

利用迁移学习和焦点损失卷积神经网络的石墨分类^①



徐小平¹, 余香佳¹, 刘广钧¹, 王峰²

¹(西安理工大学 理学院, 西安 710054)

²(西安交通大学 数学与统计学院, 西安 710049)

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

摘要: 为了使得优质石墨资源得到优质优用, 提出利用迁移学习和焦点损失卷积神经网络的石墨分类识别算法. 在自建的初始数据集基础上, 通过对数据集的离线扩充与在线增强, 有效扩大数据集并减低深层 CNN 过拟合的风险. 以 VGG16、ResNet34 和 MobileNet V2 为基础模型, 重新设计新的输出模块载入全连接层, 提高了模型的泛化能力与鲁棒性; 结合焦点损失函数, 修改模型超参数并在石墨数据集上训练. 实验仿真发现, 本文所提方法的准确率均在 95% 以上, 识别准确率提高, 收敛速度加快, 模型更加稳定, 证明了所提算法的可行性与有效性.

关键词: 石墨; 图像分类; 迁移学习; focal loss; 卷积神经网络

引用格式: 徐小平, 余香佳, 刘广钧, 王峰. 利用迁移学习和焦点损失卷积神经网络的石墨分类. 计算机系统应用, 2022, 31(3): 248-254. <http://www.c-s-a.org.cn/1003-3254/8378.html>

Graphite Classification Using Transfer Learning and Focal Loss Convolutional Neural Network

XU Xiao-Ping¹, YU Xiang-Jia¹, LIU Guang-Jun¹, WANG Feng²

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

²(School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: For better use of high-quality graphite resources, this study proposes a graphite classification and recognition algorithm based on transfer learning and focal loss convolutional neural network (CNN). The offline expansion and online enhancement of the self-built initial data set can effectively expand the data set and reduce the overfitting risk of deep CNN. With VGG16, ResNet34, and MobileNet V2 as basic models, a new output module is redesigned and loaded into the full connection layer, which improves the generalization ability and robustness of the model. Combined with the focal loss function, the hyperparameters of the model are modified and trained on the graphite data set. The simulation results show that the proposed method has the accuracy improved to above 95% with faster convergence and a more stable model, which proves the feasibility and effectiveness of the proposed algorithm.

Key words: graphite; image classification; transfer learning; focal loss; convolutional neural network (CNN)

石墨具有耐高温、导电、导热、抗腐蚀等优异特性, 我国自 2016 年将晶质石墨列为战略性矿产以来, 石墨产业政策倾向由合理开发利用过渡到战略性新兴产业产品的研发^[1]. 国家“十三五”发展规划将石墨材料列为新材料发展战略重点^[2]. 一般的, 石墨可分为

两大类——天然石墨和人造石墨. 二者结构相近, 物理化学性质相似, 但使用路径大不相同. 在冶金、机械工业、化学工业等众多应用领域, 辨别石墨的类别尤为重要, 若将两者混为一谈将造成决策性错误, 带来大量的资源浪费和经济损失^[3], 因此从我国的战略需求角度

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: 2021-05-07; 修改时间: 2021-06-08; 采用时间: 2021-07-01; csa 在线出版时间: 2022-01-24

来说,研究石墨的智能分类对我国石墨资源分类管理有极大的深远意义。

传统的识别石墨类别的方法大多为显微镜识别或化学物质识别,这些方法人为影响因素较大,时间成本高,给工业带来巨大的挑战。如今,基于深度学习的人工智能研究为计算机视觉提供了最先进的解决方案。深度卷积神经网络(convolution neural network, CNN)是一种深度学习方法,目前在图像识别应用的不同学科中受到青睐。人们注意不到的小细节,可以通过 CNNs 很容易地分辨出来,这种算法只需最少的预处理,就可以直接从像素图像中提取视觉模式。CNN 结构是由 LeNet 架构引入的^[4], AlexNet^[5] 使 CNN 流行起来。从那时起,随着各种各样的设计和应用, CNN 的受欢迎程度呈指数级增长。

由于 4 类石墨图像相似度高,浅层 CNN (如 AlexNet 等网络) 识别率低,模型相对简单,可能会限制其分类精度;综合识别准确率、识别时间及模型参数大小等考虑因素,本文最终选用网络结构更复杂、层次更深的 VGG16、ResNet34 和 MobileNet V2 进行改进,这些网络在缩短培训时间的情况下,对硬件要求更弱,降低了计算量,能够有效提取图像深层特征,避免了手工提取特征的复杂方式,提高了识别准确率。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;同时重置迁移网络的载入模块,调试每层连接层的输出参数,提高网络的收敛速度,并添加 Dropout 层优化模型,使迁移网络与模型有效融合。

传统的 CNN 模型大多使用交叉熵损失函数(cross-entropy loss, CEL),但 CEL 只关心正确标签的准确性,忽略了非正确标签的差异,因此大多数研究者对损失函数进行了研究。Triplet Loss 是 Schroff 等人^[6]于 2015 年提出,主要用于训练差异性较小的样本,但该函数训练过程不稳定,收敛速度慢,有可能陷入局部最优。2017 年,针对密集物体的目标检测,He 等人^[7]提出焦点损失函数(focal loss, FL),不仅解决了样本的不平衡问题,还可以解决难易样本不平衡的问题,使网络更重视难样本。针对石墨数据集中图像差异较小、不易分辨,且难样本远多于易样本的特点,本文选用 FL 作为损失函数,调试 FL 的参数设置,提升了网络的稳定性。仿真实验结果表明,结合 FL 的改进迁移网络应用在石墨数据集上,模型的鲁棒性和泛化能力更强,识别精度

更高,取得了理想的识别效果。

1 实验数据

1.1 数据集的构建

本文实验样本采集自陕西省内乡县瑞鑫高温复合材料制品有限公司的鳞片石墨、人造石墨、膨胀石墨及炭黑样本,利用 iPhone 手机后置摄像头于正午同一位置从不同方向拍摄样本图像视频 3 分钟,视频的分辨率为 1080 p;将拍摄的石墨视频按帧提取图像,每秒 1 帧,筛选后共得到 715 张样本图像,分类保存。实验过程中实验的所有图像均从保存的数据获取。4 类石墨图像如图 1 所示。

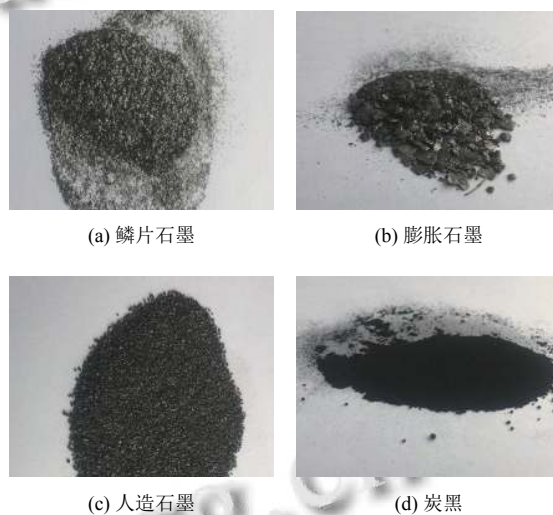


图 1 4 种石墨类别

1.2 数据集的离线扩充与在线增强

获取多种石墨类别的数据集是困难的,收集到的图像会出现本身分布不均、图像少等特点,因此对数据集进行数据增强很有必要。数据增强可以分为离线增强和在线增强。离线增强直接对数据集进行处理,常用于数据集较小的时候。而在线增强适用于大型数据集,在模型训练过程中获得 batch 数据之后,然后对这个 batch 的数据进行增强,如旋转、平移、翻折等相应的变化。本文在初始数据集较小时,采取离线扩充,对数据集直接处理,增大数据集样本量,对收集到的图像分别做垂直翻转、样本旋转 90°、绕 X 轴翻转以及仿射变换等操作,最终得到 3 575 张图像。

通过在线增强对数据进行处理,减少深层 CNN 的过拟合问题。基于石墨图像的特征,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,最后对处理过的数据集进行归一化处理,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。

按照比例随机选取数据集2860张(80%)作为训练集对网络进行训练,715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

	鳞片石墨	膨胀石墨	人造石墨	炭黑
数据集	1 020	965	715	875
训练集	816	772	572	700
测试集	204	193	143	175

2 实验原理与方法

2.1 迁移学习

近年来的成功案例^[5,8-10]将深度CNN作为解决具有挑战性的计算机视觉任务的主流方法。然而,从头开始训练一个网络需要大量的训练数据、时间和GPU。我们只有3575张训练和测试图像,小的数据集无法训练一个深入的CNN,因此本文融合迁移学习的思想进行卷积神经网络的训练。

迁移学习是指使用已经针对不同任务训练过的模型的过程,希望该模型能够有足够的广义信息来解决新的特定任务^[11],该方法利用一个巨大的数据库上的预先训练模型CNN来帮助学习目标任务^[12]。本文分类图像是石墨,具有不同于猫、狗等的视觉表现,在大图像上学习到的视觉表现可能无法很好地表示石墨图像,因此需要对预先训练好的CNN结构进行修改以适应我们的任务。Donahue等人^[8]和Zeiler等人^[13]提供了证据,证明了从预先训练的CNN中学习到的通用图像表示优于先前最先进的手工制作特征。

2.2 卷积神经网络模型

目前常见的CNN有AlexNet、VGG、ResNet、MobileNet等,由这些深层网络学习的现成图像表示是强大和通用的,已用于解决许多视觉识别问题^[8,14]。考虑到这些现成的CNN特性的良好性能,它们已经成为解决大多数计算机视觉问题的主流图像特征^[15]。

深度卷积神经网络具有深度深、识别效果好等特

点,而VGG16、ResNet34和MobileNet V2是近年来具有代表意义的CNNs,在图像分类领域取得了优异的成绩。VGG通过使用多个较小尺寸的卷积核来代替较大尺寸的卷积核,以保证在具有相同感受野的条件下可以继续增加网络深度。VGG16具有16个隐藏层(13个卷积层和3个全连接层),是深度学习框架中适用于图像分类的常用的经典CNN模型。ResNet提出残差结构并使用 3×3 卷积进行降维,使得在没有引入额外参数且不增加网络计算复杂度的前提下,ResNet在参数、深度、宽度以及计算成本上都具有优势。本文数据集不到4000张,使用ResNet34模型的深度即可很好训练样本。MobileNet V2^[16]作为2018年提出的轻量级网络,始终使用DW卷积提取特征,提出Inverted Residual,可以在实现上减少内存的使用。基于上述优点,本文以VGG16、ResNet34和MobileNet V2网络为基础建立石墨分类识别模型,其对比分析如表2。

表2 模型对比分析表

网络信息	VGG16	ResNet34	MobileNet V2
年份	2014	2015	2018
Top-1准确率(%)	71.5	—	71.7
参数数目	138 357 544	63 470 656	4 253 864
层数	16	34	—
BN	否	是	是
残差结构	否	是	是

2.3 焦点损失函数

在图像分类中,损失函数一般为交叉熵损失函数。交叉熵是指真实输出(概率)与预测输出(概率)的距离,因此交叉熵的值越小,两个概率分布越接近。针对石墨多分类问题,数据集中存在大量简单的负样本,这些样本在迭代过程中产生的损失占据了总损失的大部分,这可能会导致我们偏离正确的优化方向。为了改善这种情况,本文使用焦点损失函数,公式如式(1)所示:

$$L_{FL} = \begin{cases} -(1-y')^{\gamma} \log y', & y = 1 \\ -y^{\gamma} \log(1-y'), & y = 0 \end{cases} \quad (1)$$

首先在交叉熵的基础上加上一个因子 $(1-y')^{\gamma}$,用于减少容易分类的样本对于损失的贡献,增大难以分类的样本对于损失的贡献。由于调制因子 γ 的作用,模型减弱了易识别样本的错误贡献,这样模型在训练时可以更有效地集中于难以分类的样本。本文多次实验后,设置 $\gamma = 0.25$ 。

在此基础上,FL还引入平衡因子 α ,如式(2)所示。

该因子可以用来平衡正负样本本身的数量比例不均即类别不平衡. 由于自建的石墨数据集类别不平衡问题不是很突出, 故设置 $\gamma = \text{NAN}$.

$$L_{FL} = \begin{cases} -\alpha(1-y')^y \log y', & y = 1 \\ -(1-\alpha)y^y \log(1-y'), & y = 0 \end{cases} \quad (2)$$

2.4 改进的迁移网络模型

结合迁移学习, 本文对 3 种网络进行改进, 冻结模型架构的特征层, 重新设计新的输出模块载入网络的全连接层, 新的全连接层如图 2 所示.

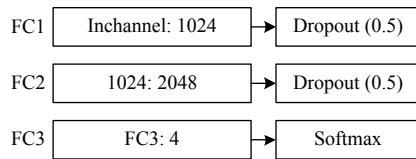


图 2 新的输出模块

在新的输出模块中, 共设置 3 层连接层, 调试连接层 FC1、FC2 的输出参数, 优化模型的稳定性与泛化能力; 由于网络的全连接层有大量参数, 因此在连接层 FC1、FC2 中均加入 Dropout, 设置随机丢失概率 $p = 0.5$, 这种技术可以减少神经元复杂的共同适应, 使模型更有效地进行融合. 改进迁移网络的结构如图 3.

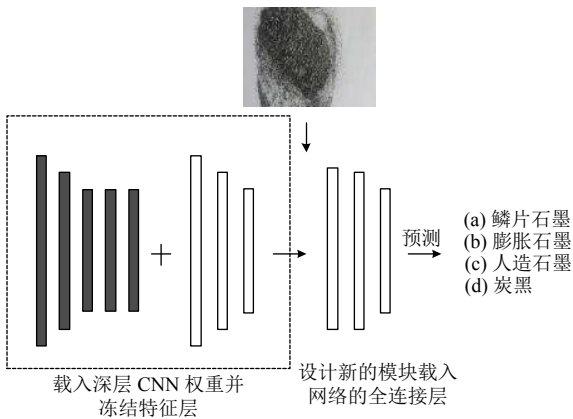


图 3 基于迁移学习的改进卷积神经网络模型

本文将此模型称为基于迁移学习的改进卷积神经网络模型, 应用于石墨分类问题, 修改优化后的网络重命名为 I-VGG16、I-ResNet34 和 I-MobileNet V2. 在上述改进基础上, 将结合 FL 损失函数的网络记为 I-VGG16 + FL、I-ResNet34 + FL 和 I-MobileNet V2 + FL.

本文设计的模型训练过程包含以下 3 个阶段: 数据增广、迁移学习、超参数调整, 模型训练过程如图 4.

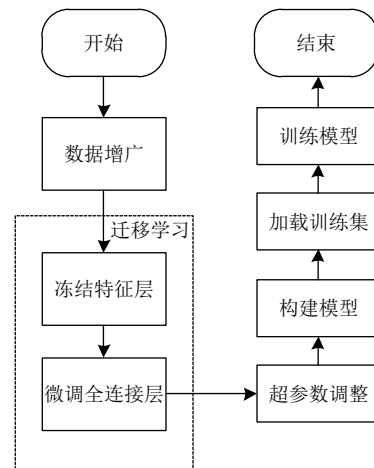


图 4 本文算法实现流程图

在载入新的输出模块前, 均添加一个全局平均池化层, 最后使用 Softmax 非线性的最终分类层, 设置最终输出为 4. 不断调试网络的超参数, 最终使用 Adam 优化器对所有模型进行 100 个 epoch 的微调, 以 0.0002 为初始学习率, Nesterov 动量设为 0.9, 批量大小为 32.

3 结果分析

3.1 评价指标与环境配置

本文以准确率 acc 、损失值 $loss$ 以及运行时间这 3 个指标评估一个网络的训练效果. 测试准确率 acc 是指模型在测试集上输出正确结果的比率, 其定义公式如式 (3) 所示:

$$acc = \frac{n_{correct}}{n} \quad (3)$$

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.

迁移学习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):

$$E = \frac{1}{n} \sum_i (\hat{y}_i^{test} - y_i^{test})^2 \quad (4)$$

本文按照整个网络模型运行结束所规定的 epoch 的时间计时. 一个 epoch 是指所有训练样本在神经网络模型中都进行了一次正向传播和一次反向传播, 即将所有训练样本训练一次的过程.

本文在 Python 3.8 的编译环境下, 使用 1.8.0 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: Windows 10 操作系统, 处

理器为 Intel(R) Core(TM) i9-10900F CPU, GPU 为 NVIDIA GeForce RTX 3060.

3.2 测试结果

为了验证本文算法的可行性与有效性, 将数据增

强后的石墨数据集导入本文所提方法中, 通过对测试集的实验, 将未改进的模型、改进后的模型与同时修改损失函数的网络模型进行对比, 最后得到模型损失与准确率的对比曲线如图 5、图 6 和图 7 所示.

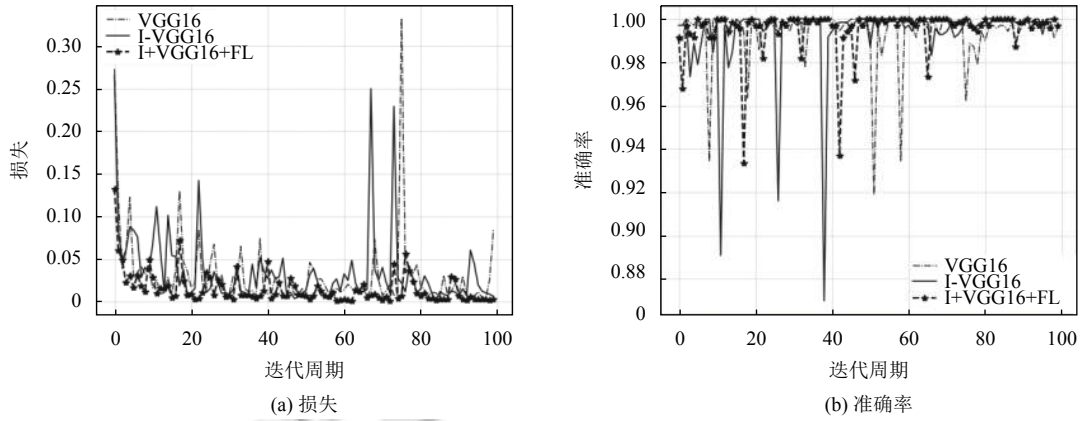


图 5 VGG16 模型损失与准确率对比图

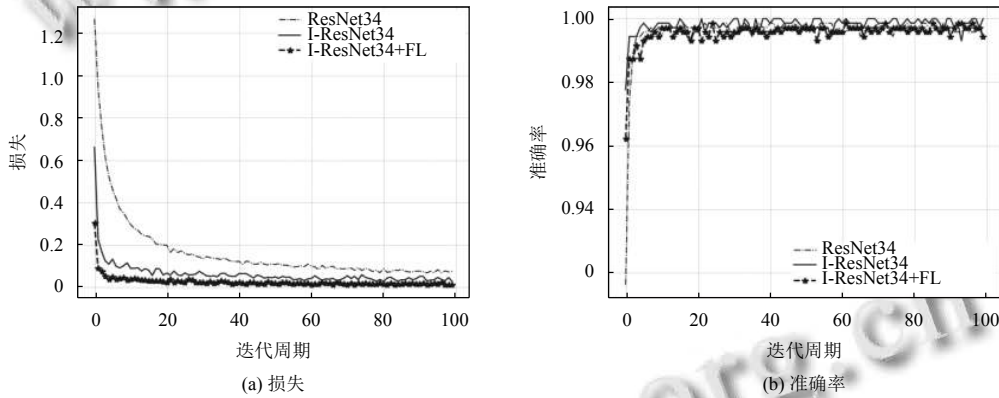


图 6 ResNet34 模型损失与准确率对比图

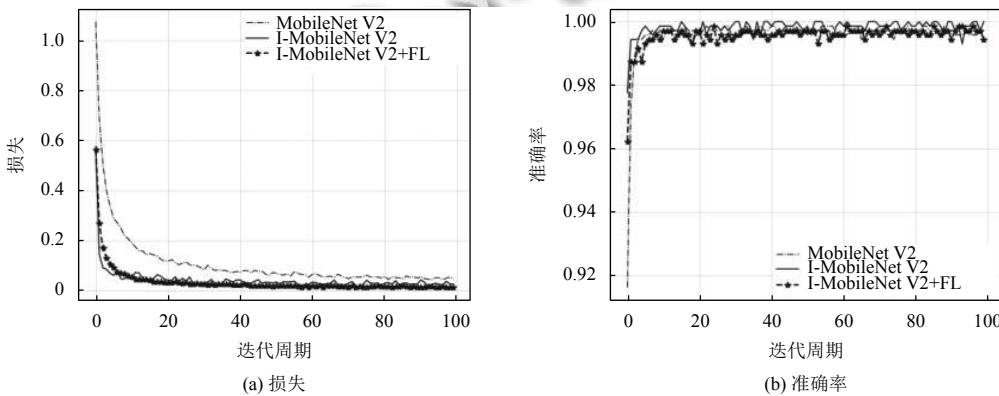


图 7 MobileNet V2 模型损失与准确率对比图

如图 5 所示, VGG16 有较大的 Loss 值震动, 模型的稳定性较差, 在对模型进行改进后, I-VGG16 与 I-

VGG16 + FL 网络性能明显更稳定、更优; 相比 VGG16, ResNet34 与 MobileNet V2 网络更深, 模型更加稳定,

可以看出改进后的网络有效地缓解了由于迁移学习的二次应用导致的提取特征能力差的问题,使用FL后进一步优化了模型,损失值均逐步降低,准确度得到了有效的提升,如图6、图7所示.最终实验结果分析如表3所示.

由表3可知,VGG16虽然准确率及训练时间相对较差,但损失达到相对最小;整体来看,ResNet34的训练时间相对较短,优化后的模型准确率也有所提高;MobileNet V2作为轻量级网络,识别准确率均能达到99%以上,且随着损失函数的改进损失值逐步减小,均证明了结合FL与改进的模型能有效侧重石墨数据集的难样本图像,说明了卷积神经网络在石墨分类中的可行性以及有效性.结合FL的改进CNN模型的损失

值与准确率对比图如图8所示,直观地展示了使用本文改进后的模型更加稳定,均能在几个epoch内快速收敛,准确率在93%内浮动,证明了改进算法的有效性.

表3 不同算法模型的损失值、准确率与训练时间对比

模型	损失值($\times 10^{-2}$)	准确率(%)	训练时间(min)
VGG16	0.36	92.93	74.73
I-VGG16	0.16	93.71	76.13
I-VGG16 + FL	0.10	95.69	75.88
ResNet34	6.75	93.29	40.13
I-ResNet34	2.61	99.18	40.72
I-ResNet34 + FL	0.97	99.84	41.27
MobileNet V2	4.37	98.57	39.10
I-MobileNet V2	1.68	99.81	45.18
I-MobileNet V2 + FL	1.23	99.57	46.82

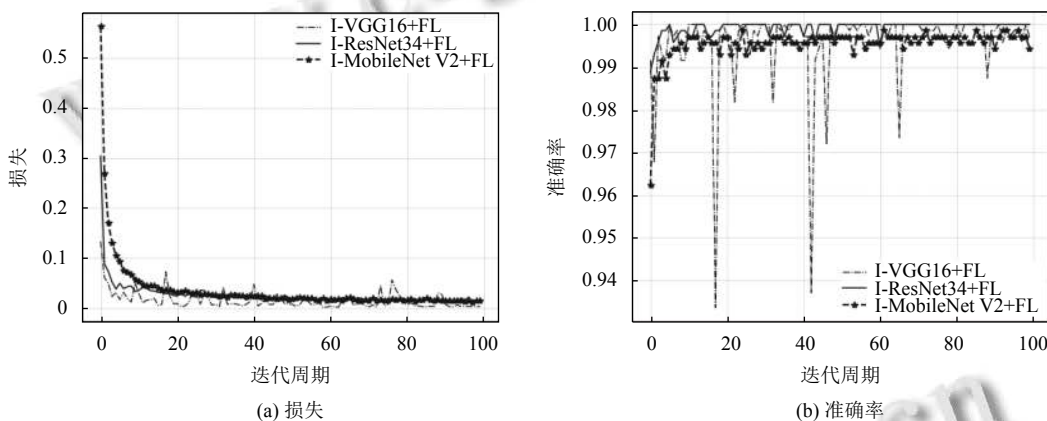


图8 结合FL的改进CNN模型对比图

4 结束语

本文结合Focal Loss,将改进的CNN应用于石墨图像的分类识别中.基于CNN建立的识别模型前期并不需要繁杂的图像预处理步骤,数据增强也可以通过训练过程中在线完成,模型建立过程更加简洁.对石墨数据集进行数据扩充和增强,解决了石墨数据集存在的训练样本不足的问题.对处理后的数据集进行仿真实验,得出重新设计的全连接层有助于提高网络的分类性能的结论.对损失函数进行优化后的模型,通过对模型超参数的调试,其网络模型的性能得到进一步改进,实现了石墨分类的相对较高准确率与较低损失值.在未来的工作中,将在已实现的3种CNN以外进行其他深层网络的尝试,因此探讨不同冻结比例、冻结层数对石墨数据集分类效果的影响是我们未来的重点研究内容.

参考文献

- 1 颜玲亚,陈正国,周雯,等.石墨作为战略性矿产的政策分析及建议.中国非金属矿工业导刊,2019,(S1):7-12.
- 2 焦丽香,高树学,郭加朋,等.我国石墨资源开发利用现状分析.中国非金属矿工业导刊,2019,(S1):1-6.
- 3 李悦,张顺艳.高纯石墨生产工艺现状及展望.科学技术创新,2018,(1):166-167.
- 4 Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278-2324. [doi: 10.1109/5.726791]
- 5 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th Annual Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012. 1097-1105.
- 6 Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. Proceedings

- of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 815–823. [doi: [10.1109/CVPR.2015.7298682](https://doi.org/10.1109/CVPR.2015.7298682)]
- 7 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. arXiv:1512.03385, 2015.
- 8 Donahue J, Jia YQ, Vinyals O, *et al.* DeCAF: A deep convolutional activation feature for generic visual recognition. <http://proceedings.mlr.press/v32/donahue14.pdf>. (2020-10-23).
- 9 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2015.
- 10 Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 580–587.
- 11 Lu J, Behbood V, Hao P, *et al.* Transfer learning using computational intelligence: A survey. Knowledge-Based Systems, 2015, 80: 14–23. [doi: [10.1016/j.knsys.2015.01.010](https://doi.org/10.1016/j.knsys.2015.01.010)]
- 12 Tajbakhsh N, Shin JY, Gurudu SR, *et al.* Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 2016, 35(5): 1299–1312. [doi: [10.1109/TMI.2016.2535302](https://doi.org/10.1109/TMI.2016.2535302)]
- 13 Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Neural Networks. European Conference on Computer Vision. Cham: Springer International Publishing, 2014. 818–833.
- 14 Razavian AS, Azizpour H, Sullivan J, *et al.* CNN features off-the-shelf: An astounding baseline for recognition. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus: IEEE, 2014. 512–519.
- 15 Azizpour H, Razavian AS, Sullivan J, *et al.* From Generic to Specific Deep Representations for Visual Recognition. DeepVision workshop at CVPR 2015. IEEE, 2015. 1–1.
- 16 Sandler M, Howard A, Zhu M L, *et al.* MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018. 4510–4520.