

利用改进 AlexNet 卷积神经网络识别石墨^①



徐小平¹, 余香佳¹, 刘广钧¹, 刘 龙²

¹(西安理工大学 理学院, 西安 710054)

²(西安理工大学 自动化与信息工程学院, 西安 710048)

通信作者: 余香佳, E-mail: yuxiangjiavere@foxmail.com

摘 要: 为促进矿业领域向信息化、智能化的方向转变, 实现对石墨的智能识别尤为关键. 针对人工识别石墨花费时间长、效率低等问题, 提出了一种改进的 AlexNet 网络应用于石墨的图像识别. 首先通过随机裁剪、依概率水平翻转和归一化处理等手段对数据集进行图像预处理达到数据增强的目的; 然后采用激活函数 ReLU6 压缩动态范围, 使算法更稳健; 运用批标准化算法进行归一化加快收敛速度; 修改卷积核的大小增强泛化能力; 在全连接层加上 dropout 正则化进一步防止过拟合. 在仿真实验中, 与已有方法进行比较, 所给方法降低了损失值, 提高了石墨的识别平均准确率.

关键词: 石墨; 识别; 卷积神经网络; 特征提取; 小样本数据集

引用格式: 徐小平, 余香佳, 刘广钧, 刘龙. 利用改进 AlexNet 卷积神经网络识别石墨. 计算机系统应用, 2022, 31(2): 376-383. <http://www.c-s-a.org.cn/1003-3254/8311.html>

Graphite Recognition Using Improved AlexNet Convolution Neural Network

XU Xiao-Ping¹, YU Xiang-Jia¹, LIU Guang-Jun¹, LIU Long²

¹(School of Sciences, Xi'an University of Technology, Xi'an 710054, China)

²(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: The intelligent recognition of graphite is particularly critical to the transformation of the mining industry to informatization and intellectualization. To address the long time and low efficiency in manually identifying graphite, this study proposes an improved AlexNet network for graphite image recognition. First, image preprocessing is performed on the data set through random cropping, horizontal flipping according to probability, and normalization to achieve data augmentation. Then, the activation function ReLU6 is employed to compress the dynamic range so that the algorithm can become more robust. The batch standardization algorithm is used for normalization to speed up the convergence, and the convolution kernel is resized to enhance the generalization ability. Finally, dropout regularization is added to the fully connected layer to further prevent overfitting. Compared with the existing method, the proposed method reduces the loss value and improves the average accuracy of graphite recognition in the simulation experiment.

Key words: graphite; recognition; convolution neural network (CNN); feature extraction; small sample data set

随着科学技术的快速发展, 非金属矿的各种矿种应用日新月异, 其战略意义逐步提高. 尤其是石墨被誉为“二十一世纪支撑高新技术发展的战略资源”, 具有良好的耐高温性、导电性、化学稳定性和抗热震性等^[1],

不仅是新能源、新材料、航天航空等领域关键原材料, 还是耐火材料、电极电刷、铅笔等传统工业领域的重要原料^[2]. 在高科技产业发展中, 石墨所占地位越来越重要, 其独特的物理特性使其能够应用于国计民生的

① 基金项目: 国家自然科学基金 (61773016); 陕西省创新能力支撑计划 (2020PT-023); 陕西省自然科学基金基础研究计划 (2018JQ1089)

收稿时间: 2021-04-21; 修改时间: 2021-05-19; 采用时间: 2021-05-28; csa 在线出版时间: 2022-01-17

各个领域, 具有十分重要的工业价值^[3,4].

以卷积神经网络 (convolution neural network, CNN) 为代表的深度学习算法在大数据时代得到快速发展, 在机器学习和计算机视觉等研究领域备受关注^[5,6]. CNN 作为一种深度前馈神经网络, 在图像识别领域取得了不小的突破^[7]. 相比于传统的分类方法, 基于卷积神经网络的全自动分类方法脱离了人力识别的繁琐, 同时利用反向传播算法自动优化模型参数, 获得当前条件最优模型. 深度学习的概念是 Hinton 等人^[8]于 2006 年提出的, 它能更抽象、更深层次描述物体的属性和特征, 并且设计出 AlexNet 网络, 从此掀起了一波深度学习的热潮. 2016 年, Ebrghimi 等人^[9]提出了一种基于层次分析法和特征映射的图像视觉方法, 利用图像识别辨别出 16 种常见矿石; 2017 年, 胡发焕^[10]采用模糊支持向量机 (NP-FSVM) 对矿石图像进行初步识别; 同年, 伍云霞^[11]将 BP 神经网络与小波变换相结合, 取得较高的识别率; 2019 年, 吕红杰^[12]使用 AlexNet 卷积神经网络用于煤岩识别并取得较好的图像识别结果; 2020 年, 王李管等人^[13]使用改进的卷积神经网络应用于钨矿分类问题.

作为图像识别领域中的强大算法^[6,7], 目前鲜有研究者将 CNN 应用到石墨的图像识别任务中. 与 GoogLeNet、ResNet 等更先进的模型相比, AlexNet 具有更简单的网络结构和更少的网络参数^[7]; 相比于现有的轻量级网络深度深, 训练难度不会很大, 表征能力强, 更方便做出不同的改进, 可以节省大量模型训练时间, 更不易过拟合^[6]. 因此, 本文给出利用 AlexNet 和深度学习的图像识别技术进行石墨与非石墨识别分选. 针对石墨图像本身特征及数据集过小, 通过对数据集进行图像预处理达到数据增强的目的, 同时对激活函数、归一化层、卷积核、Dropout 层和优化器进行替换更新. 仿真实验结果表明, 与相关算法相比, 本文所提方法复杂度

低、参数量少, 提高了网络的收敛速度, 减低了损失率, 有效地实现了石墨的图像识别目标.

1 传统 AlexNet 网络模型

1.1 卷积神经网络

AlexNet 属于 CNN 的一种, 而 CNN 属于人工神经网络, 已经广泛应用于图像识别等机器学习领域^[6]. 由于它的网络结构采用了权值共享方式, 更加类似于生物神经网络, 其优点相对于一般神经网络, 减少了网络的权值数量. CNN 最大的优势在于通过感受野和共享权重减少网络的参数, 即参数减少和权值共享. 该方法使得训练速度更快, 而且训练需要的样本更少. CNN 基本网络结构如图 1 所示, 分为 4 个部分: 输入层 (input), 卷积层 (convolution, Conv), 全连接层 (full connected layer, FC) 和输出层 (output).

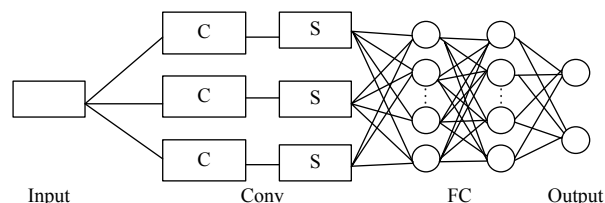


图 1 基本 CNN 网络结构

卷积神经网络与其他神经网络模型最大的区别在于, 卷积神经网络在神经网络的输入层前面连接了卷积层, 这样卷积层就变成了卷积神经网络的数据输入层. 卷积层是卷积神经网络中独特的网络结构, 主要进行图像的特征提取; 池化层针对特征图进行稀疏处理, 减少数据的运算量; 全连接层在整个卷积神经网络中起到“分类器”作用, 它将在已学到的“内容”映射到样本标记空间 (把前面提取到的特征综合起来). CNN 的连接过程如图 2 所示.

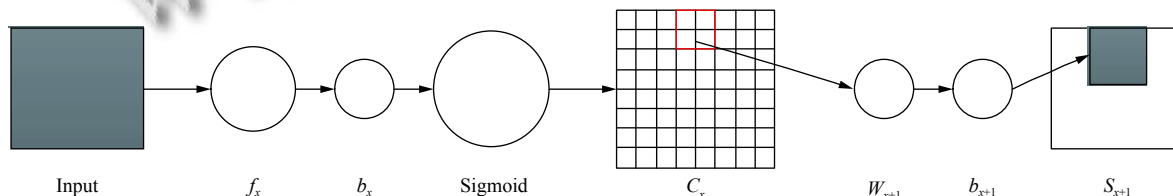


图 2 CNN 连接过程

图 2 中, f_x 表示数字滤波器; b_x 表示偏置; C_x 表示卷积层的特征图; W_{x+1} 表示下采样的权值; b_{x+1} 表示对应的权重; S_{x+1} 表示下采样层.

1.2 AlexNet 卷积神经网络

AlexNet 是 ISLVR 2012 (ImageNet large scale visual recognition) 竞赛的冠军网络, 分类准确率提升到

了80%以上,其网络结构如图3所示。

该网络包含8层结构,其中5层卷积层、3层全连

接层。在 AlexNet 网络中还包含3个池化层 (Pooling), 卷积层和池化层交替连接使用。

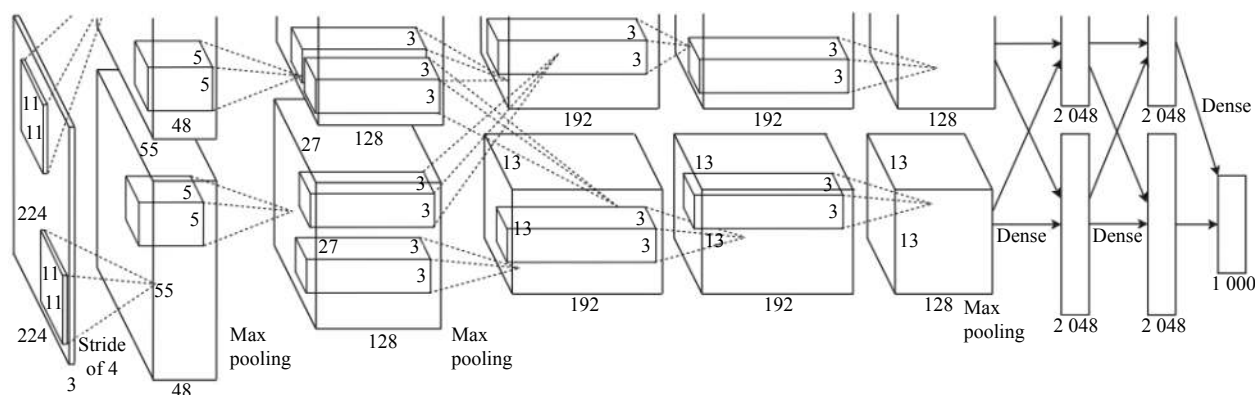


图3 传统的 AlexNet 网络结构

在 AlexNet 网络中,5 层卷积层分别使用 11×11 、 5×5 、 3×3 、 3×3 、 3×3 的卷积核进行特征提取。AlexNet 舍弃此前普遍使用的 Sigmoid 或 tanh 函数,采用 ReLU 作为激活函数;由于经过 ReLU 函数得到的值域没有区间,因此 AlexNet 提出了局部响应归一化 (local response normalization, LRN),对 ReLU 得到的数据进行归一化处理,抑制小的神经元,提高模型的泛化能力。

2 改进的 AlexNet 模型

2.1 数据增强

为了提高模型的泛化能力,使用大样本数据集进行训练是最好的办法。对于小样本数据集,在深度学习中可以采用数据增强的方法来解决数据不足的问题^[14]。这里利用 PyTorch 的图像预处理包 Transform 对数据集中的图片进行如图4所示的预处理操作。

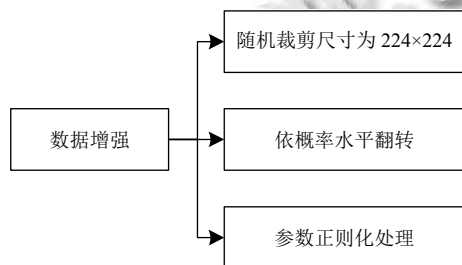


图4 数据增强操作

首先将测试集进行随机裁剪得到图像像素为 224×224 ,并依概率(设置 $p = 0.5$)对图像进行水平翻转,最后对处理过的数据集进行归一化处理,达到数据增强的目的。

2.2 激活函数

在 AlexNet 网络中,虽然将 ReLU 函数作为激活函数有效克服了 Sigmoid 和 tanh 函数梯度消失、收敛速度慢等缺点,但是 ReLU 在 $x > 0$ 的区域使用 x 进行线性激活,有可能造成激活后的值太大,影响模型的稳定性,这会使得网络的训练速度下降,甚至可能会降低网络的泛化性能。为抵消 ReLU 的线性增长部分,本文使用 ReLU6 函数。

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

相比于 ReLU 函数,模型中使用 ReLU6 作为非线性层,在低精度计算时能压缩动态范围,算法更稳健。

2.3 批量归一化算法

原始 AlexNet 中使用了 LRN 对第一、二层进行归一化操作,以增强模型的泛化性能,但 LRN 算法对模型的实际改善效果有限,同时极大增加模型的训练时间。文献[15]中提出批量归一化 (batch normalize, BN) 算法能够减小经过激活函数而导致的数据偏移,有效解决训练过程中数据分布不一致的问题。BN 算法通过计算小批次中的均值和方差达到特征归一化的目的。

给定一个尺寸为 n 的批量 $B = \{x_1, x_2, \dots, x_m\}$,在训练过程中,式(2)和(3)记录各个批量的均值和方差:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (3)$$

最后在实际的测试中求得整个数据集的均值与方差,进行归一化处理,即:

$$\hat{x} = \frac{x_i - \hat{\mu}_B}{\sqrt{\hat{\sigma}_B^2 + \varepsilon}} \quad (4)$$

其中, $\hat{\mu}_B$ 表示整个数据集的均值, $\hat{\sigma}_B^2$ 表示整个数据集的方差; ε 常表示系数,用于增加数值的稳定性.

BN 算法的关键之处在于进行变换重构,引入可拉伸参数 γ 与偏移参数 β 进行如式 (10) 的修正处理.

$$y_i = \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i) \quad (5)$$

其中, $\gamma^{(i)} = \sqrt{D[x_{(i)}]}$, $\beta^{(i)} = E[x_{(i)}]$, 两个参数随着网络权重的更新而更新,最终输出归一化样本值 y_i .

2.4 卷积层

VGG (visual geometry group) 网络是牛津大学计

算机视觉组和 Google DeepMind 研发的深度卷积神经网络. 该网络结构使用 3×3 大小的卷积核来扩大通道数,以提取更复杂和更具有表达力的特征,且具有扩展性强、泛化能力好等特点^[16]. 因此本文借鉴 VGG 网络的优点,将 AlexNet 网络中的卷积核大小均修改为 3×3 ,大大减少参量数目,提取更为具体的特征.

2.5 改进的 AlexNet 模型结构

本文选用 Adam 优化器,以 0.000 2 为初始学习率;选择交叉熵作为损失函数,避免均方误差带来的学习速率逐渐降低的问题. 由于在 AlexNet 网络的全连接层引入了大量的参数,因此在全连接层后加入 Dropout,这种技术减少了神经元复杂的共同适应,使模型更有效的进行融合. 改进后的 AlexNet 网络模型如图 5 所示.

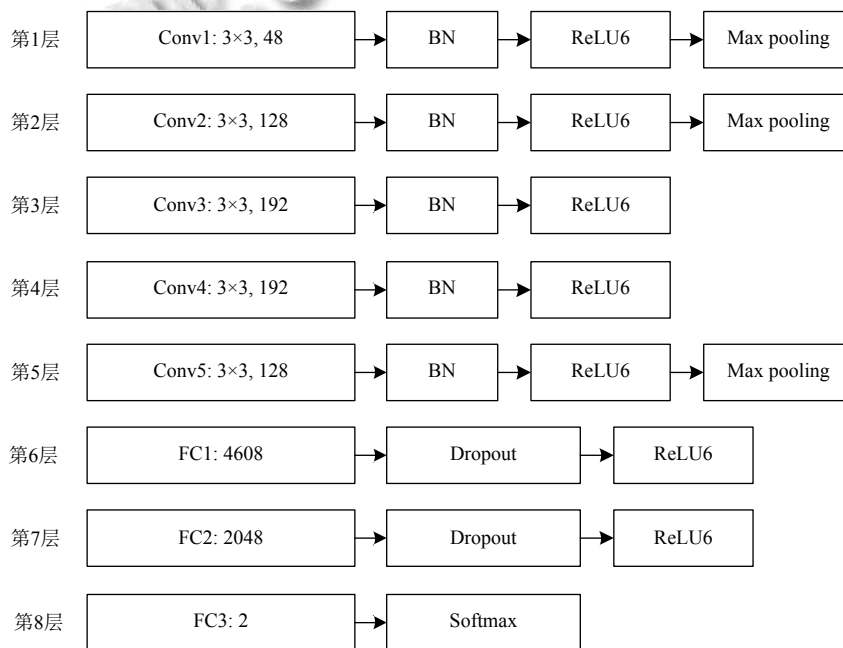


图5 改进的 AlexNet 网络模型

本文采用改进的 AlexNet 网络模型完成数据集的自动分类任务. 模型的总体框架如图 6 所示.

3 实验参数与数据集

3.1 实验环境及参数设置

本文在 Python 语言的编译环境下,使用 1.5.0 版本 PyTorch 框架完成实验仿真,通过 Transform 实现图像数据的预处理. 实验环境为: Mac OS 10.14.6 版本操作系统,处理器为 1.6 GHz Intel Core i5,未使用 GPU

加速,运行一次网络平均消耗 9 000–10 000 s. 设置权重衰量为 0.01,用于防止过拟合;迭代周期 epoch 为 100,一个 batch 大小设置为 32;所有权重参数初始化为均值为 0、标准差为 0.01 的高斯分布,同时将训练集图片在输入前随机打乱,减少图片顺序对模型的影响.

3.2 样本采集与扩充

本文图像样本采集自实验室现有石墨样品,将其清洗干净后进行干燥操作,利用 iPhone 手机后置摄像头从不同方向及角度拍摄样本图像视频,随后对视频

进行分帧处理,共得到 166 张样本图像.同时在网络上收集非石墨矿石图像 180 张,最终将图像分为石墨与非石墨两种类别,共 346 张彩色图片.图 7 为石墨与非石墨的样本图像.

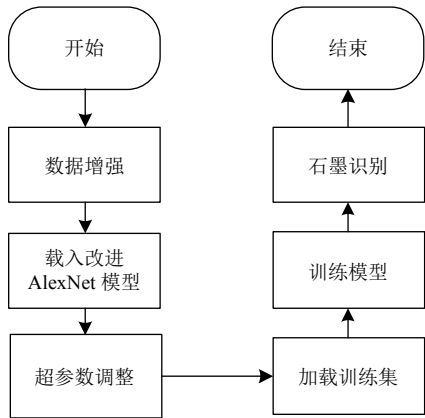


图 6 利用改进的 AlexNet 的图像分类流程图

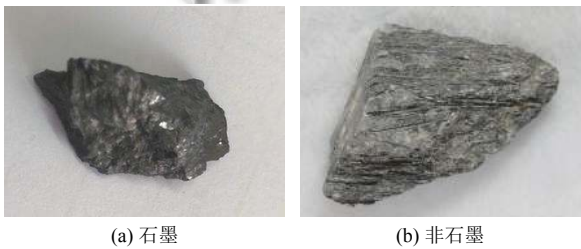


图 7 石墨与非石墨样本图像

一般的,数据集越充足,模型训练的效果越好,其泛化能力得到增强.为解决训练样本不足的问题,本文对以上数据集进行了数据扩充,对收集到的图像分别做中心裁剪、样本旋转(180°、90°)、绕 X 轴翻转以及均值模糊操作,最终得到 1 857 张图像,按照比例随机选取数据集中 955 张(80%)作为训练集对网络进行训练,902 张(20%)作为测试集验证网络的性能.将训练集与测试集分开处理.数据集分类情况如表 1 所示.

表 1 石墨数据集

类别	石墨	非石墨
石墨数据集	955	902
训练集	764	722
测试集	191	180

3.3 数据集预处理

为降低样本的不平衡性和解决数据集过小容易造

成的过拟合现象,本文对数据集进行数据增强(data augmentation, DA)操作,提高网络分类准确率.在 DA 操作中,常见方法包括尺度变换、图像对比度调整、图像裁剪、仿射变换、水平或垂直翻转等,针对不同的数据集可选择不同的增强方法.基于矿石图像的特征,将训练集图片在输入前随机打乱,减少图片顺序对模型的影响;将训练集中的每张图片进行随机裁剪,统一图像像素为 224×224,并依概率(设置 $p=0.5$)对图像进行水平翻转,最后以 0.5 的均值与 0.5 的标准差分别对每个通道的数据正则化处理,达到数据扩充与数据增强的目的.

4 实验结果与分析

4.1 评价指标

本文以准确率(Accuracy, acc)和损失(Loss)这两个指标评估一个网络的训练效果.测试准确率是指模型在测试集上输出正确结果的比率,能反映一个网络的使用效果,是很重要的指标.其定义公式如式(6)所示:

$$acc = \frac{n_{\text{correct}}}{n} \quad (6)$$

其中, n_{correct} 表示测试集中网络识别正确的个数, n 表示测试集样本个数.

神经网络的训练过程是最小化损失函数的过程, Loss 即为损失函数的值.实际上,损失函数计算的是模型在测试集上的均方误差(MSE, E):

$$E = \frac{1}{n} \sum_i (\hat{y}_i^{\text{test}} - y_i^{\text{test}})^2 \quad (7)$$

4.2 实验结果

本文使用改进的 AlexNet 模型与传统的 AlexNet 模型分别对数据集进行训练,图 8 展示了使用 ReLU6 函数与 ReLU 函数的损失与准确率的变化曲线.通过实验仿真发现,使用 ReLU6 函数可使得损失值由 0.092 5 降为 0.077 2,达到相对较小.

针对归一化层,本文对比使用 BN 算法与未使用 BN 算法两种情况进行仿真,结果如图 9 所示.

图 9 结果表明,使用 BN 算法有效降低模型测试的损失值,同时避免了模型陷入过拟合风险.为了进一步观察卷积层在使用 BN 算法前后特征提取的情况,随机提取出数据集中的一张图片,如图 10 所示,分类为石墨.

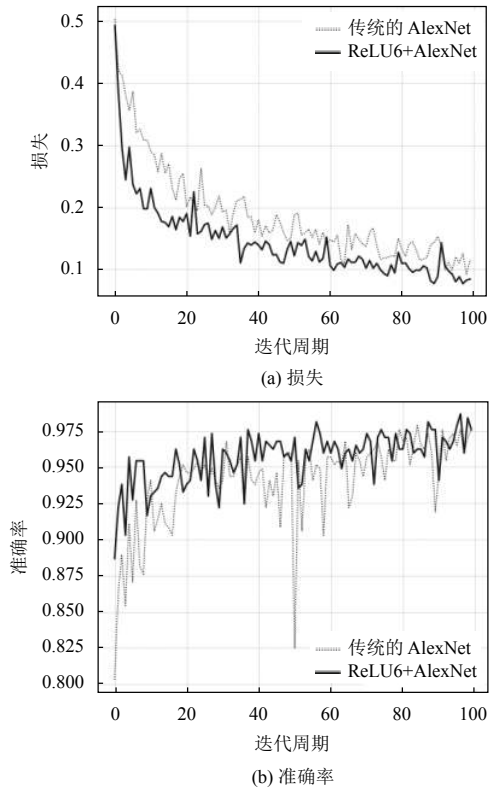


图8 不同激活函数模型的对比曲线

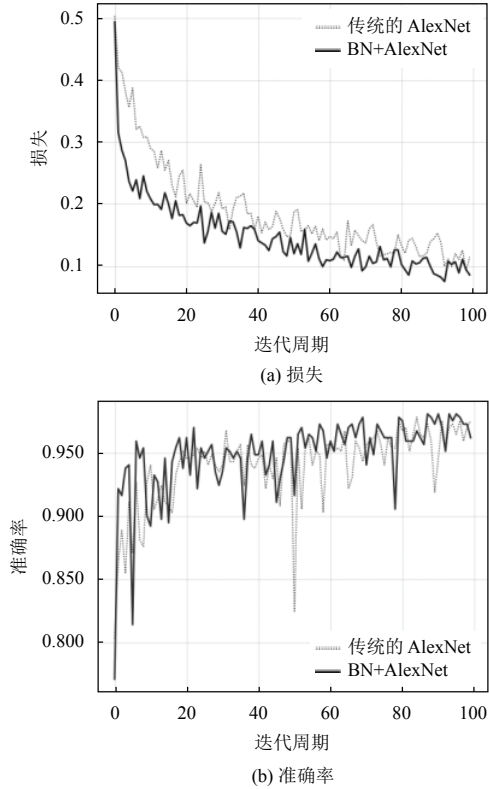


图9 使用 BN 算法的模型对比曲线

依次打印出图 10 在 Conv1, Conv2 与 Conv3 的前 12 个通道的特征提取图, 如图 11 所示。



图 10 特征提取石墨样本图像

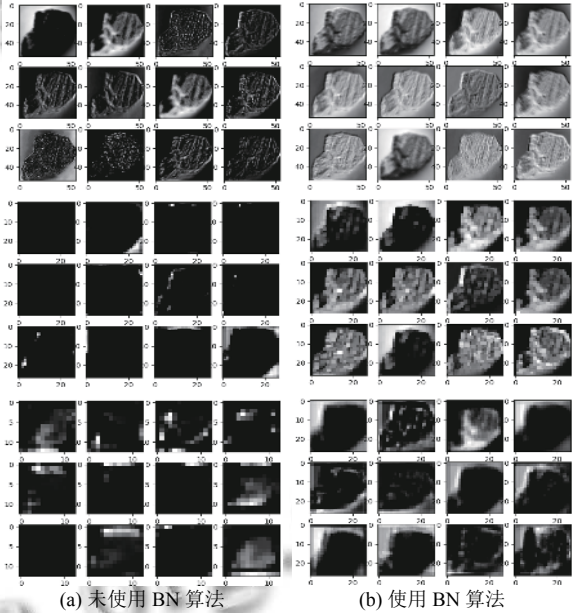


图 11 使用 BN 算法的卷积层特征提取情况

由图 11 可知, 传统的 AlexNet 网络在第 3 卷积层所提取的特征已经较为模糊, 学习能力不是很好; 相比使用 BN 算法后的模型, 特征提取能力增强, 可保证网络进行快速、充分的学习. 使用 BN 算法前后的图像识别损失与准确率如表 2 所示。

表 2 加入 BN 算法的损失与准确率对比		
方法	损失	平均准确率 (%)
未加入BN (传统的AlexNet)	0.092 5	94.037 7
BN+AlexNet	0.075 0	94.897 6

除激活函数与归一化层的改进, 本文还对卷积核大小做出修改, 如图 12 所示。

图 12 展示了在以上改进的基础上, 使用 3×3 小卷积核的损失值与准确率对比情况. 明显看出使用小卷

核提高了模型的泛化能力,减低了模型在测试集上的损失值.

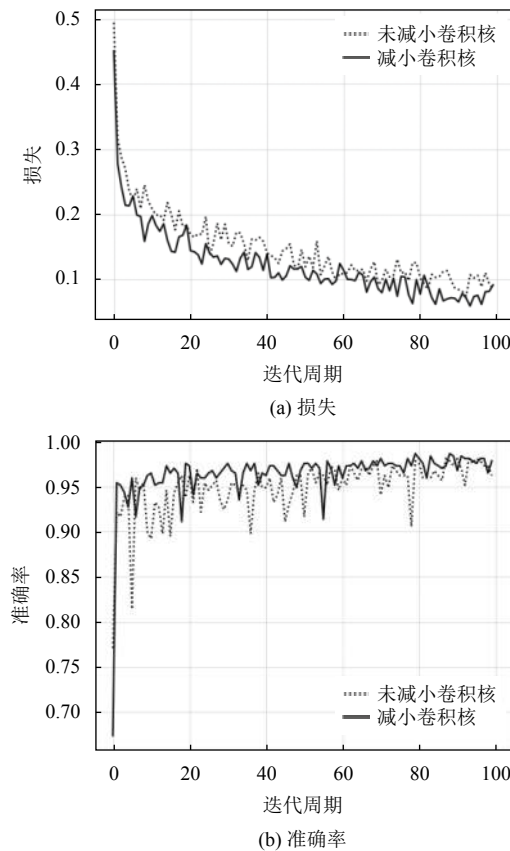


图 12 修改卷积核的模型对比曲线

为了评估改进的 AlexNet 网络对石墨分类数据集的识别性能,将传统的 AlexNet 网络、文献 [7]、文献 [17] 的 AlexNet 网络作为对照组,提取出 Conv2 的情况,如图 13 所示.

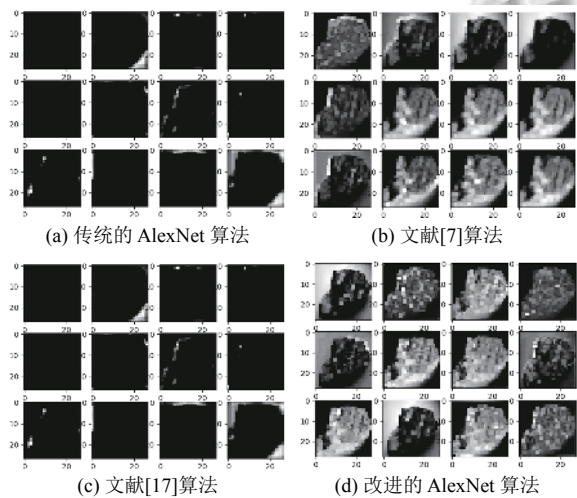


图 13 Conv2 在不同模型下的特征提取情况

通过图 13 可以看出,使用改进的 AlexNet 网络在第 2 卷积层中的特征情况较其他两种模型更为清晰,提取学习到更多的特征.最后得到模型损失与准确率的对比曲线如图 14 所示,进一步验证本文提出的改进 AlexNet 网络的有效性.

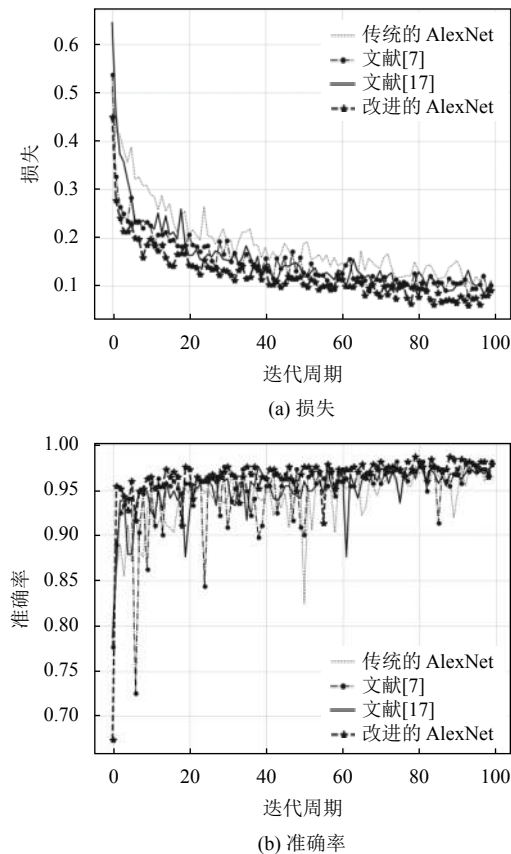


图 14 传统的 AlexNet 与改进的 AlexNet 模型对比曲线

通过对测试集的实验,将传统的 AlexNet 网络、文献 [7]、文献 [17] 的方法作为对照组,与本文方法的损失与准确率对比如表 3 所示.

表 3 损失率与准确率对比			
方法	损失	最高准确率 (%)	平均准确率 (%)
传统的 AlexNet	0.092 5	97.843 7	94.037 7
文献[7]	0.080 6	98.113 2	94.770 9
文献[17]	0.086 7	97.843 7	94.997 3
改进的 AlexNet	0.059 9	98.652 3	96.363 9

由表 3 可知,使用本文提出的改进的 AlexNet 网络较其他两种情况效果更好,有效验证了本文所提出的网络模型的可行性与有效性.

5 结语

石墨数据收集难度较大,该领域研究的专业性较强,导致石墨的智能识别发展不易.本文使用改进的 AlexNet 卷积网络实现石墨识别,此过程不需要人工提取石墨的图像特征,智能化程度较高,训练过程较为简单.通过对 AlexNet 的激活函数、归一化层和超参数的优化,提高了整体模型的鲁棒性,在不增加模型计算复杂度的同时避免了过拟合问题.实验结果表明,本文改进算法较传统 AlexNet 的分类平均准确率得到提高,损失值降低,模型具有较强的泛化性.将改进 AlexNet 模型应用于石墨图像识别,增加了选矿流程的自动化程度,减少了人工拣选的工作量.但本文使用数据集为小数据集且石墨样本为块状,增加石墨种类对数据集进行扩充,对市面上绝大部分的流动石墨进行识别是下一步的研究方向.

参考文献

- 1 孙斌,孔庆友.解读地球密码系列:平凡的奇迹——石墨.济南:山东科学技术出版社,2016.
- 2 刘之能,申士富,刘海营,等.我国石墨选矿技术及装备进展.现代矿业,2020,36(8):143–146,162. [doi: [10.3969/j.issn.1674-6082.2020.08.042](https://doi.org/10.3969/j.issn.1674-6082.2020.08.042)]
- 3 张旭东,邱杨率,吴益民,等.原料特性和制备工艺对柔性石墨膜导热性能的影响.硅酸盐通报,2019,38(12):3988–3992,4012. [doi: [10.16552/j.cnki.issn1001-1625.2019.12.041](https://doi.org/10.16552/j.cnki.issn1001-1625.2019.12.041)]
- 4 高惠民,张凌燕,管俊芳,等.石墨、石英、萤石选矿提纯技术进展.金属矿山,2020,(10):58–69. [doi: [10.19614/j.cnki.jsks.202010006](https://doi.org/10.19614/j.cnki.jsks.202010006)]
- 5 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012. 1097–1105.
- 6 石翠萍,谭聪,左江,等.基于改进 AlexNet 卷积神经网络的人脸表情识别.电讯技术,2020,60(9):1005–1012. [doi: [10.3969/j.issn.1001-893x.2020.09.002](https://doi.org/10.3969/j.issn.1001-893x.2020.09.002)]
- 7 甘岚,郭子涵,王瑶.基于径向变换和改进 AlexNet 的胃肿瘤细胞图像识别方法.计算机应用,2019,39(10):2923–2929.
- 8 Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786): 504–507. [doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647)]
- 9 Ebrahimi M, Abdolshah M, Abdolshah S. Developing a computer vision method based on AHP and feature ranking for ores type detection. Applied Soft Computing, 2016, 49: 179–188. [doi: [10.1016/j.asoc.2016.08.027](https://doi.org/10.1016/j.asoc.2016.08.027)]
- 10 胡发焕.基于多特征融合的钨矿石初选方法研究[博士学位论文].南昌:南昌大学,2017.
- 11 伍云霞.矿用图像识别与应用.天津:天津科学技术出版社,2017.30–40.
- 12 吕红杰.卷积神经网络在煤岩图像识别中的应用研究.科技创新导报,2019,16(9):137–139. [doi: [10.16660/j.cnki.1674-098X.2019.09.137](https://doi.org/10.16660/j.cnki.1674-098X.2019.09.137)]
- 13 王李管,陈斯佳,贾明滔,等.基于深度学习的黑钨矿图像识别选矿方法.中国有色金属学报,2020,30(5):1192–1201.
- 14 Faris PD, Ghali WA, Brant R, *et al.* Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses. Journal of Clinical Epidemiology, 2002, 55(2): 184–191. [doi: [10.1016/S0895-4356\(01\)00433-4](https://doi.org/10.1016/S0895-4356(01)00433-4)]
- 15 Ioffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift. arXiv: 1502.03167v3, 2015.
- 16 李校林,钮海涛.基于 VGG-NET 的特征融合面部表情识别.计算机工程与科学,2020,42(3):500–509. [doi: [10.3969/j.issn.1007-130X.2020.03.016](https://doi.org/10.3969/j.issn.1007-130X.2020.03.016)]
- 17 Wang JT, Zheng H, Huang Y, *et al.* Vehicle type recognition in surveillance images from labeled web-nature data using deep transfer learning. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(9): 2913–2922. [doi: [10.1109/TITS.2017.2765676](https://doi.org/10.1109/TITS.2017.2765676)]