

基于卷积神经网络岩石样本识别方法

陈洪建¹, 向滔¹, 柳春源¹, 宋先璐², 寇喜鹏¹

¹重庆科技学院数理与大数据学院, 重庆

²西南石油大学电气信息学院, 四川 成都

收稿日期: 2022年6月18日; 录用日期: 2022年7月15日; 发布日期: 2022年7月25日

摘要

岩石种类繁多, 人类现已发现的岩石种类有好多种。从现实意义来说, 岩石样本识别在油气勘探、水资源勘探、矿物勘探、能源勘探、工程建设中是一项既基础又重要的环节。近年来, 随着人工智能技术的发展, 采用图像深度学习的方法建立岩石样本自动识别分类模型成为主导, 相较于传统岩石分类减少了设备影响与人为因素干扰, 更为快速且准确地识别。本研究数据集来自Big Data Mining Race dataset2021, 针对该数据集存在岩石种类不均匀、图片大小不一且数据量少的问题, 通过图片裁剪、旋转、饱和度调整等图片增强手段, 上述问题得以缓解。本研究针对岩石剖面, 结合迁移学习, 通过冻结与微调预训练模型的部分层, 分别建立基于InceptionV3网络、ResNet50网络、VGG16网络的逐步细分串联模型与单个模型, 通过准确率、召回率等值判断模型的好坏, 最终得出结论。

关键词

岩石识别, 卷积神经网络, ResNet50, InceptionV3, VGG16

Rock Sample Recognition Method Based on Convolution Neural Network

Hongjian Chen¹, Tao Xiang¹, Chunyuan Liu¹, Xianlu Song², Xipeng Kou¹

¹School of Mathematics and Big Data, Chongqing University of Science and Technology, Chongqing

²School of Electrical Engineering and Information, Southwest Petroleum University, Chengdu Sichuan

Received: Jun. 18th, 2022; accepted: Jul. 15th, 2022; published: Jul. 25th, 2022

Abstract

There are many kinds of rocks that have been proven by human beings. From a practical point of view, rock samples identification is a basic and important link in oil and gas exploration, water resources exploration, mineral exploration, energy exploration, and engineering construction. In recent years, with the development of artificial intelligence technology, the method of image deep

learning has become the dominant method to establish an automatic identification and classification model of rock samples. Compared with traditional rock classification, it reduces the influence of equipment and human factors, and can identify more quickly and accurately. The data set of this study is from Big Data Mining Race dataset2021. The data set has the problems of uneven rock types, different image sizes and small data volume. The problems can be solved by image enhancement methods such as image cropping, rotation, and saturation adjustment. In this paper, aiming at the rock profile, combined with transfer learning, by freezing and fine-tuning some layers of the pre-training model, a step-by-step subdivision series model and a single model based on the InceptionV3 network, ResNet50 network, and VGG16 network were established respectively. Through the equivalence of accuracy and recall rate, we can judge the quality of the model and finally draw a conclusion.

Keywords

Rock Identification, Convolutional Neural Network, ResNet50, InceptionV3, VGG16

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

我国的岩性识别的研究相对较晚,起源于20世纪后期,而国外较早,起源于20世纪初。传统的岩性识别方法主要经历了从地质填图到地球物理、岩石薄片分析等的过程[1],但这些方法往往受实验设备、工作环境与研究经验的限制,十分依赖分类人的经验和设备精确度,造成了低识别率、低效率等缺点[2]。目前,地质灾害、矿物勘探等系列工作产生了巨大的数据量,为克服传统岩性识别带来的低识别率、低效率、外界条件影响大等问题,将地质科学与智能算法结合成为一种新热潮,岩石图像分类开始朝着计算机自动化分类的方向发展,采用图像深度学习的方法建立岩石样本自动识别分类模型是一条新的途径。

近年来,随着人工智能技术的发展,采用图像深度学习的方法建立岩石样本自动识别分类模型成为主导,相较于传统岩石分类减少了设备影响与人为因素,更为快速且准确地识别。本研究采用深度学习技术与迁移学习相结合的方法,通过冻结在庞大数据集下训练出的模型的部分层并微调调整部分层的方式[3],分别建立基于InceptionV3网络、ResNet50网络、VGG16网络的逐步细分串联模型与单个模型,进行岩石的智能识别与分类。

2. 研究资料

对于岩石图像数据集来说,岩石类型多样、非均质性、某些岩石种类较高的相似性以及存在环境对岩石外观的破坏,限制了岩石的图像采集和识别。最初,支持向量机因克服高维表示缺陷的优点,被广泛运用到岩石分类中[4]。虽然机器学习方法已经取得了很好的效果,但是岩石特征的提取仍然是手工进行,且数据集多为标准的岩石薄片少杂质,造成图像数据集不易获取、无法应用于工程实际等问题[5]。利用卷积神经网络提取岩石特征、利用图像处理技术处理背景噪声,开启了岩石识别与分类的新大门。卷积神经网络(CNN)卷积神经网络能够自动从海量数据中提取出特征,复杂的分类或预测问题则可由逐层的特征变化来解决[6]。近五年来,大量的关于利用卷积神经网络的研究涌现:2018年,张野等基于InceptionV3网络,对三种岩石进行分类[7],2021年,许振浩等利用迁移学习,建立了的ResNet模型,准确率在90%以上[8]。但由于卷积神经网络在特征的自动提取过程中可能会造成特征冗余,识别到过多噪声,研究人员也会选择将特征工程与卷积神经网络结合研究[9][10],基于随机森林、支持向量机、人工神经网络的岩性识别研究

依然在进行中[11], Ren 等人运用神经网络进行岩石识别[12], 2020 年 Wang 等人建立了基于隐式马尔可夫模型与随机森林的岩性识别模型[13], LI 等人采用拉普拉斯支持向量机实现半监督学习[14]。

本研究数据集来自 Big Data Mining Race dataset2021, 该数据集是采用工业相机在录井现场对岩屑和岩心样品进行拍照, 拍摄的相片是用于提取颜色、纹理、粒度等特征识别岩性。数据集共有共两种图片格式: bmp 与 jpg 格式, bmp 格式图片大小统一为 4096×3000 , jpg 格式图片大小统一为 2448×2048 。本研究是利用深度学习方法对岩石样本进行识别分类。

3. 研究方法

3.1. 数据集分析

数据集中分别包含荧光图片与自然光图片 315 张, 分为黑色煤、灰黑色泥岩、灰色泥质粉砂岩、灰色细砂岩、浅灰色细砂岩、深灰色粉砂质泥岩和深灰色泥岩七类, 如图 1 所示。各类别数据数目不均衡, 如图 2 所示, 图片最多的类别为浅灰色细砂岩, 共 85 张, 占比 26.98%, 最少的类别为灰色细砂岩, 仅 18 张, 占 5.72%。

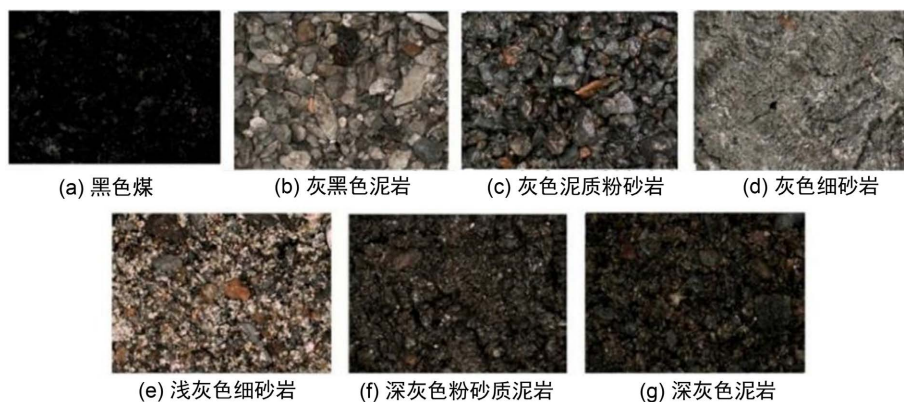


Figure 1. Seven rock sample images

图 1. 七种岩石样本图像

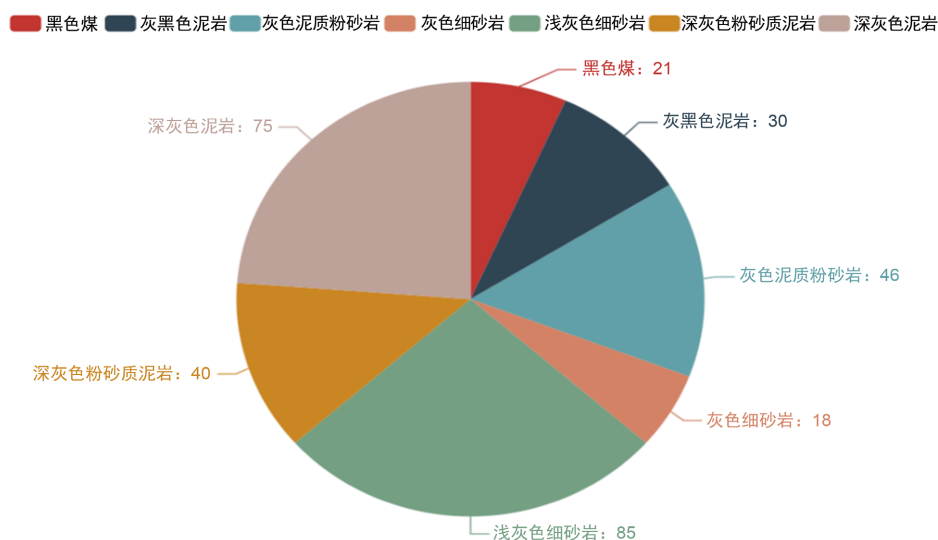


Figure 2. Number of seven rock samples

图 2. 七种岩石样本数量

岩石样本图片的分辨率有 4096×3000 (bmp 格式)与 2448×2048 (jpg 格式)两种,如图 3 所示。在 2448×2048 分辨率的图片中,矿石部分处在图片中间,四周含有背景。

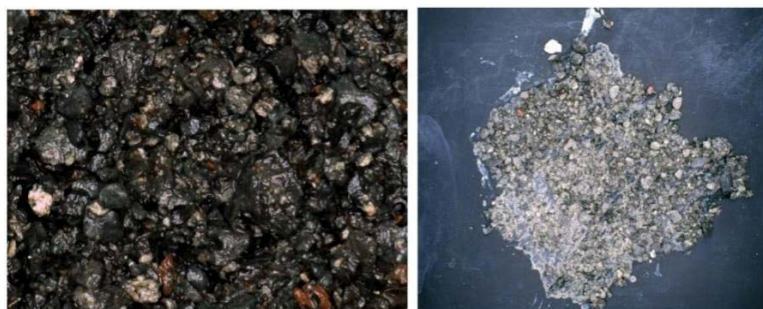


Figure 3. Rock samples with two different resolutions
图 3. 两种不同分辨率的岩石样本

该数据集中,受光照、潮湿度等的影响,存在不同种类岩石间图片太相似的问题,如图 4 所示。

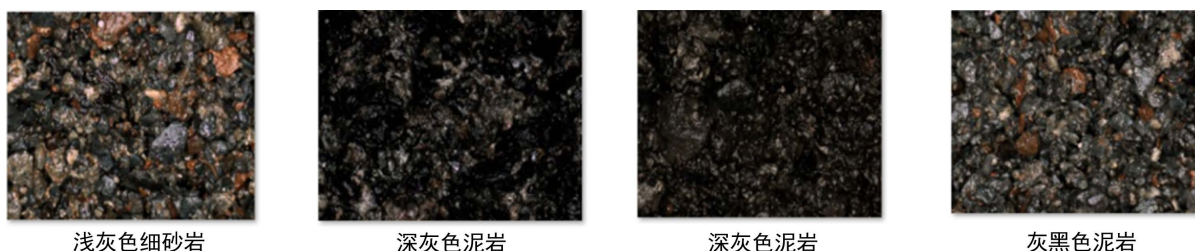


Figure 4. Similar image display
图 4. 相似图像展示

3.2. 图像预处理

同一类岩石可能因光照、潮湿程度等环境因素改变颜色、纹理。为增强模型的泛化能力,选择图像旋转、随机饱和度调节、随机亮度调节、添加高斯扰动、仿射变换五种基础增强方案。

3.2.1. 图像切割

由于原数据集图片数量较少,对图片进行训练提取特征数量不够,会致使模型过拟合,泛化能力不够。首先将图片进行切割处理,增加图片数据集,让模型能够有足够的数据进行训练。由上文可知,原数据集共两种图片格式: bmp 格式与 jpg 格式,对应两种不同图片大小与不同图片形式,两种不同图片形式如图 3 所示。分析上述两种图片格式,提出两种不同的裁剪方案。针对 jpg 格式,其像素大小为 2448×2048 ,四周背景明显,且岩石占比不大,因此选择将 jpg 图片底色部分舍去,定位 jpg 格式图片的中心点,切割出一块大小为 500×500 的图片,减少图片背景对训练的影响,如图 5 所示。针对 bmp 格式,其大小为 4096×3000 ,无明显背景,可直接对图片进行裁剪,遍历图片像素点,将一张原图分割成 48 张 500×500 大小的图片,如图 6 所示。此次采用 OpenCV-python (即 CV2, 实现了图像处理和计算机视觉方面的很多通用算法),对图片数据集分割。

3.2.2. 图像旋转

对原数据集完成图片切割后,统计发现,得到的数据集仍是极不均匀的,样本的数据量增大,但数据集不均匀的问题仍需解决。选用图片旋转改善上述问题。

为保证图片格式相同，500 × 500 的大小不变，并且图片内不留白，选择旋转 90°、180°、270° 三个角度。因此，当数量最多的岩石类别大于数据量最小的岩石类别的三倍时，选择让每张图片在三个角度内旋转，使其扩大三倍，尽量让数据集均匀。

对切割后每个岩石类别的数量统计后，以数量最多的类别为基准，其他岩石类别根据需扩充数量做随机旋转，使每个岩石类别数量均匀。旋转后各类岩石图片数量如图 7 所示。

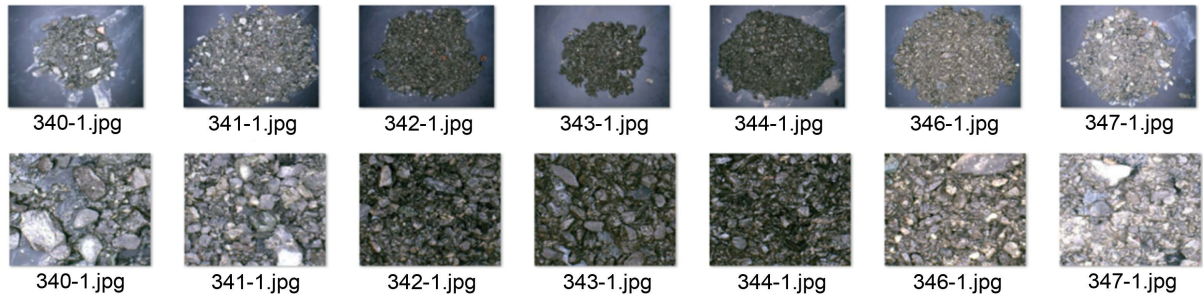


Figure 5. JPG format image cutting
图 5. JPG 格式图像切割

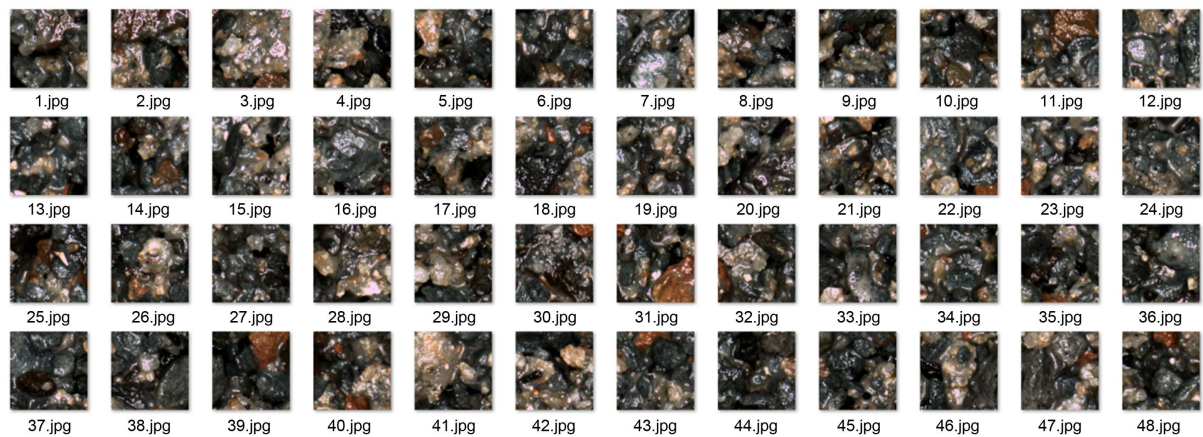


Figure 6. BMP format image cutting
图 6. BMP 格式图像切割

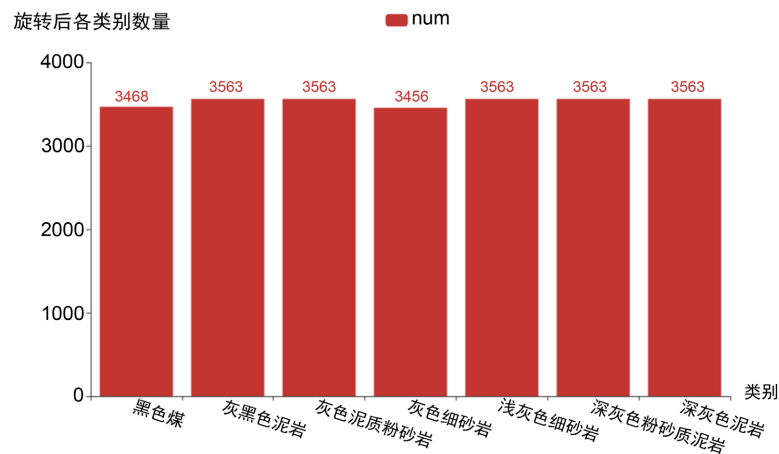


Figure 7. Comparison of the number of rock samples after rotation
图 7. 旋转后各岩石样本数量对比

3.2.3. 饱和度、明亮度调节

为更有利于串联逐步细分模型的训练，泥岩类与砂岩类图片各有 10,689 张，黑色煤共 3468 张，为保证该三类数据集均匀，采用随机饱和度与明亮度改变，增加黑色煤数量。如图 8 所示，为经过图片翻转、明亮度与饱和度调节后的效果。

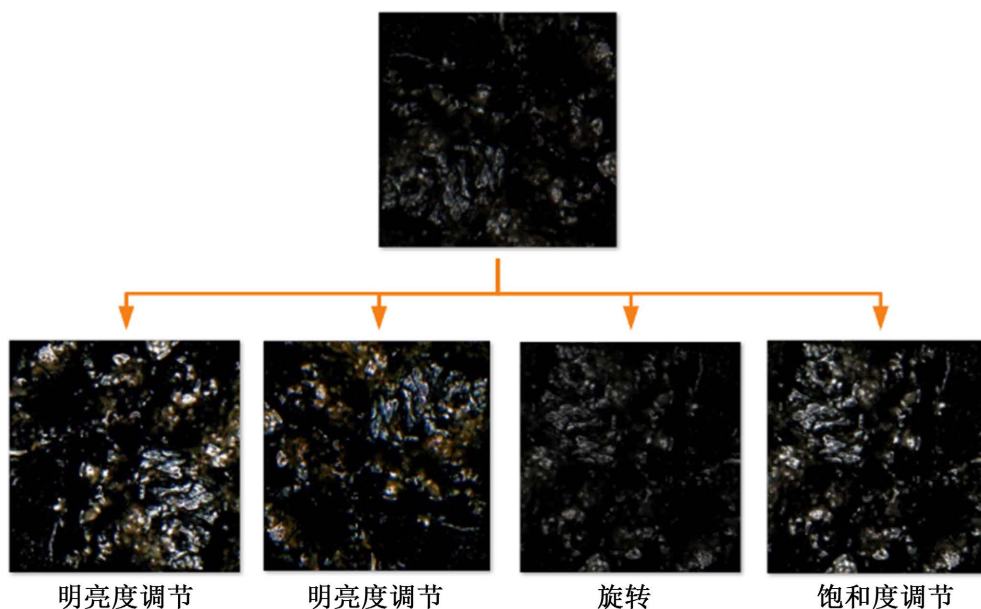


Figure 8. Flipping, brightness and saturation adjustment
图 8. 旋转、明亮度及饱和度调节

3.2.4. 高斯扰动、防射变换

Python 中有封装好了的 `imgaug` 库，用作图像增强，本次设计中，导入 `imgaug` 库，利用 `GaussianBlur` 函数为图像增添高斯扰动、`Affine` 函数为图像进行仿射变换、`Flipud` 函数垂直翻转输入图像。将以上三种变换放入 `Sequential` 列表，对输入的每张图片随机选取变换进行图像增强。图 9 为一张图片进行的随机处理：

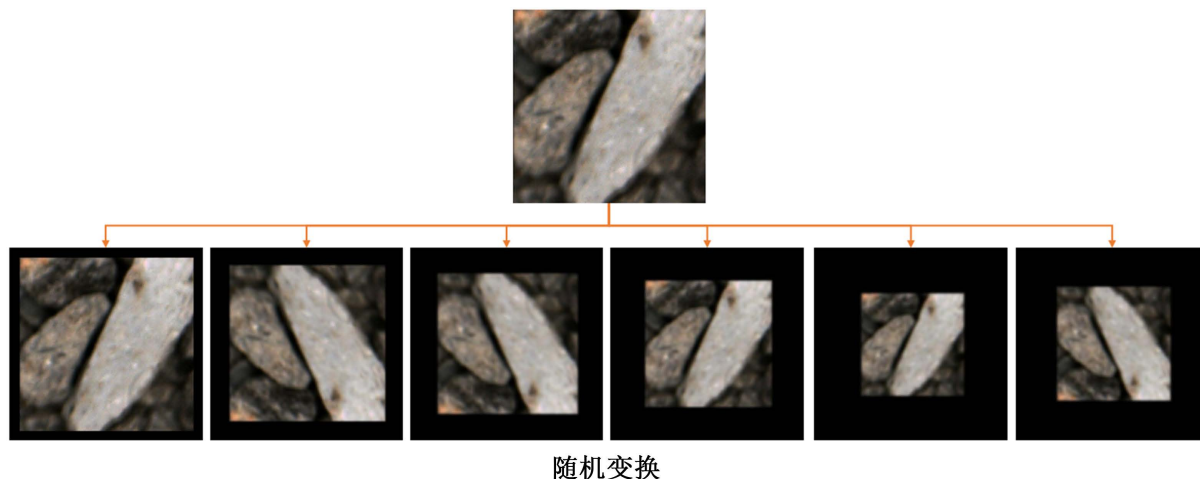


Figure 9. Random transformation process
图 9. 随机变换过程

3.3. 模型构建

本次设计共两种类型的模型：单个模型与逐步细化模型。将增强后的图片，转换成张量形式，并进行归一化处理。图片打乱后传入预训练模型进行迁移学习的训练。

3.3.1. 逐步细化模型

该数据集共七类岩石，根据岩石类别，可将岩石分为黑色煤、泥岩、砂岩三大类。黑色煤类包含一小类，泥岩类包含三小类：深灰色泥岩、深灰色粉砂质泥岩与灰黑色泥岩，砂岩类共包含三小类：灰色泥质粉砂岩、灰色细砂岩与浅灰色细砂岩。

逐步细化模型的主要思想就是将数据集先分为三大类，而后再在三大类中逐步细分，意味着它将运用到三个子模型，三个子模型分别对应三个不同的任务，如图 10 所示。

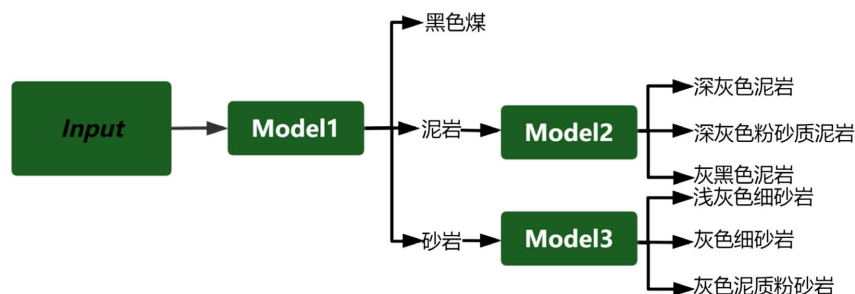


Figure 10. Series model

图 10. 串联模型

训练模型一只需对黑色煤、泥岩、砂岩三大类的全局特征有整体把握，如颜色，对该七类岩石进行三大类的分类。模型二与模型三需分别学习泥岩和砂岩的局部特征——纹理特征，才能更好地区分差异不大的小类。而后将上述三个模型串联，实现全局信息与局部信息的串联，如图 11 所示。

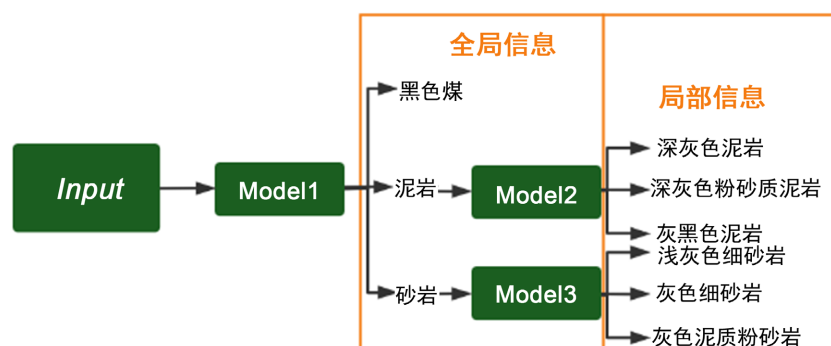


Figure 11. Series model

图 11. 串联模型

因此需要分别训练三个以 ResNet50 为主干网络的模型，实现三种不同的分类，而后将三个模型组合以实现逐步识别的功能。

模型一用到的数据集，要求对原始数据图中的七种岩石图片进行裁剪、翻转、明亮度调节、饱和度调节，以实现黑色煤、泥岩、砂岩图片数量的均匀，而后将得到的数据集打乱，再按照 6:2:2 的比例，划分训练集、验证集、测试集，同时保证三类数据集图片互不冲突。最终，划分出的训练集中：黑色煤共 6768 张，泥岩共 6771 张，砂岩共 6771 张；验证集中：黑色煤共 2258 张，泥岩共 2259 张，砂岩共 2259

张；测试集中：七类岩石的数量各 753 张。根据每张图片的类别添加数据标签，按照黑色煤 = 0、泥岩 = 1、砂岩 = 2 的映射关系对每张图片数据标签进行一一映射，使数据标签数值化。得到的训练集与验证集数据还需进行打乱、数据增强、归一化等处理，才能传入模型中进行训练。训练时，只会用到训练集进行模型权重的更新，验证集只用于评测每次权重更新后的模型的准确率。测试集用于最终模型的评测。

传入模型二的数据集：数据增强后泥岩类数据集按照 8:2 的比例划分，同时保证深灰色泥岩、深灰色粉砂质泥岩、灰黑色泥岩三类在训练集与验证集中的数量均匀。传入模型三的数据集：利用数据增强后砂岩类数据集按照 8:2 的比例划分，保证浅灰色细砂岩、灰色细砂岩、灰色泥质粉砂岩三类在训练集与验证集中的数量均匀。模型二与模型三用到的训练集不能与测试集有包含关系。最终，模型二用到的训练集：三类泥岩数量各 2408 张；验证集：三类泥岩的数量各 602 张。根据每张图片的类别添加数据标签，按照灰黑色泥岩 = 0、深灰色粉砂质泥岩 = 1、深灰色泥岩 = 2 的映射关系对每张图片数据标签进行一一映射，使数据标签数值化；模型三用到的训练集：三类砂岩数量各 2408 张；验证集：三类砂岩的数量各 602 张。根据每张图片的类别添加数据标签，按照灰色泥质粉砂岩 = 0、灰色细砂岩 = 1、浅灰色细砂岩 = 2 的映射关系对每张图片数据标签进行一一映射，使数据标签数值化。

3.3.2. 单个模型

单个模型，就是将该七类岩石，运用一个模型，提取岩石的各类特征，一次性完成七种岩石的分类任务。区别于串联模型的数据集划分，此处需保证七种岩石图片数量均匀，如图 12 所示。

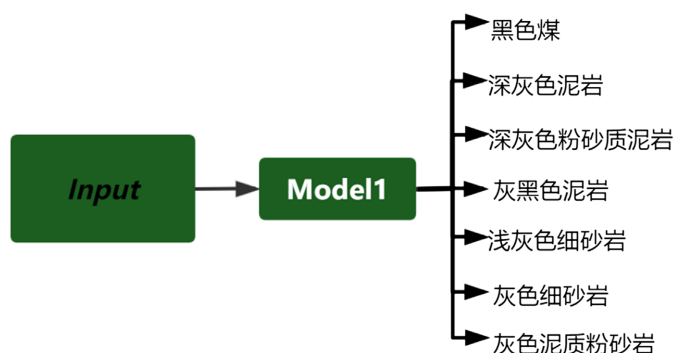


Figure 12. Single model

图 12. 单个模型

将七类岩石按照 6:2:2 的比例，划分训练集、测试集、验证集。最终，划分出的训练集中：黑色煤共 2080 张，灰色细砂岩共 2073 张，其他种类岩石各 2137 张；验证集中：黑色煤共 694 张，灰色细砂岩共 691 张，其他种类岩石各 713 张；测试集中：黑色煤共 694 张，灰色细砂岩共 691 张，其他种类岩石各 713 张。根据每张图片的类别添加数据标签，按照黑色煤 = 0、灰黑色泥岩 = 1、灰色泥质粉砂岩 = 2、灰色细砂岩 = 3、浅灰色细砂岩 = 4、深灰色粉砂质泥岩 = 5、深灰色泥岩 = 6 的映射关系对每张图片数据标签进行一一映射，使数据标签数值化。

3.4. 模型调整

为提高模型的准确率，采用改变优化器、改变模型、改变学习率、改变 batch_size、改变激活函数等方法。

3.4.1. 调整优化器

损失函数的是否收敛关系着模型拟合的好坏。本次设计中，共试用了两种优化器以使损失函数收敛：

SGD 优化器与 ADAM 优化器。以 VGG16 模型为例，固定模型结构、batch_size 为 32、激活函数为 Relu，举出一些关于调整优化器后的实验数据，如表 1 示。

Table 1. System resulting data of standard experiment
表 1. 标准试验系统结果数据

起始学习率	学习率下降	每次下降一次	迭代次数	优化器	准确率
0.001	0.5	10	30	SGD	0.712
0.001	0.1	10	30	SGD	0.730
0.01	0.1	10	30	SGD	0.7069
0.0001	0.5	10	30	SGD	0.7103
0.001	0.5	10	30	ADAM	0.6860
0.001	0.1	10	30	ADAM	0.7108
0.01	0.1	10	30	ADAM	0.7266
0.0001	0.5	10	30	ADAM	0.7041

3.4.2. 调整学习率

学习率的大小影响着学习模型的训练进度，太大的学习率模型的训练速度快，但可能会致使模型的准确率和损失函数震荡，收敛速度过快，适用于模型训练初期；过小的学习率模型的训练速度慢，可能会致使模型的收敛速度缓慢，导致过拟合，适用于模型训练后期。本文中，利用 Keras 中提供的 Learning Rate Scheduler 函数，对模型的学习率作调整，训练初期，模型的学习率为一个较大的值，当迭代一定周期后，学习率乘以一个小于 1 的数值，以将学习率降低。在此基础上，设置早停(Early Stopping)，防止模型过拟合。以 VGG16 模型为例，固定模型结构、batch_size 为 32、激活函数为 Relu，举出一些关于调整学习率后的实验数据，如表 1 示。可知，学习率不能太大也不能太小，模型的准确率都会被影响到。

3.4.3. 调整模型

本次设计采用了 ResNet50 模型、VGG16 模型、InceptionV3 模型。首次在 Keras 中采用 ResNet50 模型进行迁移学习时，模型的准确率只在 0.3 左右徘徊，不论采用何种方法添加 Dropout 层、L2 正则化、BN 层，都无法改变效果。经查阅资料后，发现利用 Keras 冻结时，ResNet50 的 BN 层并没有被冻结住，可以将 BN 层的 training 参数利用 set_learning_phase 设置。经以上调整，模型准确率从 0.3 提升到 0.7 左右。

如下表 2 所示，固定 batch_size 为 32、激活函数为 Relu、优化器为 SGD，学习率从 0.001 每 10 次迭代下降为原来的 0.1 倍，共迭代 30 次，举出一些关于调整模型的实验数据。

Table 2. The influence of the number of nodes and layers in the full connection layer on the accuracy
表 2. 全连接层的节点数和层数对准确率的影响

模型	层数	全连接层结构	准确率
ResNet50	1	Dense(1024, activation='relu') BatchNormalization()	0.7422
ResNet50	1	Dense(1024, activation='relu') BatchNormalization()	0.5932
	2	Dense(512, activation='relu') BatchNormalization()	
ResNet50	1	Dense(1024, activation='relu') BatchNormalization()	0.6105
	2	Dense(512, activation='relu') BatchNormalization()	
	3	Dense(512, activation='relu') BatchNormalization()	

Continued

InceptionV3	1	Dense(1024, activation='relu') BatchNormalization()	0.5128
InceptionV3	1	Dense(1024, activation='relu') BatchNormalization()	0.4983
	2	Dense(512, activation='relu') BatchNormalization()	
InceptionV3	1	Dense(1024, activation='relu') BatchNormalization()	0.5112
	2	Dense(512, activation='relu') BatchNormalization()	
	3	Dense(512, activation='relu') BatchNormalization()	
VGG16	1	Dense(1024, activation='relu') BatchNormalization()	0.7059
VGG16	1	Dense(1024, activation='relu') BatchNormalization()	0.7327
	2	Dense(512, activation='relu') BatchNormalization()	
VGG16	1	Dense(1024, activation='relu') BatchNormalization()	0.7400
	2	Dense(512, activation='relu') BatchNormalization()	
	3	Dense(512, activation='relu') BatchNormalization()	

3.4.4. 调整全连接层节点数和层数

全连接层的节点数与层数也会影响模型的准确率,关于它的调整也需要参考实验数据。如表 2 所示,固定 batch_size 为 32、激活函数为 Relu、优化器为 SGD,学习率从 0.001 每 10 次迭代下降为原来的 0.1 倍,共迭代 30 次,举出一些关于调整全连接层节点数和层数的实验数据。由数据可知,当全连接层数到一定值后,模型准确率不升反降。

3.4.5. 调整 batch_size

batch_size 指模型采样一次,同时传入模型中的图片数量,其大小会影响模型的正确率、训练速度。太小的 batch_size 会导致模型训练速度缓慢,被同一批次的某个样本带偏;太大的 batch_size 会降低模型的泛化能力。当验证集中的损失函数与准确率震荡明显时,可以将超参数 batch_size 调大。如下表 3 所示,固定学习率从 0.001 每 10 次迭代下降为原来的 0.1 倍,共迭代 30 次,以 VGG16 模型为例,举出一些关于调整 batch_size 后的实验数据。

Table 3. Influence of activation function, optimizer and batch_size on accuracy

表 3. 激活函数、优化器与 batch_size 对准确率的影响

Batch_size	激活函数	起始学习率	迭代次数	优化器	准确率
32	relu	0.001	30	SGD	0.7014
32	tanh	0.001	30	SGD	0.730
64	relu	0.001	30	SGD	0.7098
64	tanh	0.001	30	SGD	0.7212
32	relu	0.001	30	ADAM	0.6866
32	tanh	0.001	30	ADAM	0.7176
64	relu	0.001	30	ADAM	0.6858
64	tanh	0.001	30	ADAM	0.7108

3.4.6. 调整激活函数

本次设计中共采用两种激活函数: Relu 激活函数与 Tanh 激活函数。如表 3 所示, 固定学习率从 0.001 每 10 次迭代下降为原来的 0.1 倍, 共迭代 30 次, 以 VGG16 模型为例, 举出一些关于调整激活函数后的实验数据。

4. 模型结果与分析

4.1. 实验环境

本项目在内存为 64G, CPU 为 i9-9900KF, GPU 为 NVIDIA GeForce RTX 2080 的设备上进行模型的训练, 使用 Keras 2.3.1, Opencv-python 4.4.4.64, Python 3.7.12 的训练环境。

4.1.1. 模型检验

本次模型检验采用精确率、准确率、F1 值、召回率[15]。先了解混淆矩阵, 混淆矩阵中包含四个概念: TP (实际为正样本也被预测为正样本的数量)、FP (实际为负样本但被预测为正样本的数量)、TN (实际为负样本也被预测为负样本的数量)、FN (实际为正样本但被预测为负样本的数量), 如图 13 所示。

混淆矩阵		真实值	
		P	N
预测值	P'	TP	FP
	N'	FN	TN

Figure 13. Confusion matrix

图 13. 混淆矩阵

图 13 中, P 代表正样本, N 代表负样本。TP + FP = P', 即实际为正样本也被预测为正样本的数量加上实际为负样本但被预测为正样本的数量等于被预测为正样本的数量; FN + TN = N', 即实际为正样本但被预测为负样本的数量加上实际为负样本也被预测为负样本的数量等于被预测为负样本的数量。本次设计检验时可当作简单的二分类模型。

精确率是指预测到正样本个数中实有正样本个数的百分比, 反映分类器不标负样本到正样本个数的能力, 其计算公式如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

准确率是指正确分类的样本量占总样本量的比例, 其计算公式如下:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

召回率是指实际正样本在预测正样本中所占的比例, 反映分类器发现全部正样本的能力, 其计算公式如下:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

F1 值(F1 Score), 它是精确率与召回率加权平均值, 反映了该模型的稳健性, 其计算公式为:

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

4.1.2. 串联模型

先把要分类的照片做一个切割处理，若分类照片为 bmp 格式则把它切割成 500×500 大小的照片，见图 14，把切割好的照片转换成矩阵，归一化。然后引入该模型对其进行识别分类，并对所有结果进行比对，得到一个最优的识别结果。若识别图像为 jpg 格式时，首先确定其图像的中心点位置，然后取出中心点周围 500×500 尺寸的图像(见图 15)，将裁剪好的图像变换成矩阵，归一化处理，然后传入模型进行识别，得出识别结果，见图 16，识别结果会以文本的形式显示在原图上。

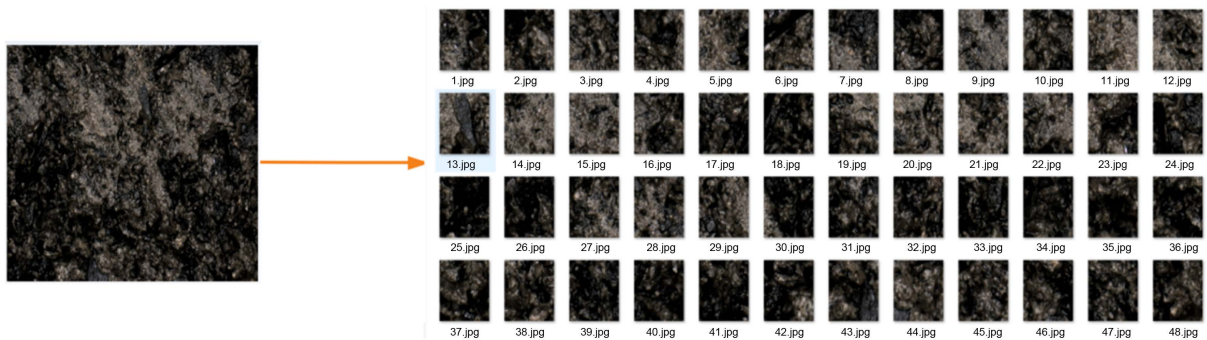


Figure 14. BMP cut image
图 14. BMP 切割图像

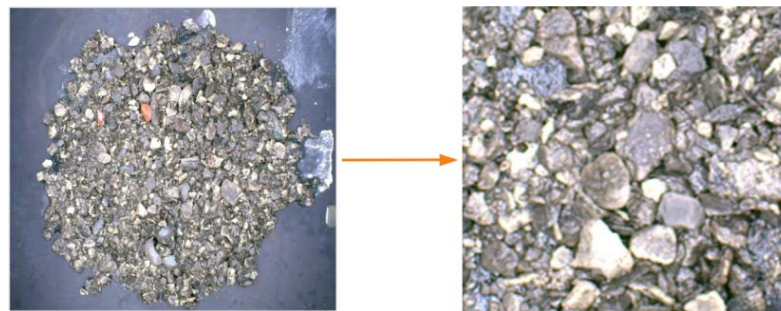


Figure 15. JPG cut image
图 15. JPG 切割图像

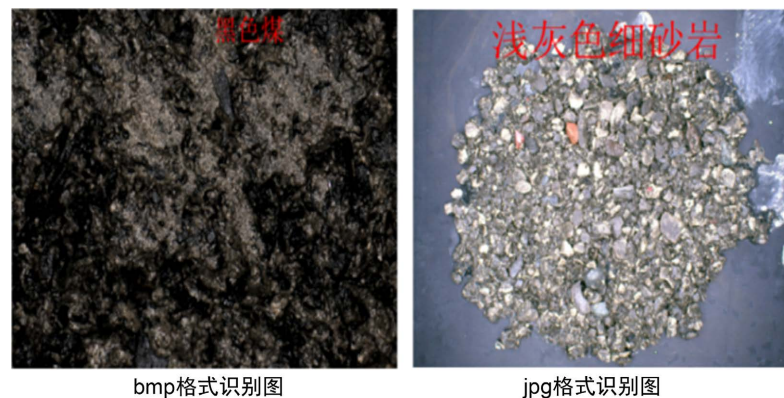


Figure 16. Identification results
图 16. 识别结果

本次设计的串联模型中,模型一采用 ResNet50 网络用以区分黑色煤、泥岩、砂岩三大类岩石,修改全连接层,增加了三个隐藏层;模型二与模型三采用 VGG16 网络用以区分除黑色煤外的其他六小类岩石,修改全连接层,增加了两个隐藏层。将该多分类问题看作二分类问题,在此串联模型的基础上,各类岩石的精确率、准确率、F1 值、召回率结果如表 4 所示。

Table 4. Evaluation of all kinds of rocks by series progressive subdivision model

表 4. 串联逐步细分模型各类岩石评价

岩石种类	精确率	准确率	召回率	F1 值
黑色煤	0.1468	0.7505	0.1556	0.1511
灰黑色泥岩	0.3098	0.7968	0.3440	0.3260
灰色泥质粉砂岩	0.3909	0.8471	0.1262	0.1908
灰色细砂岩	0.3156	0.8414	0.0943	0.1452
浅灰色细砂岩	0.2965	0.8397	0.0890	0.1369
深灰色粉砂质泥岩	0.2667	0.8465	0.0425	0.0733
深灰色泥岩	0.2989	0.8372	0.1036	0.1538

4.1.3. 单个模型

先把要识别的照片做一个切割处理,若识别照片为 bmp 格式则把它切割成 500×500 大小的照片,见图 17,把切割好的照片转换成矩阵,归一化。然后引入该模型对其进行识别,并对所有识别结果进行比对,得到一个最优的识别结果。若识别图像为 jpg 格式时,首先确定其图像的中心点位置,然后取出中心点周围 500×500 尺寸的图像(见图 18),将裁剪好的图像转换成矩阵,归一化处理,然后传入模型进行识别,得出识别结果,见图 19,识别结果会以文本的形式显示在原图上。

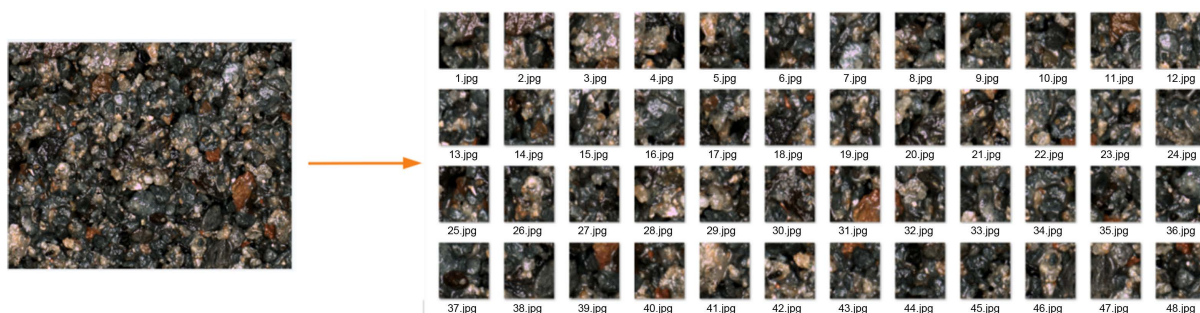


Figure 17. BMP cut image

图 17. BMP 切割图像

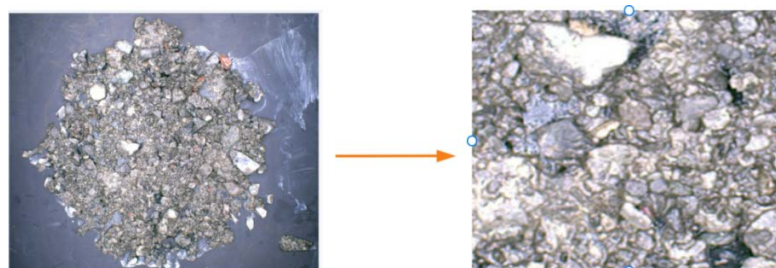


Figure 18. JPG cut image

图 18. JPG 切割图像

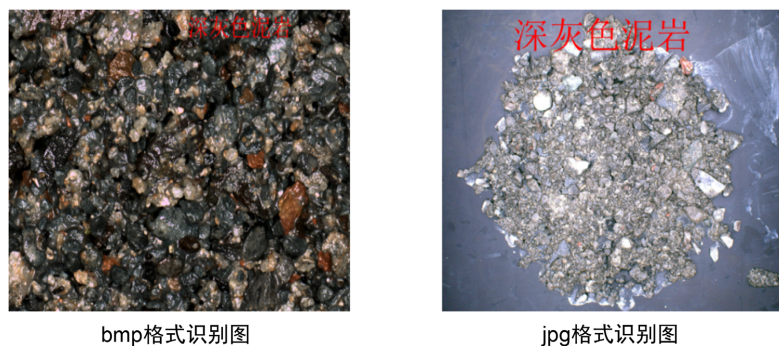


Figure 19. Identification results
图 19. 识别结果

将该多分类问题看作二分类问题，本次设计的单个模型采用 VGG16 网络，修改全连接层，增加三个隐藏层。各类岩石的精确率、准确率、F1 值、召回率结果如表 5 所示。

Table 5. Evaluation of various rocks in a single model
表 5. 单个模型各类岩石评价

岩石种类	精确率	准确率	召回率	F1 值
黑色煤	0.1291	0.7461	0.1412	0.1349
灰黑色泥岩	0.1386	0.7485	0.1431	0.1408
灰色泥质粉砂岩	0.1534	0.7538	0.1571	0.1552
灰色细砂岩	0.2932	0.6524	0.1594	0.2065
浅灰色细砂岩	0.1528	0.7651	0.1388	0.1455
深灰色粉砂质泥岩	0.1436	0.7437	0.1571	0.1500
深灰色泥岩	0.1217	0.7744	0.0912	0.1043

4.1.4. 模型泛化性检验

针对于模型的泛化性检测，训练的图片分为 bmp 和 jpg 格式，并且拍摄环境是一样的，但是在实际生活中，将会是不同的拍摄环境，在不同的拍摄环境下模型是否具有可用性呢。如图 20 所示，两幅图是从国家岩矿化石标本资源共享平台下载的两幅图片，类别都是灰黑色泥岩，但为不同的岩石样本，训练时的图片是一堆同类别岩石散落在一起的，但这些图片是单个岩石样本。利用训练好的模型对其进行识别，都得到了正确的结果，初步验证了模型的泛化性。

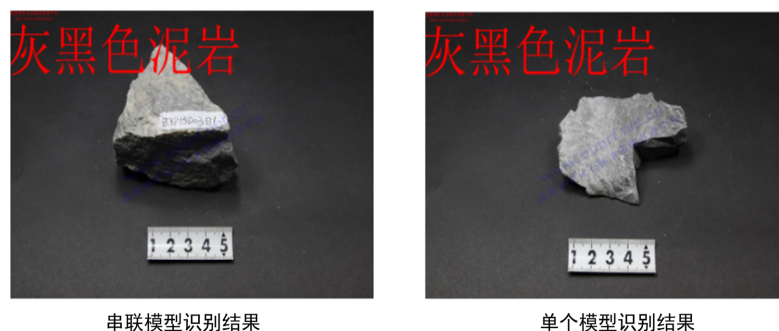


Figure 20. Model identification results
图 20. 模型识别结果

以上两张图片是专业拍摄的岩石样本，基本验证了模型泛化性，还需采用随意拍摄的图片评判模型是否具有泛化性。如图 21 所示，两张图片是随意找寻的黑色煤图片，可以看出图片都是随意拍摄的，截取出岩石样本部分用来评判模型，得到了正确的结果，识别结果以文本的形式显示在图上，判定模型有较好的泛化性。



Figure 21. Model identification results

图 21. 模型识别结果

5. 总结

本次设计中，共使用了串联逐步细分模型与单个模型，串联逐步细分模型较单个模型性能较好，但串联模型存在：无论怎样调整超参数，随着迭代次数增加，训练集的准确率逐渐提高，但是验证集的准确率却在一个数值左右波动，过拟合越来越严重的问题。

浏览数据集时发现，数据预处理过程中，黑色煤部分岩石被切割后的图片特征不明显，在其上进行数据增强如明亮度调节、饱和度调节、添加高斯扰动后，可能给数据集带来了比较多的噪声，如图 22 所示，造成特征冗余，测试集泛化能力弱，即使增大了数据集也无法使模型的准确率提升。在以后的学习中，可以思考建立特征工程，进行特征的选择。

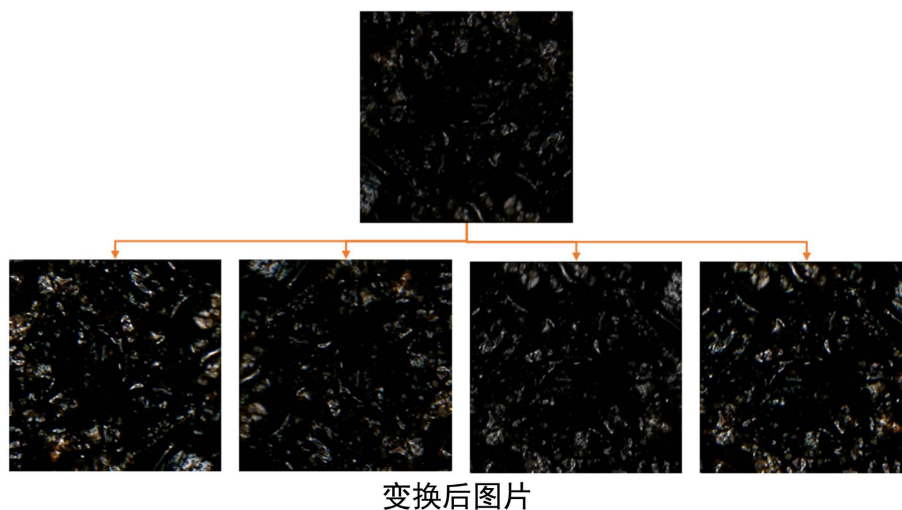


Figure 22. Noise image

图 22. 噪声图像

利用深度学习进行岩石样本智能识别，对今后的岩石考察有着重要的作用，今后发现新的岩石品种，可以采用图像识别方式进行拍照，可以识别出该新品种与其他原有品种的特征，在外太空岩石采集也同样有着类似作用。

参考文献

- [1] 付光明, 严加永, 张昆, 胡浩, 罗凡. 岩性识别技术现状与进展[J]. 地球物理学进展, 2017, 32(1): 26-40.
- [2] 刘晨, 赵晓晖, 梁乃川, 张永新. 基于 ResNet50 和迁移学习的岩性识别与分类研究[J]. 计算机与数字工程, 2021, 49(12): 2526-2530+2578.
- [3] 陈志豪. 基于深度迁移卷积神经网络的旋转机械故障诊断方法研究[D]: [硕士学位论文]. 广州: 广东技术师范大学, 2021.
- [4] 吴施楷, 曹俊兴. 基于连续限制玻尔兹曼机的支持向量机岩性识别方法[J]. 地球物理学进展, 2016, 31(2): 821-828.
- [5] Zhang, Q., Lin, G., Zhang, Y., Xu, G. and Wang, J.-J. (2018) Wildland Forest Fire Smoke Detection Based on Faster R-CNN Using Synthetic Smoke Images. *Procedia Engineering*, **211**, 441-446. <https://doi.org/10.1016/j.proeng.2017.12.034>
- [6] 安鹏, 曹丹平. 基于深度学习的测井岩性识别方法研究与应用[J]. 地球物理学进展, 2018, 33(3): 1029-1034.
- [7] 张野, 李明超, 韩帅. 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, 2018, 34(2): 333-342.
- [8] 许振浩, 马文, 林鹏, 石恒, 刘彤晖, 潘东东. 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, 2021, 29(5): 1075-1092.
- [9] Yin, Y., Alqahtani, Y., Feng, J., Chakraborty, J. and McGuire, M.P. (2021) Classification of Eye Tracking Data in Visual Information Processing Tasks Using Convolutional Neural Networks and Feature Engineering. *SN Computer Science*, **2**, Article No. 59. <https://doi.org/10.1007/s42979-020-00444-0>
- [10] Giancardo, L., Arevalo, O., Tenreiro, A., Riascos, R. and Bonfante, E. (2018) MRI Compatibility: Automatic Brain Shunt Valve Recognition Using Feature Engineering and Deep Convolutional Neural Networks. *Scientific Reports*, **8**, Article No. 16052. <https://doi.org/10.1038/s41598-018-34164-6>
- [11] 邹琪. 测井岩性识别集成模型的研究与构建[D]: [硕士学位论文]. 上海: 东华理工大学, 2020.
- [12] Ren, X., Hou, J., Song, S., Liu, Y., Chen, D., Wang, X., *et al.* (2019) Lithology Identification Using Well Logs: A Method by Integrating Artificial Neural Networks and Sedimentary Patterns. *Journal of Petroleum Science and Engineering*, **182**, Article ID: 106336.
- [13] Wang, P., Chen, X., Wang, B., Li, J. and Dai, H. (2020) An Improved Method for Lithology Identification Based on a Hidden Markov Model and Random Forests. *GEOPHYSICS*, **85**, 27-36. <http://doi.org/10.1190/geo2020-0108.1>
- [14] Li, Z., Kang, Y., Feng, D., Wang, X.-M., Lv, W., Chang, J., *et al.* (2020) Semi-Supervised Learning for Lithology Identification Using Laplacian Support Vector Machine. *Journal of Petroleum Science and Engineering*, **195**, Article ID: 107510. <https://doi.org/10.1016/j.petrol.2020.107510>
- [15] 孟寒姣. 基于知识图谱的铁路电务事故智能问答系统的设计与实现[D]: [硕士学位论文]. 石家庄: 河北科技大学, 2020.