReLU6 of the MV2 module in MobileViT as the activation function of the module, which effectively solves the problem of neuron death and improves the convergence speed of the model. Finally, the training set and the test set are divided, the cosine annealing algorithm is used to automatically update the learning rate, and the transfer

收稿日期: 2022-10-05; 修订日期: 2022-12-08

资助项目: 黑龙江省自然科学基金项目《基于分布式算法和多源异构井筒数据驱动的页岩储层有利区评价研究》(编号: LH2023F007)

作者简介: 王琼(1969-), 女, 教授, 从事图像处理和智慧油田等方面的研究。E-mail: 765575098@qq.com

* 通信作者: 霍凤财(1976-), 男, 副教授, 从事智能控制和图像处理等方面的研究。E-mail: huofc@126.com

learning is used to speed up the training process, so as to realize the automatic identification of rock slice images. The experimental results show that the accuracy of the improved MobileViT for lithology identification is 82.8%, and the model parameters are only 7.66M, which has good robustness.

Key words: rock thin section; MobileViT; lithology identification; cosine annealing; light weight

岩石是地球物质的主要组成成分,根据成因分为火成岩、沉积岩和变质岩三大类(赖文等, 2020)。岩性识别是石油勘探与开发过程中不可或缺的一环。根据岩石特有的颜色、成分、结构等岩性特点,可将储集岩划分为不同的单元,依据相似的岩性,同一单元的储集岩有着类似的地质和成岩条件(付光明等, 2017)。这些条件与孔隙度和含油量密切相关,同时也是石油储层特征研究、储量计算和地质建模的前提。薄片鉴定是岩性识别的重要方法之一,在传统岩石薄片的岩性识别与分类中,主要以人眼观察、手工操作和经验分类为主,受人为主观因素的影响大,需要一定的先验知识,专业程度较高,难以普及(宫清顺, 2012)。

近年来,大数据与人工智能算法正改变着地质学,作为人工智能算法的核心,机器学习在矿床地质、土壤分析、岩性识别等领域取得了突破性的进展(徐述腾等, 2018; 周永章等, 2018, 2021)。因此,将岩石薄片图像与机器学习相结合可以克服上述难题。多层感知机是机器学习中的分支之一。早年, Marmo et al. (2005)对来自不同海洋环境的 1000 多张碳酸盐岩薄片图像使用多层感知机进行建模分析,该模型使用岩石薄片的灰度图作为输入,得到 23 个特征值后,用这些特征值进行分类,最终取得了 93.3% 和 93.5% 的准确率; Singh et al. (2010)利用多层感知器神经网络对玄武岩岩石薄片的图像进行了分类识别,在测试集上得到 90% 的识别准确率。但多层感知机存在一定的局限性,随着网络层数的增加,模型变得难以训练,且还伴随着能梯度爆炸、梯度消失等问题,并不适合大型的图像分类任务(He et al., 2016)。随着 AlexNet、VGG(Simonyan 等, 2014)、ResNet(马泽栋等, 2022)等深度卷积神经网络的兴起,机器学习的另一子集深度学习逐渐在计算机视觉和自然语音处理中得到广泛应用(盖荣丽等, 2021)。与传统机器学习算法,如 SVM、随机森林等模型需要对图片进行额外的预处理和特征提取不同,深度学习中的深度卷积神经网络通过卷积核直接提取图片特征;通过卷积运算替换传统的矩阵

乘法运算,极大地减少了运算参数;通过残差连接和反向传播在保证网络层数增加的同时,学习任务的精度也随之增加(张珂等, 2021)。近年,学者专家逐渐将深度卷积神经网络引入到岩性识别中。张野等(2018)使用 Inception V3 网络结合迁移学习,将在 ImageNet 上预训练的权重迁移到花岗岩、千枚岩及角砾岩图片的岩性识别上,测试集的准确率达到 85% 以上;熊越晗等(2021)使用 Inception V3 模型与迁移学习相结合,对砂岩、泥岩、泥质砂岩和砂质泥岩进行建模分析,经 1000 轮迭代后,在训练集、验证集上分别取得了 92.77% 和 76.31% 的准确率;许振浩等(2021)使用 DSOD 算法对图像进行岩石检测,从而建立岩石图像数据集,然后使用 ResNet101 结合迁移学习,对花岗岩、流纹岩等常见的岩石图像进行特征提取,岩性识别的正确率达 90.21%;徐述腾等(2018)设计有针对性的 Unet 网络(陈国军等, 2023)实现对矿相显微镜矿石矿物的分类识别,测试集中的矿物识别准确率达到 90%。综上,将机器学习与深度学习引入岩性识别,可以自动识别岩性,降低主观因素的影响,但仍存在数据集不全面和模型参数多的缺点。本文旨在设计一种通用的轻量级模型,实现岩石薄片图像的全自动识别,采用改进的、参数量仅为 7.66M 的 MobileViT 模型对岩石薄片图像进行特征提取和识别(Mehta et al., 2021)。为提高模型的训练速度和准确率,引入迁移学习,将在 ImageNet 数据集中训练好的权重迁移到岩石薄片图像的识别中。经验证,在训练集上的准确率达 100%,测试集的准确率达 82.9%。

1 岩石薄片图像岩性自动识别方法

卷积神经网络(刘强等, 2022)使用固定尺寸大小的卷积核对图片进行特征提取,使得模型对图片的特征提取是局部的,只能通过大量的网络堆叠进一步获取图片的特征信息,随着网络层数的叠加,极有可能导致梯度消失、梯度爆炸等问题(Dosovitskiy et al., 2020)。Vision Transformer(ViT)中的 Attention 机制可以对图片进行全局特征提取

(Dosovitskiy et al., 2020),但又由于缺少权值共享和稀疏连接的特性,使模型需要大量的参数进行迭代,导致模型的训练时间长,收敛速度慢。因此,本文用卷积与 ViT 相结合的轻量级模型 MobileViT 对岩石薄片图像进行特征提取,在兼顾准确率的基础上,减少参数。整体技术路线如图 1 所示。

首先,为了让模型学习到更多各类岩石所包含的特征,对岩石薄片图像进行随机旋转、翻转、缩放等操作,达到扩充数据集、增加样本量的目的(周志华, 2016),并将图片进行归一化处理。其次,将在 ImageNet 数据集上预训练好的权重迁移到岩石薄片图像的岩性识别中,经 MobileViT 提取特征。最后,经过 Softmax 分类器得到各岩石薄片图像各自对应的岩性。

1.1 MobileViT 模型

MobileViT 模型流程如图 2 所示,其核心模块

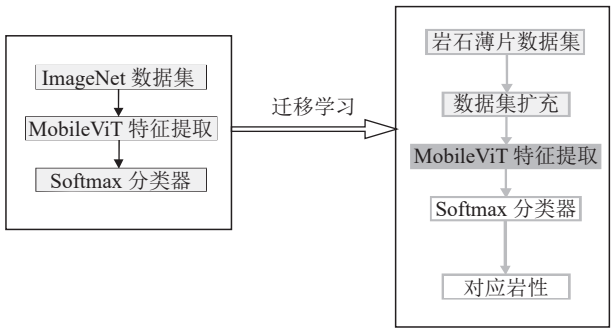


图 1 技术路线

Fig. 1 Technical route

为 MobileNetV2(MV2) 模块(Sandler M 等, 2018)和 MobileViT(MViT)模块。

当卷积步长(Stride)为 2 时, MV2 模块如图 2-a 所示,包含 2 个大小为 1×1 的卷积核,用以改变输入特征的通道,一个大小为 3×3 的卷积核,对输入特征进行深度可分离卷积;当步长为 1 时,为防止输出特征过大导致梯度爆炸、梯度消失等问题,引入残差连接,使网络参数可以跨层传播。首先将输入数据下采样,降低数据维度;然后使用卷积操作提取特征;最后进行上采样,使数据维持输入时的维度操作。MV2 模块使用逆操作,首先对输入数据进行上采样;然后使用深度可分离卷积进行特征提取,最后下采样,使数据维持输入时的维度,极大地减少了运算量及参数量。

MViT 模块如图 2-b 所示。由基于卷积的局部特征提取模块、基于 Transformer 的全局特征提取模块和特征融合模块三部分组成, L 代表 Transformer 的个数。MViT 提取图像特征分为以下 3 步: ①使用卷积操作对输入数据进行下采样;②将下采样后的数据进行分块,并将每个小块展平;③从左至右取对应位置的像素点进一步经 Transformer 提取特征。这样在减少计算量的同时兼顾了全局特征提取。

1.2 MV2 模块的改进

在深度学习模型中,为了使模型具备更好的学习能力,必须引入激活函数,使模型具有非线性拟合的能力(Hinton G, 2010)。

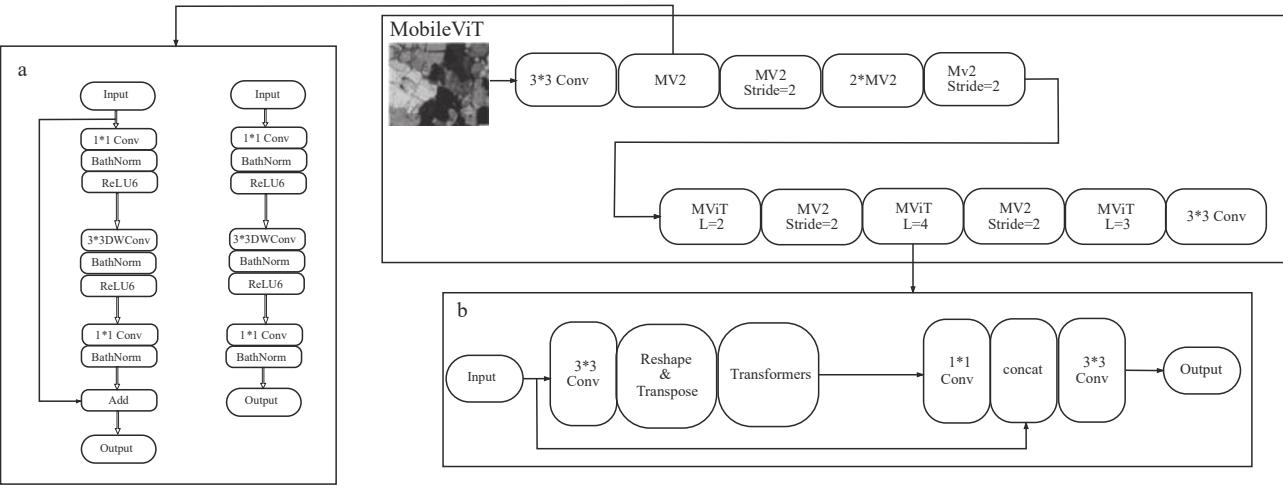


图 2 MobileViT 特征提取流程

Fig. 2 MobileViT feature extraction process

a—MV2 模块; b—MViT 模块

常规 MV2 模块采用 ReLU6 作为激活函数, 其形式如公式 (1) 所示, 函数图像如图 3 所示。

$$\text{ReLU6}(x) = \min(6, \max(0, x)) \quad (1)$$

当输入数据小于 0 或大于 6 时, 输出为 0。ReLU6 通过门控机制对输入数据进行激活, 对输入数据的符号和数值大小高度敏感, 当且仅当输入数据处于 (0,6) 范围时, 神经元才会被激活, 缺乏随机因素。高斯误差线性单位 (Gaussian Error Linear Units, GELU) 的表达式如公式 (2) 所示, 其图像如图 4 所示。

$$\text{GELU}(x) = xP(X \leq x) = x\Phi(x) \quad (2)$$

其中, X 服从标准正态分布, $P(X \leq x)$ 是 X 小于或等于给定值 x 的概率, $\Phi(x)$ 为正态分布的概率函数, 表达式如公式 (3) 所示, 其中, μ 和 σ 为均值和方差, 都是可优化的参数。

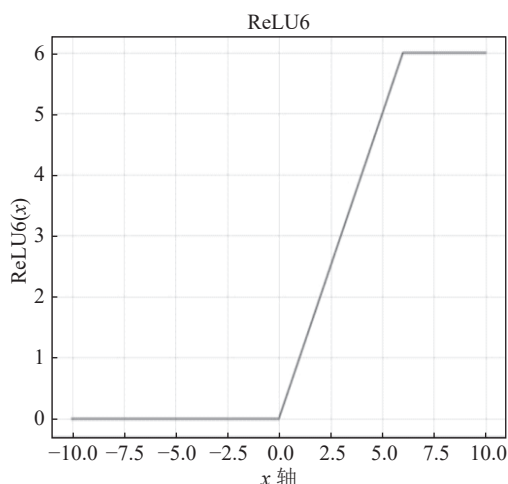


图 3 ReLU6 函数图像

Fig. 3 Image of ReLU6 function

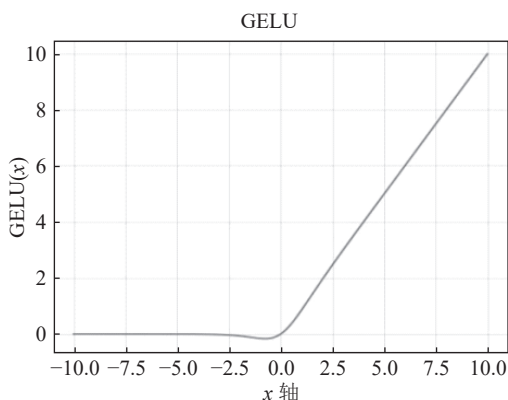


图 4 GELU 激活函数图像

Fig. 4 Image of GELU function

$$xP(X \leq x) = x \int_{-\infty}^x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dX \quad (3)$$

最终可近似为公式 (4)。

$$0.5x(1 + \tanh[\sqrt{2/\pi}(x + 0.044715x^3)]) \quad (4)$$

即为输入数据 x 添加了一个 (0,1) 范围内的乘数。由于神经网络常使用批量归一化 (Batch Normalization) 处理输入数据, 使得神经元的输入 x 近似遵循正态分布。所以, 随着 x 减小, $\Phi(x)$ 的值也随之减小, 被丢弃的可能性则随之增加; 反之, 随着 x 增加, $\Phi(x)$ 的值随之增加, 被保留的可能性也随之增加。相当于在激活函数中添加了随机正则。

与 ReLU6 不同, GELU 随机依赖输入数据实现对神经元的激活, 对输入数据的符号和数值大小不再敏感, 不易造成神经元死亡, 能有效避免梯度消失的问题 (Hendrycks et al., 2016)。

1.3 余弦退火算法

由于鞍点的存在, 深度神经网络难以训练。如图 5 所示, 灰点所在位置对于 x 轴来说是最小值, 对于 z 轴来说是最大值。模型进行梯度下降时, 该点的梯度为 0, 模型认为此时已经收敛到最优, 不再进行反向传播更新参数, 因此陷入局部最优。

为了防止模型陷入局部最优, 可以加大模型的学习率, 从而加大梯度下降的步长, 使模型跳出鞍点 (李嘉璇, 2017)。但随着学习率的加大, 神经网络在进行梯度下降时会在驻点附近震荡, 无法收敛至最优。使用余弦退火 (Cosine annealing) 算法 (Loshchilov 等, 2016) 对学习率进行调整可以很好地

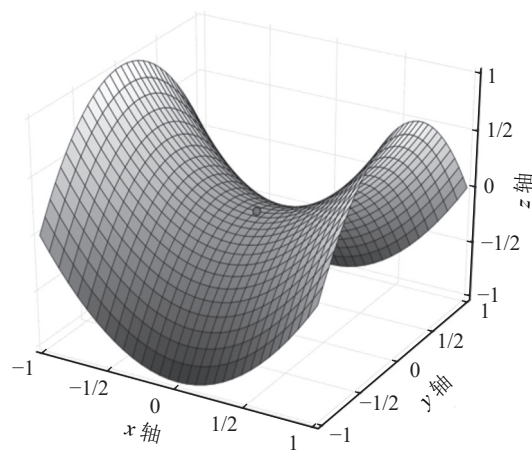


图 5 马鞍面及鞍点

Fig. 5 Saddle face and saddle point

避免上述问题。

余弦退火算法的原理如公式 (5) 所示。

$$\eta_i = \eta_{min}^i + \frac{1}{2}(\eta_{max}^i - \eta_{min}^i)(1 + \cos(\frac{T_{cur}}{T_i}\pi)) \quad (5)$$

式中, η_i 表示当前学习率; i 代表重启的次数, 通过提高学习率实现; η_{min}^i 与 η_{max}^i 是学习率的最小值与最大值, 用来定义学习率的范围; T_{cur} 表示重启后经过的训练轮次 (Epoch); T_i 表示当前总共训练轮次。

以 i 等于 3, 训练轮次等于 100 为例。设置学习率的范围为 (0,1), 学习率随训练总轮次的变化率如图 6 所示。在开始训练时, 学习率首先缓慢降低, 经过一段时间后加速降低, 最后再次缓慢降低。当降低至最小值时, 学习率更新为最大值, 近似呈周期变化, 在提升模型训练速度的同时, 能够有效避免模型陷入局部最优。

1.4 迁移学习

对于深度学习, 从零开始训练一个模型往往要耗费大量的人力物力。ImageNet 数据库中包含 2 万多个图像类别, 1400 多万张图像, 是世界上种类最全、数量最多的图像数据库。迁移学习可以利用模型间的相关性, 将在 ImageNet 上预训练好的权重迁移到小样本的深度学习任务中, 小样本的学习过程中不必重新提取特征, 仅需对主干网络的全连接层和 Softmax 层进行适当调整, 在提升模型训练速度的同时能够有效地提升模型的准确率(Pan et al., 2010)。

2 实验结果与分析

2.1 实验数据集

本文采用的数据集来自南京大学岩石教学薄片

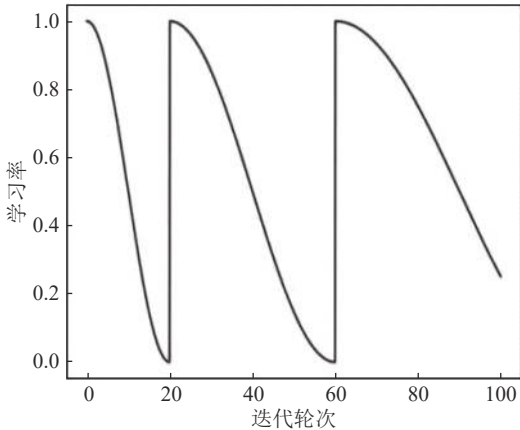


图 6 余弦退火算法学习率
Fig. 6 Learning rate for Cosine Annealing Algorithm

显微图像, 岩石薄片图像的种类和数量如表 1 所示。该数据集非常全面, 包含了 5 种火山碎屑岩、6 种砂岩、6 种泥页岩和粉砂岩等共计 28 种 699 张沉积岩石薄片图像; 7 种基性岩、7 种中性岩、11 种酸性岩等共计 40 种 963 张火成岩岩石薄片图像; 9 种片岩、6 种片麻岩、3 种大理岩等共计 40 种 972 张变质岩岩石薄片图像, 总计 2634 张图像。为了使模型更好地学习到各类岩石薄片图像中所包含的独有的特征, 随机对图像进行翻转、旋转和缩放, 将数据集扩大至原来的 2 倍(图 7)。

2.2 模型的训练与测试分析

实验分别采用 ResNet50、ResNeXt50、VGG、MobileViT 和改进后的 MobileViT 对岩石薄片图像进行建模分析。实验中, 以 8 : 2 的比例划分训练集和测试集, 训练轮次 (Epoch) 设置为 50 轮, Batch Size 的大小为 4, 优化器为 Momentum, 改进后的 MobileViT 采用余弦退火算法自动更新学习率, 学习率的范围为 [0.00001, 0.02], 重启次数 i 等于 3, ResNet50、ResNeXt50、VGG 和 MobileViT 的学习率为固定值 0.00001。训练集的准确率和训练时长如表 2 所示, 训练过程的准确率 (Accuracy) 如图 8 所示。

对于准确率, 使用了迁移学习的 MobileViT 和改进后的 MobileViT 都能达到 100%, 而未使用迁移学习的 VGG、ResNet50 和 ResNeXt50 均未达到 100%, 且在训练的全过程 MobileViT 和本文算法的准确率都明显优于其他 3 种算法。这是 MobileViT 自身的 Attention 机制能够捕捉图片的全局信息, 使图像的感受野不再受限于卷积核的大小, 随着感受野的增加, 正确率也随之提升, 在此基础上由于改进

表 1 岩石薄片图像种类及数量分布
Table 1 Types and quantity distribution of rock slice images

岩石类型	岩性	种类	数量
沉积岩	火山碎屑岩、砂岩、泥页岩和粉砂岩、灰岩、白云岩、硅质岩、蒸发岩、其他内源沉积岩	28	699
	超基性(超镁铁质)岩、基性岩、中性岩、酸性岩、碱性岩及相关岩石	40	963
火成岩	糜棱岩、角岩、矽卡岩、大理岩、蛇纹岩、云英岩、板岩、千枚岩、片岩、片麻岩、变粒岩、斜长角闪岩、麻粒岩、榴辉岩、混合岩、碎裂岩、含变余构造的变质岩	40	972

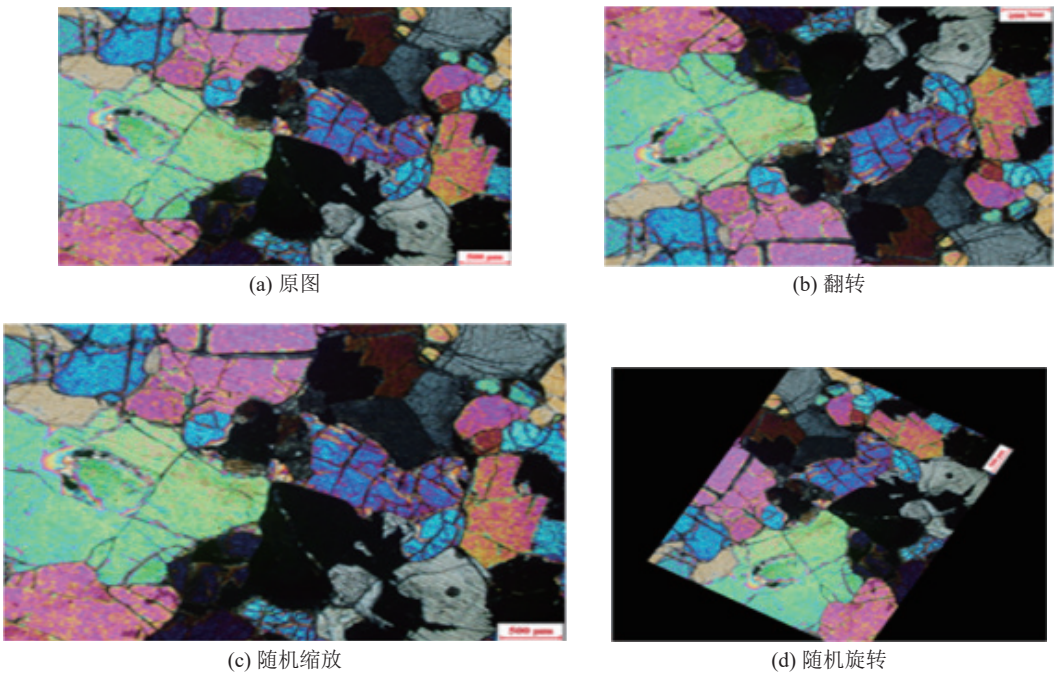


图 7 图像扩充

Fig. 7 Image expansion

了激活函数,能够有效避免神经元的死亡,改进后的 MobileViT 在第 15 轮的训练中逐渐收敛至最优,而 MobileViT 在第 22 轮后的训练中才逐步收敛至最优。对于训练时长,由于 VGG、ResNet50、ResNeXt50 未使用迁移学习,需要重新开始训练,因此训练时间较长,使用了迁移学习的 MobileViT 训练时间为 505.4 min,较未使用迁移学习的几种算法训练时长几乎缩短了 1/2,而改进后的 MobileViT 较 MobileViT 训练时长又缩短了 83.1 min,可见余弦退火算法能够提升模型的训练速度。

在测试集上,三者的正确率和参数量如表 3 所示。对于准确率,由于未采用迁移学习且采取固定的学习率,模型容易陷入局部最优,ResNet50、

ResNeXt50 和 VGG 在训练集的准确率虽然能达到 90.2%、88.1% 和 62.3%,但在测试集的准确率只有 59.8%、56.9% 和 40.1%,而采用了迁移学习的

表 2 训练集的准确率与训练时长对比

Table 2 Comparison of training set accuracy and training time

模型	最高准确率	是否迁移学习	训练时长/min
ResNet50	90.2%	否	905.2
ResNeXt50	88.1%	否	988.9
VGG	62.3%	否	1034.0
MobileViT	100%	是	505.4
Improved_MobileViT	100%	是	422.3

表 3 测试集的准确率与模型参数对比

Table 3 The accuracy of the test set compared with the model parameters

模型	准确率	参数量
ResNet50	59.8%	97.49M
ResNeXt50	56.9%	95.48M
VGG	40.1%	527.79M
MobileViT	80.2%	7.66M
Improved_MobileViT	82.9%	7.66M

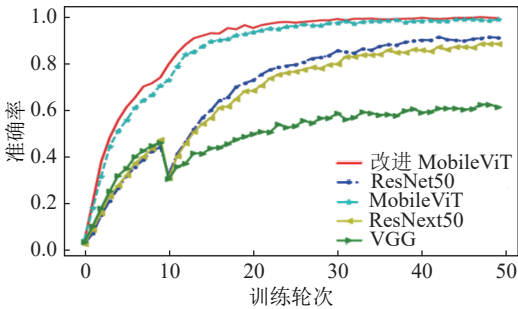


图 8 训练过程准确率

Fig. 8 Accuracy plots for the training process

MobileViT 和改进 MobileViT 在测试集上的准确率为 80.2% 和 82.9%。

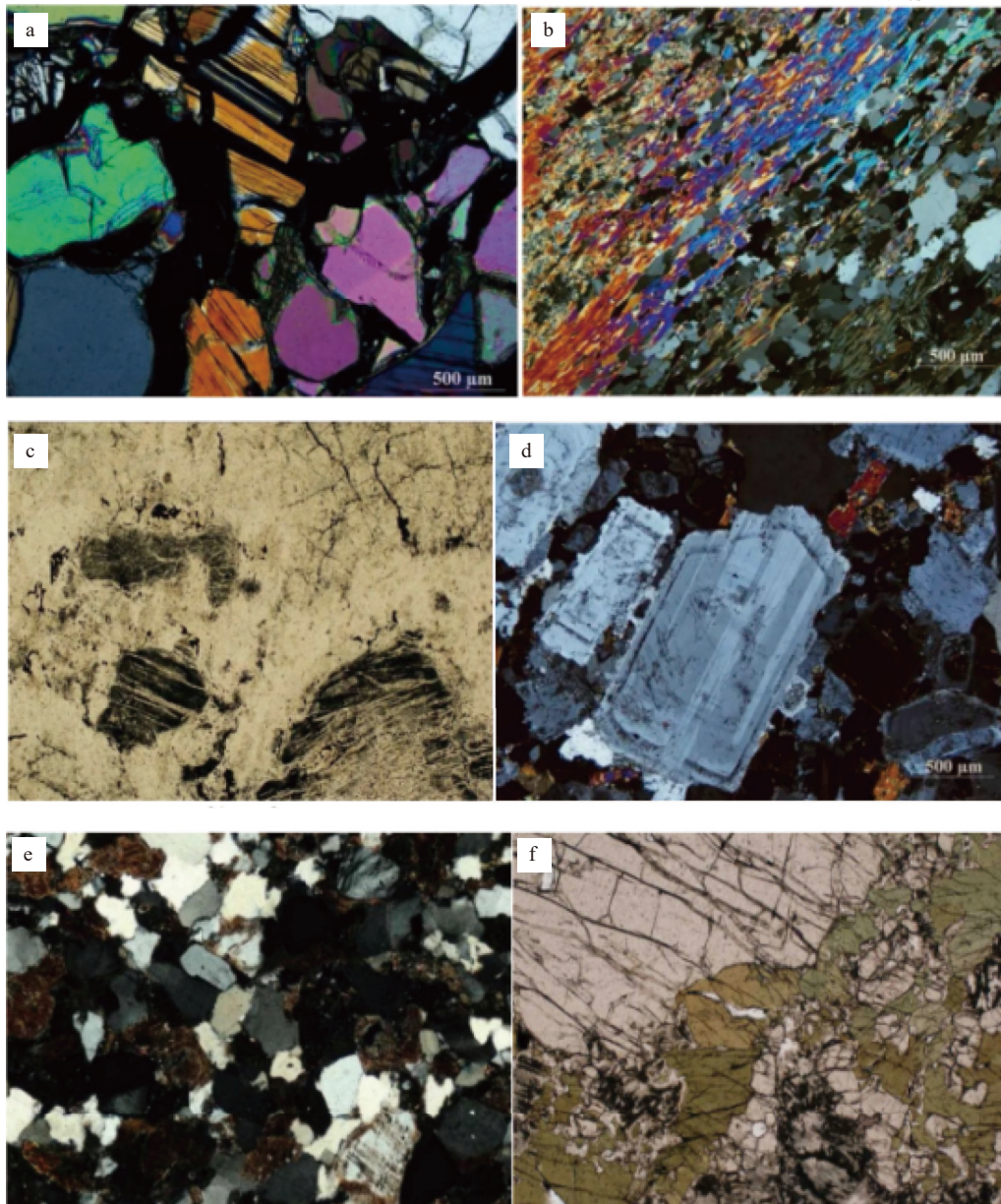
由于使用 GELU 作为激活函数,大量神经元得以存活,改进后的 MobileViT 对岩性识别的准确率较改进前提升了 2.7%。对于模型的参数量,MobileViT 和改进后的 MobileViT 参数量仅为 7.66M。因此,综合对比,本文算法性能更好,更能满足岩石薄片图像岩性识别的综合需求。

2.3 模型的鲁棒性验证

为验证模型的鲁棒性,选取源自网络的高清岩

石薄片图像:1 张橄榄岩、4 张辉绿岩、2 张蛇纹岩、2 张斜长岩、1 张长石砂岩、1 张麻粒岩、2 张长石石英砂岩、3 张粗面岩、3 张大理岩、4 张花岗斑岩、4 张珍珠岩、2 张霞石正长岩、2 张白云母片岩、1 张红柱石角岩、1 张堇青石角岩、1 张糜棱岩共 16 类 34 张图像,部分图像如图版 I 所示,用于验证。经验证,模型仅误识别 1 张长石石英砂岩。模型认为,被误判的长石石英砂岩为沉积石英岩的概率最高,为 36.8%;认为长石石英砂岩的概率次之,为 33.6%,二者非常接近。这是因为二者同属沉积岩且皆含石

图版 I Plate I



a~f. 验证图像合集; a. 橄榄岩; b. 辉绿岩; c. 蛇纹岩; d. 斜长岩; e. 长石砂岩; f. 麻粒岩

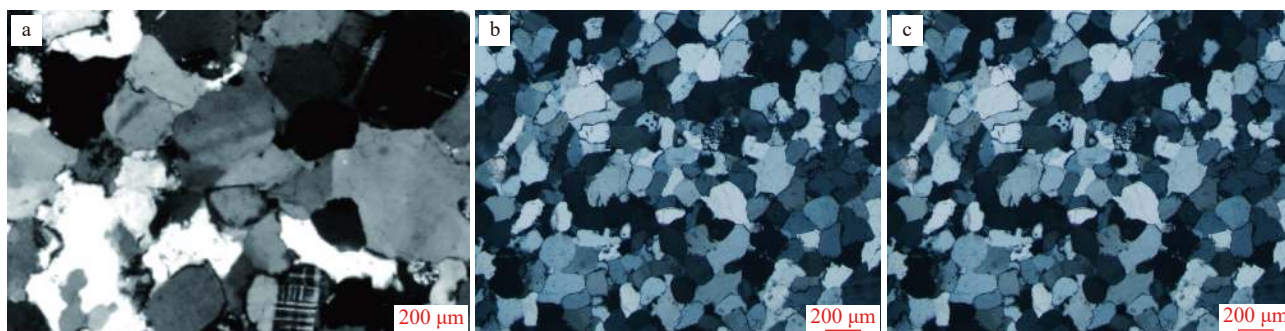


图 9 验证图像与数据集中图像对比

Fig. 9 Comparison between validation and data set images

a—验证集中的长石石英砂; b—数据集中的沉积石英岩; c—数据集中的长石石英岩

英等矿物,在数据集中,长石石英砂岩的薄片图像与沉积石英砂岩的薄片图像高度相似(图 9),而验证集中的长石石英砂岩薄片图像为灰度图,某些纹理更接近数据集中的沉积石英砂,因此模型产生误判。综上,模型能够准确识别数据集外岩石薄片图像的岩性,具有较好的鲁棒性。

3 结 论

本文通过增加岩石薄片图像的数量,采用 GELU 替换 ReLU6,作为 MobileViT 中 MV2 模块的激活函数,使用迁移学习将在 ImageNet 数据集上训练好的参数迁移到 MobileViT 上,使用余弦退火算法自动更新学习率,实现对 108 种岩石薄片图像的自动岩性识别,在测试集上的准确率达到 82.9%,对数据集以外的图像也能正确识别。得出以下结论。

(1) MobileViT 自身的 Attention 机制可以使模型对图像的感受野更大,能够在短时间的训练中快速提升准确率。

(2) 使用 GELU 替代 ReLU6 作为 MobileViT 模型的激活函数能够加快模型的收敛过程。

(3) 使用余弦退火算法和迁移学习能够缩短训练周期,避免模型陷入局部最优。

(4) 经过改进的 MobileViT 模型在保持参数不变的情况下能够有效提升岩性识别的准确率。

(5) 模型已充分学习到各类岩石薄片所特有的特征,具有较好的鲁棒性。

致谢: 审稿专家对文章提出了宝贵的意见和建议,在此表示衷心感谢。

参考文献

Dosovitskiy A, Beyer L, Kolesnikov A, et al. 2020. An image is worth

16×16 Words: Transformers for image recognition at scale[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.48550/arXiv.2010.11929.

Hendrycks D, Gimpel K. 2016. Bridging nonlinearities and stochastic regularizers with gaussian error linear Units[J]. abs, 1606.08415.

He K, Zhang X, Ren S, et al. 2016. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.48550/arXiv.1512.03385.

Hinton G. 2010. A practical guide to training restricted boltzmann machines[J]. Momentum, 9(1): 926–947.

Loshchilov I, Hutter F. 2016. SGDR: Stochastic Gradient Descent with Warm Restarts[J]. CoRR, 2016, arXiv preprint arXiv: 1608.03983.

Marmo R, Amodio S, Tagliaferri R, et al. 2005. Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples[J]. Computers & geosciences, 31 (5): 649–659.

Mehta S, Rastegari M. 2021. MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.48550/arXiv.2110.02178.

Pan S J, Qiang Y. 2010. A Survey on Transfer Learning[J]. IEEE Transactions on Knowledge and Data Engineering, 22(10): 1345–1359.

Sandler M, Howard A, Zhu M, et al. 2018. Mobile NetV2: Inverted residuals and linear bottlenecks[C]//IEEE. Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.48550/arXiv.1801.04381.

Simonyan K, Zisserman A. 2014. Very deep convolutional networks for large-Scale image recognition[J]. Computer Science, arXiv preprint arXiv: 1409.1556.

Singh N, Singh T N, Tiwary A, et al. 2010. Textural identification of basaltic rock mass using image processing and neural Network[J]. Computational Geosciences, 14: 301–310.

陈国军, 姜朕, 尹冲, 等. 2023. 基于深度学习的岩石孔隙分割方法[J]. 计算机与数字工程, 51(5): 1157–1162.

盖荣丽, 蔡建荣, 王诗宇, 等. 2021. 卷积神经网络在图像识别中的应用研究综述[J]. 小型微型计算机系统, 42(9): 1980–1984.

付光明, 严加永, 张昆, 等. 2017. 岩性识别技术现状与进展[J]. 地球物理学进展, 32(1): 26–40.

- 宫清顺, 黄革萍, 孟祥超, 等. 2012. 三塘湖盆地火山岩岩性识别方法[J]. *中国石油勘探*, 17(3): 37–41, 6.
- 李嘉璇. 2017. TensorFlow 技术解析与实战[M]. 人民邮电出版社: 314.
- 刘强, 滕奇志, 何海波. 2022. 岩石薄片多聚焦图像融合的深度学习算法研究[J]. *电子测量技术*, 45(3): 125–130.
- 赖文, 蒋璟鑫, 邱检生, 等. 2020. 南京大学岩石教学薄片显微图像数据集[J]. *中国科学数据 (中英文网络版)*, 5(3): 26–38.
- 马泽栋, 马雷, 李科, 等. 2022. 基于岩石图像深度学习的多尺度岩性识别[J]. *地质科技通报*, 41(6): 316–322.
- 熊越晗, 刘东燕, 刘东升, 等. 2021. 基于岩样细观图像深度学习的岩性自动分类方法[J]. *吉林大学学报 (地球科学版)*, 51(5): 1597–1604.
- 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.
- 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.
- 张珂, 冯晓晗, 郭玉荣, 等. 2021. 图像分类的深度卷积神经网络模型综述[J]. *中国图象图形学报*, 26(10): 2305–2325.
- 张野, 李明超, 韩帅. 2018. 基于岩石图像深度学习的岩性自动识别与分类方法[J]. *岩石学报*, 34(2): 333–342.
- 周永章, 陈烁, 张旗, 等. 2018. 大数据与数学地球科学研究进展——大数据与数学地球科学专题代序[J]. *岩石学报*, 34(2): 255–263.
- 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.
- 周志华. 2016. 机器学习[M]. 北京: 清华大学出版社.