

基于改进 InceptionV3 模型的常见矿物智能识别

彭伟航^{1,2}, 白 林^{2,3,4}, 商世为⁵, 唐小洁⁵, 张哲远⁵

PENG Weihang^{1,2}, BAI Lin^{2,3,4}, SHANG Shiwei⁵, TANG Xiaojie⁵, ZHANG Zheyuan⁵

1. 成都理工大学网络安全学院, 四川 成都 610059;
2. 自然资源部地质信息技术重点实验室, 北京 100037;
3. 成都理工大学数学地质四川省重点实验室, 四川 成都 610059;
4. 成都理工大学管理科学学院, 四川 成都 610059;
5. 成都理工大学地球科学学院, 四川 成都 610059

1. College of Network Security, Chengdu University of Technology, Chengdu 610059, Sichuan, China;
2. Key Laboratory of Geological Information Technology, Ministry of Natural Resources, Beijing 100037, China;
3. Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu 610059, Sichuan, China;
4. College of Management Science, Chengdu 610059, Sichuan, China;
5. College of Earth Sciences, Chengdu University of Technology, Chengdu 610059, Sichuan, China

摘要:以常见的 16 类矿物作为研究对象,收集每一类矿物约 1000 张图像,按比例划分为训练集、验证集和测试集,通过图像随机选取增加数据的多样性,建立矿物识别 InceptionV3 模型,训练 7 万次在测试集上获得 81% 的识别正确率。通过对损失函数的改进,引入 Center Loss 损失函数,训练 40 万次识别准确率提高到 86%。对分类的混淆矩阵分析发现,孔雀石等外观特征明显的矿物识别正确率很高,而闪锌矿等与其他矿物容易混淆导致正确率较低。从特征图分析看出,模型很好地提取了孔雀石的放射状特征,矿物图像特征向量聚集程度很高,也说明了模型的可靠性。

关键词:矿物图像;矿物识别;人工智能;深度学习

中图分类号:P585;P628 **文献标志码:**A **文章编号:**1671-2552(2019)12-2059-08

Peng W H, Bai L, Shang S W, Tang X J, Zhang Z Y. Common mineral intelligent recognition based on improved InceptionV3. Geological Bulletin of China, 2019, 38(12):2059-2066

Abstract: To study 16 kinds of common minerals, the authors collected 1000 images for each type, and then divided them into training set, validation set and test set. Before putting the images into the model, the authors selected a random area of each image for data augmentation. After training the InceptionV3 model with 70000 steps, the authors obtained an 81% accuracy in the test set. Through improving the loss function and introducing the Center Loss, the authors raised the accuracy to 86% after training 400000 steps. The obfuscation matrix shows that, the recognition accuracies for the minerals with obvious appearance characteristics such as malachite are higher while those for other minerals like sphalerite are less due to the obfuscation with other minerals. The analysis of the feature map shows that the model extracts the radial feature of malachite perfectly, and the feature vector of mineral image aggregate is in a high degree, which also can prove the reliability of the model.

Key words: mineral image; mineral recognition; artificial intelligence; deep learning

收稿日期:2019-04-17; **修订日期:**2019-07-29

资助项目:国家重点研发计划《基于地质云的地质灾害基础信息提取与大数据分析挖掘》(编号:2018YFC1505501)和成都理工大学国家级大学生创新创业训练计划项目《基于人工智能方法的岩石识别技术研究》(编号:201810616003)

作者简介:彭伟航(1999-),男,在读硕士生,从事人工智能相关研究。E-mail: pangdarren@outlook.com

通讯作者:白林(1980-),男,博士,讲师,从事地质信息化研究工作。E-mail: bailin@mail.cdut.edu.cn

矿物识别是一项重要的地质工作,它是根据对矿物的形态和物理性质特征进行观察以达到初步鉴定矿物的目的^[1]。当今科学研究的发展趋势正在朝着用数据说话的方向发展,地质学也不例外。地球科学的发展应适应当今时代的发展潮流,将大数据、区块链、人工智能等新一代的信息技术运用于地质工作中,研究并建立地质数学模型,再利用数学模型解决更多的地质问题^[2-4]。目前在很多网站可以收集到大量的矿物图像,采用人工智能技术从矿物图像中学习矿物识别的规律,使得矿物图像的自动识别正成为可能。

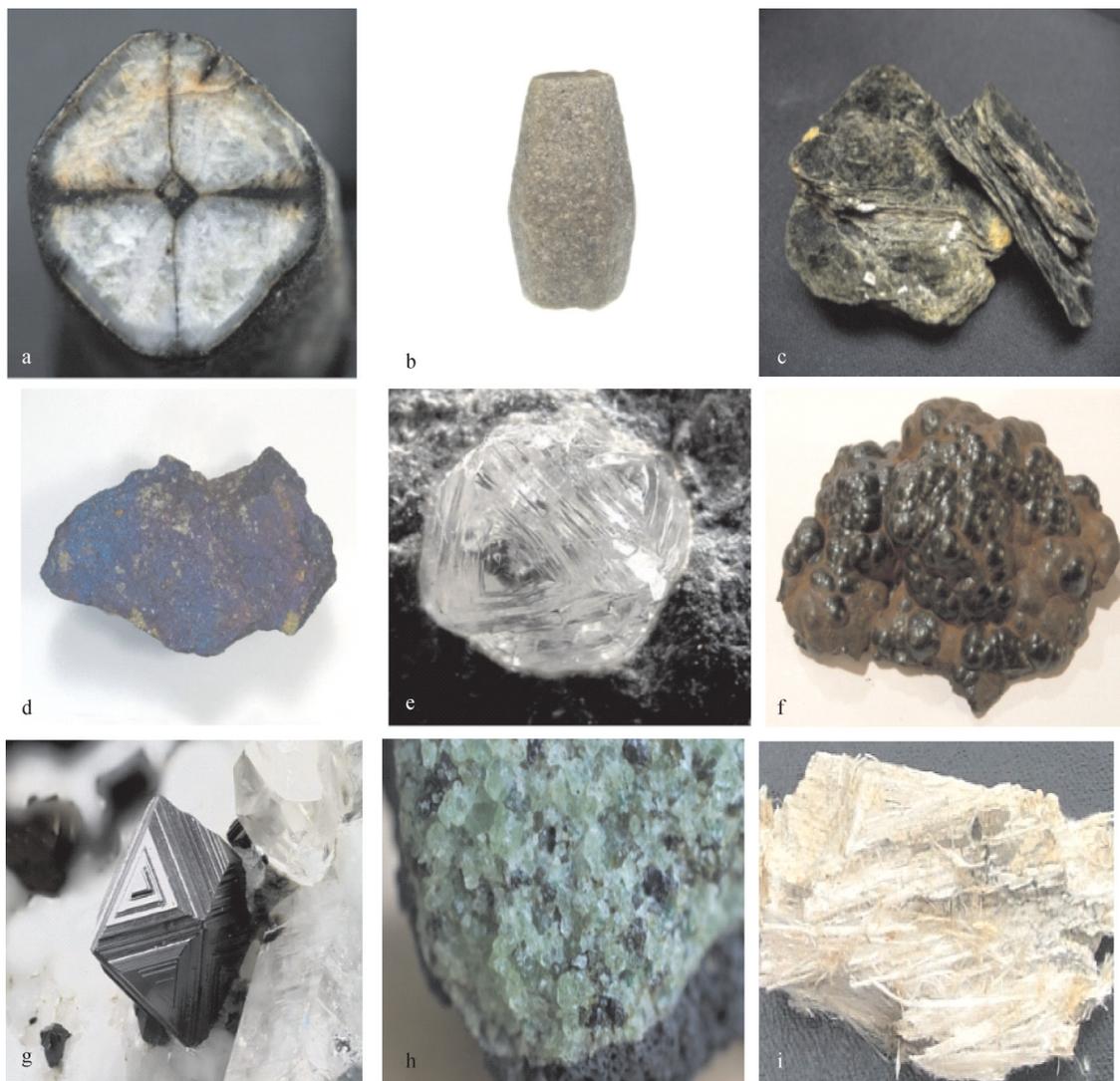
矿物肉眼鉴定通常先区分其光泽、颜色、条痕、硬度,再结合其形态、解理、断口、密度判断矿物种

类^[5]。例如,图版 I 的矿物图像来自互联网中最大的矿物数据库平台 MinDat。红柱石在其近似正方形的横截面上可见到捕获的炭质和粘土形成的黑十字特征;刚玉单晶常具有特有的腰鼓状形态;云母单晶常呈假六方板状,集合体常为片状集合体,且其解理面具有珍珠光泽;斑铜矿的蓝紫铜色十分典型;金刚石单晶常为浑圆状,且断面可见油脂光泽等(图版 I)。

1 矿物识别研究现状

一般的图像识别对象通常拥有较固定的颜色、形状等特征,而矿物图像颜色、形状等特征因其成分、生长环境等不同而变化较大,识别难度较大。

图版 I Plate I



a. 红柱石;b. 刚玉;c. 黑云母;d. 斑铜矿;e. 金刚石;f. 赤铁矿;g. 磁铁矿;h. 橄榄石;i. 石棉

Mengko 等^[6]提出,将颜色分析、纹理分析等图像技术用于矿物鉴定,将矿物图像数据处理后建立 RGB 和 HIS 颜色模型进行颜色分析,用其中的量化数据解释相关物理现象,并用该数据进行矿物识别和分类。Thompson 等^[7]从薄片数据中提取了一组基于色相、饱和度、强度和纹理测量值参数,并用神经网络对已知样品进行训练。对之前未见过的矿物样品进行网络测试,获得了高达 93% 的成功率。Baykan 等^[8]利用人工神经网络和颜色空间进行矿物识别,实验基于红、绿、蓝(RGB)颜色空间和色调、饱和度、明度(HSV)颜色空间提取纹理分割镜下的矿物,利用人工分类的矿物样本图像的像素值,训练了具有 k 次交叉验证的神经网络。对未见过的矿物样品进行测试,获得了 81%~98% 的成功率。叶润青等^[9]提出了基于图像分类的矿物含量测定及精度评价方法,该方法通过统计图像中每种矿物的像元数量测定矿物含量,并采用混淆矩阵评价含量测定精度。Ghiasi-Freez 等^[10]利用图像分析和智能判别分类器对薄片图像进行半自动化孔隙度识别。Aprile 等^[11]提出了一种结合图像分析技术和人工神经网络的方法,利用光学数字图像对考古陶片中的矿物包裹体和孔隙进行自动分类,特别是对石英、钙质团聚体和次生孔隙度的自动识别。Aligholi 等^[12]选取 CIElab 色彩空间处理薄片的数字图像,该方法有效利用了其基于色彩的特征而定义的色差度量,重新定义了矿物的 7 种基本光学性质,将局部二值模式算子和建模纹理集成到矿物识别方案中,识别矿物微观图像中的同质区域。Baklanova 等^[13]应用聚类分析算法对矿相显微镜拍摄的矿物图像中经过识别和颜色编码的矿石矿物的颜色和纹理特征进行提取,再对特征进行学习,之后根据学习的特征对矿物进行识别。白林等^[14]基于卷积神经网络构建了岩石识别的深度学习模型,使用深度学习的方法对 15 种常见岩石的图像进行识别,达到 63% 的识别准确率。

综上所述,对于矿物识别,国内外研究多是根据镜下图像或利用光谱对岩石中的矿物进行识别,手标本尺度的矿物图像直接识别的情况几乎没有。目前已经有学者采用深度学习的方法对岩石和镜下矿物进行分类研究,本文采用深度学习的方法对常见矿物进行分类识别。

表 1 Inception V3 模型结构

Table 1 The model structure of Inception V3

处理过程	卷积核大小 /步长或者图注	图像大小	备注
conv	3×3/2	399×399×3	输入矿物图像
conv	3×3/1	199×199×32	
conv padded	3×3/1	197×197×32	
pool	3×3/2	197×197×64	
conv	3×3/1	98×98×64	
conv	3×3/2	98×98×80	
conv	3×3/1	47×47×192	矿物特征提取
3×Inception	图 1-a	47×47×288	
5×Inception	图 1-b	23×23×768	
2×Inception	图 1-c	11×11×1280	
pool	8×8	11×11×2048	
linear	logits	1×1×2048	
softmax	classifier	1×1×1000	输出矿物类别

2 矿物识别 Inception V3 模型

Inception V3 模型是谷歌公司在 2015 年提出的,这里建立矿物识别的 Inception V3 模型。表 1 为 Inception V3 的结构,其输入为一个大小为 399×399×3(像素)的矿物图像,经过 6 个设计好的卷积后得到一个 47×47×192 的特征图,再依次经过 3 个 Inception 模块和全连接层后得到矿物分类结果。

Inception V3 模型实现了一种新的卷积模块,此模块使网络的宽度多样化。在每个模块中,分多条支路对输入的特征图进行特征提取,每条支路的深度、感受野、池化参数都不同,最后的输出为之前每条支路特征的融合,得到一个富含不同尺度信息的特征。输入特征图在经过多种卷积的计算后再在第三维度融合得到输出特征图(图 1),其中 1×1 的卷积用于特征图不同通道的融合。

3 实验过程

将常见的 16 类矿物作为研究对象,包括刚玉、黄铁矿、黄铜矿、金、金红石、孔雀石、蓝铜矿、菱铁矿、绿帘石、闪锌矿、石英、天青石、萤石、自然铜、自然银、方铅矿,收集每一类矿物约 1000 张图像,按照 8:1:1 的比例划分为训练集、验证集和测试集。训练集用于模型的训练,验证集用于监视模型训练时

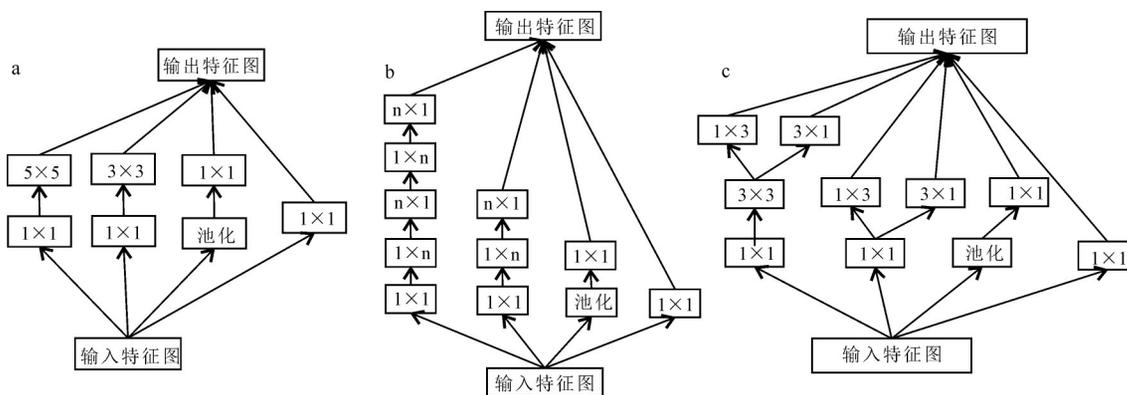


图1 Inception模块结构

Fig. 1 The structure of Inception modular

的状况进而调整训练参数,测试集单独用于最终得出模型的准确率。

3.1 图像预处理方法

通过在图像里随机选取一个区域,取区域内的图像进行训练。Random_i(i=1,2...4)是(0,1)上服从均匀分布的随机数,Area是输入图像内一个随机的区域。预处理过程如图2所示,输入图像经过随机函数随机选取一个区域,然后裁剪区域内图像缩放到相同的大小。

$$\text{Random}_i = U(0,1), i=1,2,\dots,4 \quad (1)$$

$$\text{Area} = (\text{Random}_1 \times \text{height}, \text{Random}_2 \times \text{width}, \text{Random}_3 \times \text{height}, \text{Random}_4 \times \text{width}) \quad (2)$$

3.2 Center Loss

模型的训练使用 Center Loss 代替原始的 Softmax Loss 以增加模型的内聚性。假设 $x_j \in R^d (j=1, 2, \dots, m)$ 表示模型最后输出的特征向量,输出类别为 $y_i (i=1, 2, \dots, m)$, d 为特征向量的维度, m 为每一个批次训练图像的数量。 $W_j \in R^d (j=1, 2, \dots, n)$ 是最后一个全连接层权重 $W \in R^{d \times n}$ 的第 j 列, n 为类别总数, b 为偏置项参数。

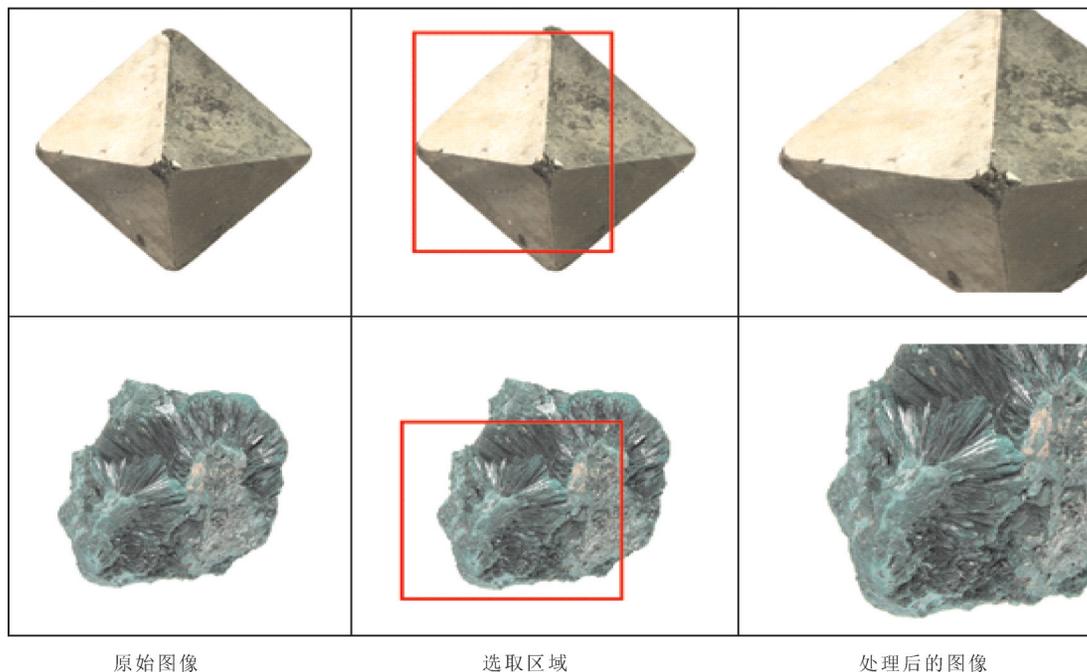


图2 预处理过程图

Fig. 2 Preprocessing diagram

原始的 Softmax Loss 的公式如下:

$$\text{Soft max Loss} = - \sum_{i=1}^m \log \frac{e^{w_j^T x_i + b_j}}{\sum_{j=1}^n e^{w_j^T x_i + b_j}} \quad (3)$$

Center Loss 在 Softmax Loss 的基础上,加上一个限制模型参数扩张的变量,使最终的模型能够在内聚性上有一定的表现。公式如下:

$$\text{Center Loss} = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (4)$$

$c_{y_i} \in R^d$ 第 y_i 类的中心($i=1,2,\dots,n$)。总的损失函数 Loss 公式为:

$$\text{Loss} = \text{Soft max Loss} + \gamma \times \text{Center Loss} \quad (5)$$

中心点在每个批次训练后更新,其更新公式为

$$\Delta C_j = \frac{\sum_{i=1}^m \delta(y_i=j) (C_j - x_j)}{1 + \sum_{i=1}^m \delta(y_i=j)} \quad (6)$$

其中当 x 为真时, $\delta(x)$ 的值等于 1, 否则等于 0。

3.3 模型训练

设定训练的正则系数为 0.000002, 输入图像大小为 399×399(像素), 学习率 0.0003, 学习率衰减系数 0.94, 学习率衰减周期 3000, Center Loss 的系数 $\gamma=0.5$ 。

训练过程中的损失函数和正确率变化如图 3 所示。图 3 中, 蓝色线表示训练数据结果, 黄色线表示验证数据结果。Center Loss 为中心损失, Softmax

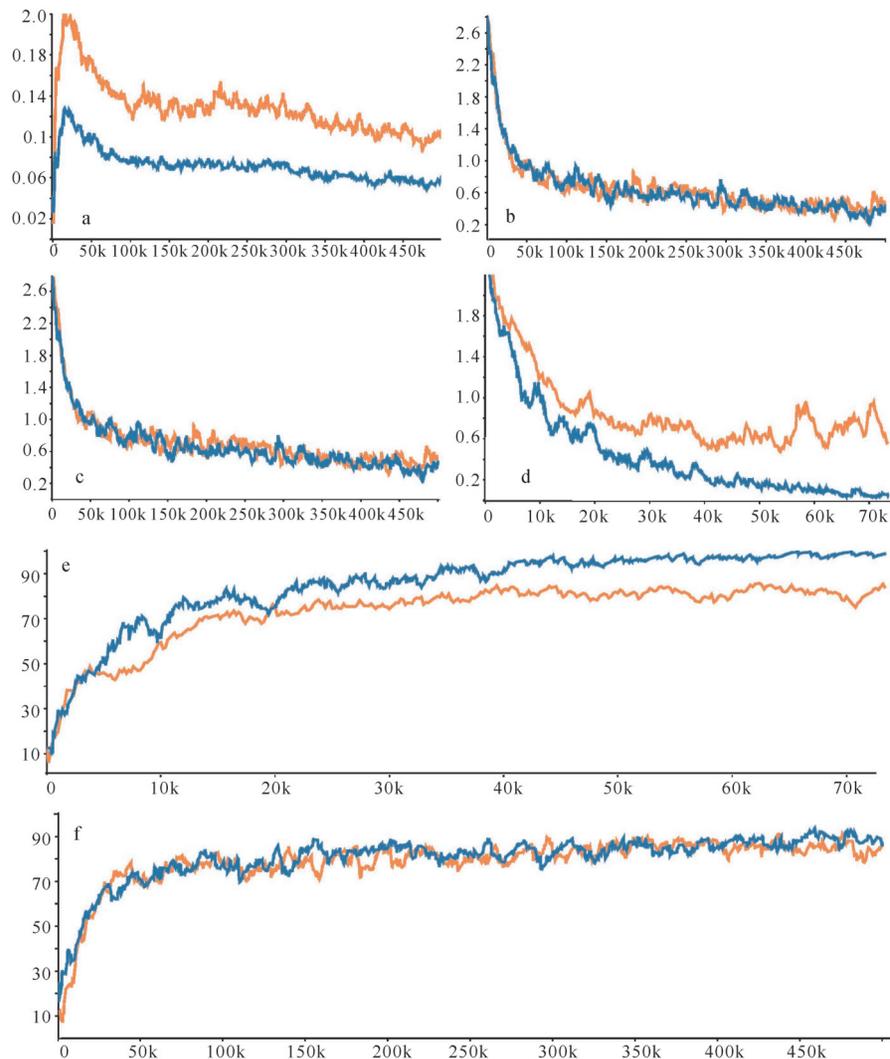


图 3 训练过程中的损失函数图与正确率图

Fig. 3 Training Loss and accuracy

a—Center Loss; b—Softmax Loss; c—Total Loss; d—Single Softmax Loss;
e—仅使用 Softmax Loss 的 Accuracy Top_1; f—加入 Center Loss 后的 Accuracy Top_1

Loss 为原始损失, Total Loss 为总的损失, Accuracy top_1 为模型的正确率。图 3-a、图 3-b、图 3-c 分别为加入 Center Loss 后的 Center Loss 函数图、Softmax Loss 函数图及总的损失函数图, 图 3-d 为单使用 Softmax Loss 的损失函数图。图 3-e、图 3-f 分别为仅使用 Softmax Loss 的训练图和加入 Center Loss 后的训练图。

对比图 3-c 和图 3-d 发现, 在加入 Center Loss 后, 训练难度加大, 使模型的损失函数很难下降, 训练步数远大于仅使用 Softmax Loss 的步数, 一定程度上反映了 Center Loss 对模型的约束性, 以及其对模型的正则效果。同时由于模型的鲁棒性的提高, 验证集和训练集的损失值和正确率都很接近。

图 4 给出了不同损失函数的矿物图像特征向量分布图, 发现 Softmax Loss 无法为模型最终的结果提供足够的内聚性, 而使用 Center Loss 训练的模型特征向量向中心聚集, 在保证类间距离足够大的同时类内距离也足够小。

4 结果分析

表 2 给出了测试集中矿物识别的混淆矩阵。矿物整体识别率在 86% 左右, top_5 达到 99% 的效果, 并在最终结果中表现出很强的鲁棒性, 每种矿物的识别率为 74%~98%。

表 2 中识别率最高的 2 种矿物为孔雀石(98%)

和蓝铜矿(98%)。两者都是外观特征明显的矿物。孔雀石主要以孔雀绿色为鉴定特征, 蓝铜矿晶体则具有深蓝色的特征。

矿物的成分变化、结晶环境不同等因素导致的外观多样性是影响矿物识别率的主要原因。识别率较低的矿物有闪锌矿(74%)、天青石(75%)和方铅矿(76%)。闪锌矿的透明度随其铁含量增高而逐渐降低, 测试集中 9% 的闪锌矿被识别为金红石, 查验发现该部分闪锌矿为红色半透明至黑色不透明略带红色调, 外观上与金红石类似。无色透明的柱状天青石易被识别为石英。类似的成分变化导致的颜色变化是矿物鉴定过程中的难题之一。

此外, 计算机对矿物图像的切割和选取也可能导致识别错误, 如闪锌矿测试集中 4% 被识别为黄铜矿, 经查验发现为闪锌矿与黄铜矿在同一张图像所致。被识别成闪锌矿的方铅矿测试图像中也存在方铅矿和闪锌矿共生的现象。

5 矿物图像中的形态特征提取

将 Inception 模型训练过程中产生的卷积图输出, 可以观察到卷积过程中是否有效提取矿物图像的特征。

图 5 是不同的放射状孔雀石原始图像和特征图的对比, 可以看到计算机很好地对具有放射状形态特征的不同孔雀石矿物图像进行了特征提取, 说明

表 2 验证集中矿物识别混淆矩阵

Table 2 Confusion matrix for mineral recognition in validation set

矿物	刚玉	黄铁矿	黄铜矿	金	金红石	孔雀石	蓝铜矿	菱铁矿	绿帘石	闪锌矿	石英	天青石	萤石	自然铜	自然银	方铅矿
刚玉	0.88	0.01	0	0	0.01	0	0	0.02	0	0	0.04	0	0.02	0	0.01	0.01
黄铁矿	0.01	0.85	0.01	0	0.03	0	0	0.01	0	0.03	0.02	0	0	0	0.02	0.02
黄铜矿	0	0.05	0.88	0	0.01	0	0	0	0	0.04	0	0	0	0	0.02	0
金	0	0.01	0.06	0.87	0	0	0	0.04	0	0	0	0	0	0.02	0	0
金红石	0.01	0.03	0	0	0.84	0	0	0	0.01	0	0.04	0	0	0	0.04	0.04
孔雀石	0	0	0	0	0	0.98	0	0	0	0	0	0	0.01	0	0.01	0
蓝铜矿	0	0	0	0	0.01	0	0.98	0	0	0	0	0	0	0	0	0.01
菱铁矿	0.02	0.01	0.06	0	0.02	0	0	0.78	0.01	0.02	0.01	0	0	0.06	0	0.01
绿帘石	0	0.01	0.03	0.01	0	0	0	0.02	0.92	0.01	0	0	0	0	0	0
闪锌矿	0	0.03	0.04	0	0.09	0	0	0.04	0.02	0.74	0	0.02	0	0	0	0.02
石英	0.04	0.01	0	0	0.02	0	0	0	0.01	0	0.85	0	0.06	0	0.01	0
天青石	0.02	0	0	0	0.02	0	0	0.02	0	0.02	0.09	0.75	0.06	0	0.02	0.02
萤石	0.04	0.01	0	0	0	0.02	0	0.01	0	0.02	0.03	0.02	0.86	0	0	0
自然铜	0	0	0.01	0	0.01	0	0	0.05	0	0	0	0	0	0.92	0.01	0
自然银	0	0	0.03	0	0.08	0	0	0	0	0	0	0	0	0.04	0.83	0.02
方铅矿	0.02	0	0.01	0	0.02	0	0.01	0.01	0.01	0.06	0.04	0	0.03	0	0.02	0.76

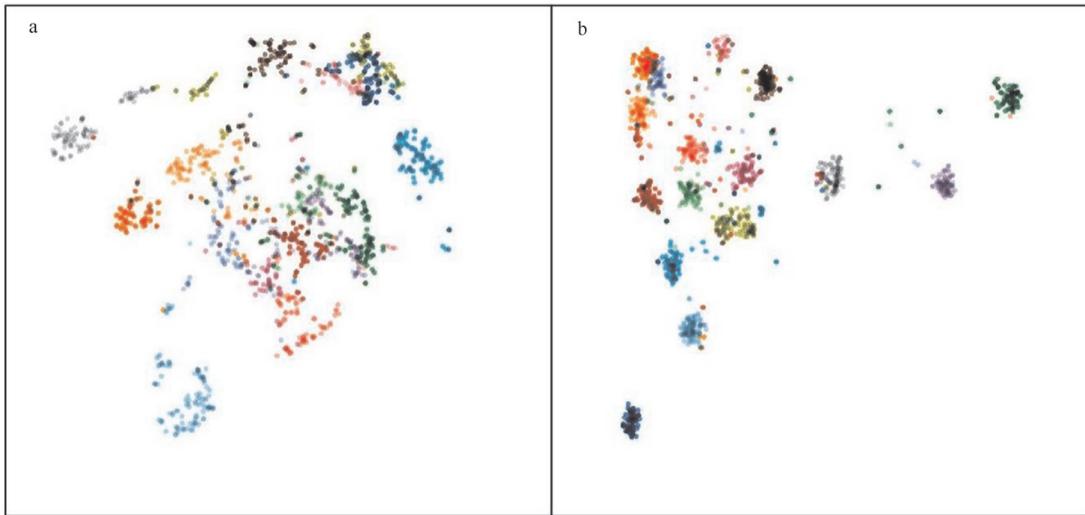


图4 不同损失函数的矿物图像特征向量分布

Fig. 4 Feature vector distribution of mineral images after using different loss functions

a—Softmax Loss;b—Center Loss

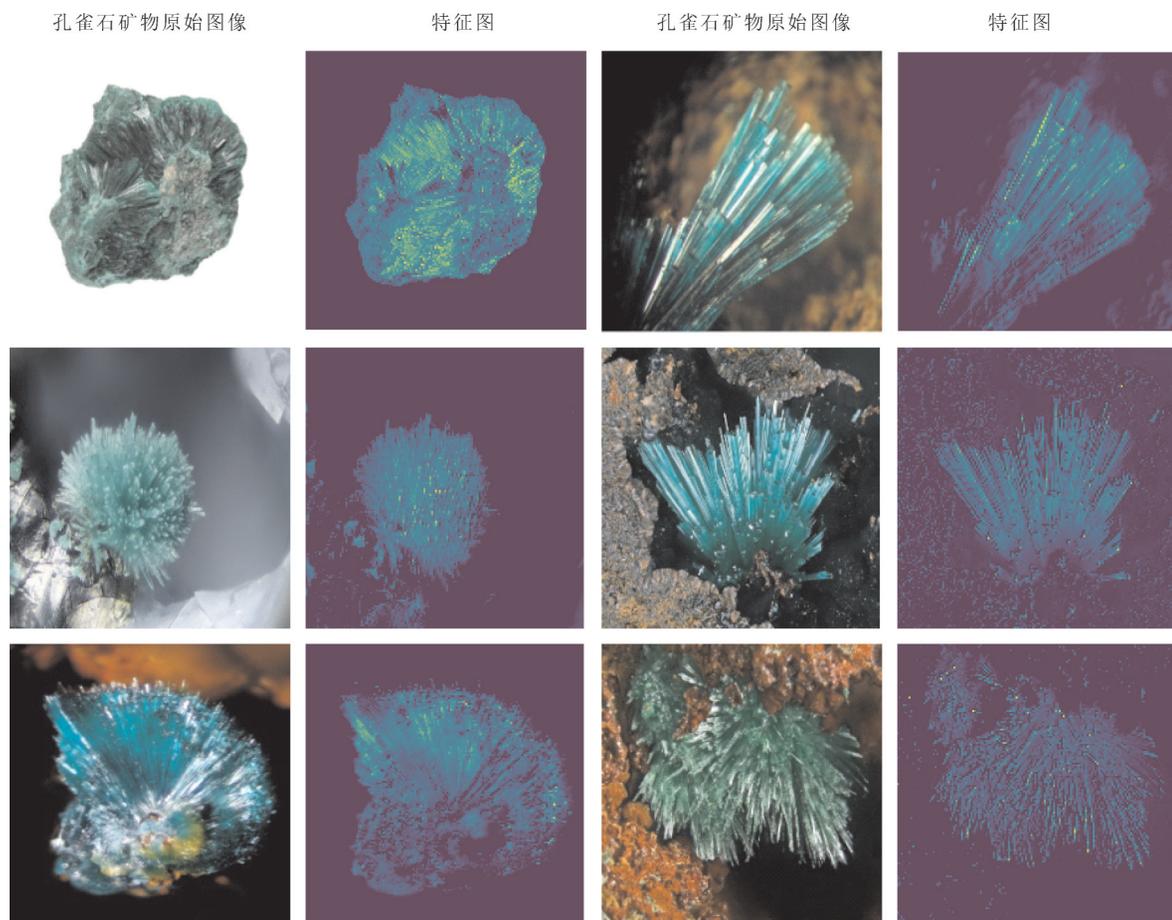


图5 孔雀石的特征提取图

Fig. 5 Feature extraction map of malachite

在深度学习中矿物形态是很重要的识别特征。

6 结 语

收集了16类常见矿物的图像,建立矿物识别的Inception V3模型,常规的Softmax函数作为损失函数,模型训练后,验证正确率为81%。加入Center Loss函数作为损失函数来限制模型的收敛条件,模型改进后的识别率达到了86%。测试集中矿物识别的混淆矩阵说明孔雀石和蓝铜矿具有较高的识别准确率,而闪锌矿、天青石和方铅矿识别准确率较低,主要原因是矿物成分变化导致的颜色变化引起。通过对矿物特征提取图的分析,发现该模型能很好地提取矿物的外观特征。目前数据来源有限,将来可以通过收集更多的数据来提高模型的精度。

致谢:本文工作得到中国地质调查局发展研究中心谭永杰总工程师、朱月琴高级工程师等多位专家和学者的指导和建议,还有许多整理数据的同学,在此表示诚挚的感谢。

参考文献

- [1]方少木. 矿物岩石肉眼鉴定[M]. 北京: 煤炭工业出版社, 1980.
- [2]赵鹏大. 大数据时代数字找矿与定量评价[J]. 地质通报, 2015, 34(7): 1255-1259.
- [3]严光生, 薛群威, 肖克炎, 等. 地质调查大数据研究的主要问题分析[J]. 地质通报, 2015, 34(7): 1273-1279.
- [4]姚震, 杜子图. 关于新时代地质调查工作的战略思考[J]. 地质通报, 2018, 37(11): 2120-2124.
- [5]赵明等. 矿物学导论[M]. 北京: 地质出版社, 2017.
- [6]Mengko T R. Digital Image Processing Technique in Rock Forming Minerals Identification[C]//IEEE Asia-Pacific Conference on Circuits and Systems, 2000:4.
- [7]Thompson S, Fueten F, Bockus D. Mineral identification using artificial neural networks and the rotating polarizer stage[J]. Computers & Geosciences, 2001, (27): 1081-1089.
- [8]Baykana N A, Yilmaz N. Mineral identification using color spaces and artificial neural networks[J]. Computers & Geosciences, 2010, (36): 91-97.
- [9]叶润青, 牛瑞卿, 张良培, 等. 基于图像分类的矿物含量测定及精度评价[J]. 中国矿业大学学报, 2011, 40(5): 810-815+822.
- [10]Ghiasi-Freez J, Soleimanpour I, Kakhodaie-Ilkhchi A, et al. Semi-automated porosity identification from thin section images using image analysis and intelligent discriminant classifiers[J]. Computers & Geosciences, 2012, (45): 36-45.
- [11]Aprile A, Castellano G, Eramo G. Combining image analysis and modular neural networks for classification of mineral inclusions and pores in archaeological potsherds[J]. Journal of Archaeological Science, 2014, (50): 262-272.
- [12]Aligholi S, Khajav R, Razmara G. Automated mineral identification algorithm using optical properties of crystals[J]. Computers & Geosciences, 2015, (85): 175-183.
- [13]Baklanova O E, Baklanov A, Shvets O Y. Methods and algorithms of computer vision for automated processing of mineral rocks images[C]//2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics. Timisoara, 2015: 449-454.
- [14]白林, 姚钰, 李双涛, 等. 基于深度学习特征提取的岩石图像矿物成分分析[J]. 中国矿业, 2018, 27(7): 178-182.