

第九届“泰迪杯”数据挖掘挑战赛

作品单位：华南师范大学

作品成员：吴东霖 曾楚斌

指导老师：李乡儒

此页信息原作品没有

B 题：岩石样本的智能识别

基于深度学习的岩石样本岩性识别与含油 面积百分含量计算

摘要

岩石不仅是地球岩石圈圈层的重要组成元素，也是陨石、月球和火星等地外物质的重要组成元素，其研究在各个领域都具有重大意义。随着人工智能技术的不断发展，各种人工智能算法开始被广泛应用于岩石样本的自动识别与分类中，为了更高效、更准确地对岩石样本图像进行识别和分类，选择合适的深度学习模型和图像处理方法是问题的关键。因此，本文主要围绕岩石样本识别研究中的两个问题开展：岩石样本岩性识别与含油面积百分含量计算。

在岩石样本岩性识别问题上，我们针对数据样本的特性，构建了基于 ResNeSt 技术的深度学习模型，引入深度残差学习、Inception、组卷积和注意力机制等技术，在相关实测数据上的测试结果表明，在裁剪九张图像用于训练时准确率可达 97.2%，表现良好。

在含油面积百分含量计算问题上，我们使用图像转换、掩膜处理等图像处理技术计算含油面积百分含量，对于含有地板背景的特殊图像，使用语义分割 U-Net 模型进行像素分类处理，得到了一个岩石图像分割的模型。实验表明，我们的模型的 DiceCoeff 达到 94.3%。

关键词：深度学习、岩石样本识别、图像分类、语义分割

Abstract

Rocks are not only important elements of the Earth's lithosphere, but also important elements of extraterrestrial materials such as meteorites, the moon and Mars, and their research is of great significance in various fields. With the continuous development of artificial intelligence technology, various artificial intelligence algorithms have been widely used in the automatic recognition and classification of rock samples. In order to recognize and classify rock sample images more efficiently and accurately, the key to the problem is to choose an appropriate deep learning model and image processing method. Therefore, this paper focuses on two aspects in the study of rock sample identification: rock sample lithology identification and oil-bearing area calculation.

For the lithology identification of rock samples, we build a deep learning model based on Resnet technology according to the characteristics of data samples, and introduce deep residual learning, Inception, group convolution and attention mechanism, etc. The test results on relevant measured data show that the accuracy rate can reach 97.2% when nine images are clipped for training.

For the calculation of oil-bearing area, image conversion, mask processing and other image processing techniques are used to calculate the percentage of oil-bearing area. For the special image containing floor background, we use semantic segmentation U-NET model to carry out pixel classification processing, and get a rock image segmentation model. Experiments show that the DiceCoeff of our model reaches 94.3%.

Key Words: deep learning, rock sample recognition, image classification, semantic segmentation

目录

摘要	I
Abstract	II
1 引言	1
1.1 研究背景与意义	1
1.2 研究目标	2
1.3 研究内容与结构	3
2 基于 ResNeSt 构建的岩石样本岩性智能识别与分类方法研究	4
2.1 相关技术	4
2.1.1 深度残差学习	4
2.1.2 Inception	5
2.1.3 组卷积	6
2.1.4 注意力机制	7
2.2 实验过程	8
2.2.1 数据预处理	8
2.2.2 实验过程	10
2.3 实验结果	12
2.4 本章小结	12
3 基于 U-Net 构建的岩石含油面积百分含量计算模型	14
3.1 U-Net 技术	14
3.1.1 语义分割	14
3.1.2 U-Net 模型	15
3.2 实验流程	16
3.2.1 数据处理	16
3.2.2 实验环境	18
3.2.3 评价指标	18
3.3 实验结果	18
3.4 本章小结	19
4 总结	21
参考文献	22

1 引言

1.1 研究背景与意义

岩石是天然产出的矿物或者矿物与其他物质（火山玻璃、生物骨骼、胶体和岩屑等）组成的固态集合体[1]。岩石不仅构成了地球的岩石圈圈层，许多地外物质陨石、月球和火星等也是由岩石组成的。岩石的种类非常丰富，在自然界中，目前被人类探明的种类达到三千多种[2]。从科研角度来说，对岩石的研究有助于了解某个地区的地质演化历史、岩石化学组成和岩石学特征；从现实意义来说，对岩石的研究有助于寻找矿产资源、油气资源和水资源等，除此之外，在进行重大工程建设时，也都要以对岩石的研究为基础。

在油气勘探中，岩石样本识别是一项即基础又重要的环节；在矿产资源勘探中，尤其是在固体金属矿产资源勘探中，岩石样本识别同样发挥着不可估量的作用；岩石样本的识别与分类对于地质分析极为重要。因此，在实际工作中首先需要对岩石进行分类定名。目前岩石样本识别的方法主要有重磁、测井、地震、遥感、电磁、地球化学、手标本及薄片分析方法等方法，传统的岩石样本识别都是以人眼观察、手工操作和经验分类为主，是由专业人员通过专用设备从岩石图像中提取有效信息特征来分类，主要依靠的是分类人的经验和设备灵敏度，这种方法存在着无法定量分析、效率较低、受人为主观因素影响较大、专业程度比较高和难以普及等问题。为了克服这些问题，岩石图像分类开始朝着计算机自动化分类的方向发展[3]。目前大数据与人工智能算法的引入使数学地球科学实现了跨越式发展，并正在改变地质学[4]，其中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。

随着机器学习领域的不断发展，各种人工智能算法开始被广泛应用于岩石样本自动识别分类中，2017 年程国建、郭文惠等[5]提出了一种基于卷积神经网络的岩石粒度自动分类方法。这一方法，基于深度学习，通过卷积网络模型实现了岩石图像特征的自动提取，该方法同时建立了模式分类器，实现了基于岩石薄片

图像的岩石粒度自动识别。这一研究岩石粒度识别准确率较高，但是标准卷积神经网络参数大，训练时间较长，而且未研究岩性识别问题。2017 年，张野等[6]基于 Inception v3——深度卷积神经网络模型，运用迁移学习方法，建立了基于岩石图像数据集的深度学习迁移模型，实现了岩石岩性的自动识别与分类。但是，这一研究的实验数据为岩石自然图像，不是新鲜剖面，未考虑岩石表面风化等问题，影响因素过多。2018 年，白林，姚钰[7]等采用深度学习方法进行岩石识别，基于卷积神经网络构建岩石识别深度学习模型，达到 63% 的识别准确率；徐术腾等[8]基于深度学习系统 TensorFlow，设计有针对性的 U-net 卷积神经网络模型，自动提取矿相显微镜下矿石矿物的深层特征信息，实现镜下矿石矿物智能识别与分类。这些模型虽然实现了基于深度学习模型的岩石识别，但是准确率低，构建的模型存在冗杂、泛化性差等问题。2019 年，冯雅兴等[9]针对岩石新鲜剖面建立基于 AlexNet 孪生卷积神经网络的岩石识别模型，数据集共 25 类，3326 张图，实验表明，测试集精度达 89.4%。其工作优点在于综合考虑了岩石的全局图像信息和局部纹理信息，缺点在于识别精度不够高，且模型庞大，最优模型大小为 495MB。本文基于数据特性并且结合上述方法和模型优点，利用深度学习 ResNeSt 技术和 U-Net 模型的优势，解决岩石样本分类和岩石含油面积百分含量计算问题，以期达到在各个参考指标上都有较好的结果。

1.2 研究目标

本项目研究目标是基于现有岩石样本数据，通过图像处理技术和深度学习算法，设计出有效的模型识别出岩石样本的岩性类别及含油气情况，实现岩石样本智能识别分类。岩石样本数据为两种类型的相片，一种是白光相片，一种是荧光相片，其是采用工业相机在录井现场对岩屑和岩心样品在暗箱内使用不同灯光拍摄生成的。白光灯下拍摄的白光相片是用于提取颜色、纹理、粒度等特征识别岩性，荧光灯下拍摄的荧光相片是用于识别岩石含油气性，石油在紫外线照射下具有的发光特征，其中发绿色光和黄色光的部分含有油气。

在岩石样本岩性识别问题上，我们将根据数据特性构建基于 ResNeSt 技术的深度学习模型完成图像分类；在岩石含油面积百分含量计算问题上，我们使用多个图像处理技术精确计算含油面积百分比，并且使用语义分割 U-Net 模型对

含有地板背景的特殊图像进行分割处理后再计算。

1.3 研究内容与结构

本文将分为岩石样本岩性识别和岩石含油面积百分含量计算两个岩石样本智能识别领域的问题来介绍，具体安排如下：

第一章介绍研究的背景、意义和目标。本章介绍了对岩石的研究在各个领域都具有重大意义，其中最关键同时也是最基础的环节就是岩石样本识别，随着人工智能技术的不断发展，各种人工智能算法开始被广泛应用于岩石样本自动识别分类中，本章回顾和讨论了几种人工智能算法在岩石样本识别问题上的应用，分析其优势和不足之处；同时阐明了研究目标，通过图像处理技术和深度学习算法，设计出有效的模型识别出岩石样本的岩性类别及含油气情况，实现岩石样本智能识别分类。

第二章研究基于 ResNeSt 构建的岩石样本岩性智能识别与分类的方法。本章中介绍了 ResNeSt 技术，阐明了其中运用到的深度残差学习、Inception、组卷积和注意力机制等技术。本章使用 ResNeSt 技术构建岩石样本岩性智能识别与分类模型，从数据预处理、模型搭建和参数调整等阶段对模型进行了详细的阐述，最后将不同裁剪方式的实验结果进行对比，效果良好。

第三章研究岩石含油面积百分含量计算问题。本章介绍了 U-Net 技术，使用图像转换、掩膜处理等图像处理技术计算含油面积百分比含量，对于含有地板背景的特殊图像，使用语义分割 U-Net 模型进行像素分类处理，得到岩石实体相对面积，进而计算含油面积，计算结果良好。

第四章总结全文，对本文实验结果作汇总，提出一些在实验过程中遇到的问题和实验的不足之处，最后提出一些在其他方面潜在的优化方案，结束本文。

2 基于 ResNeSt 构建的岩石样本岩性智能识别与分类方法研究

近年来，随着深度学习技术的发展，神经网络技术成为自动识别岩石样本的重要方法，其大大改善了以往方法中主观性强、识别成本高、错误率大等问题。目前多种神经网络模型可应用于岩石样本分类，如 VGG、ResNet、GoogleNet、ResNeSt 等[10]。本文采用先进的 ResNeSt 技术提取岩石样本特征，实现岩石样本岩性智能识别与分类。

2.1 相关技术

2.1.1 深度残差学习

深度神经网络以端到端的方式集成了高、中、低层次的特征与分类器，并且可以通过堆叠卷积层的数量来丰富图像特征的层次。然而当网络达到一定的深度后，再堆叠卷积层不仅使神经网络的收敛变得更慢，还可能降低网络的精度。为了解决这一问题，何凯明等在 2015 年提出了深度残差学习神经网络（ResNet）。ResNet 通过多个层来学习输入与输出之间的残差表示，如图 1 所示，而非直接使用卷积层来直接学习输出与输入间的映射。通过引入残差学习，构建神经网络可以很容易的通过提高网络的深度来提高精度[11]。

神经网络在前向传播过程中，随着网络深度的增加，卷积层包含的图像信息将逐步减少。引入残差块后，不同层间可跨层连接，既可以从前面的网络层接受输入，也可以向后面的神经层进行输出，使更深的卷积层也能学习到低层的特征，从而避免网络过深时丢失图像信息。

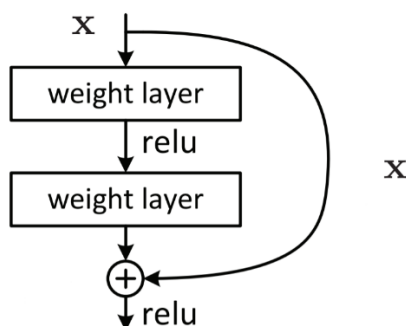


图 1 残差学习模块

2.1.2 Inception

如 2.1.1 所述，通过堆叠残差块增加网络深度可有效提高分类精度。但是，增加网络深度也会使网络的参数数量快速增长，而参数过多则更容易导致模型过拟合；并且，提高网络深度还会导致计算资源的使用急剧增加。如图 2 所示，Szegedy 等[12]提出 Inception 结构来解决这一问题：采用多尺度卷积高效学习特征。

Inception 模块采用不同大小的卷积核进行卷积之后进行拼接，从而使得网络能够学习不同感受野的信息，并将不同尺度的特征进行融合。池化层的加入在实践中被证明有效，因此 Inception 模块也嵌入了池化层。为减少计算量，Inception 模块采用 1×1 卷积核来进行降维。

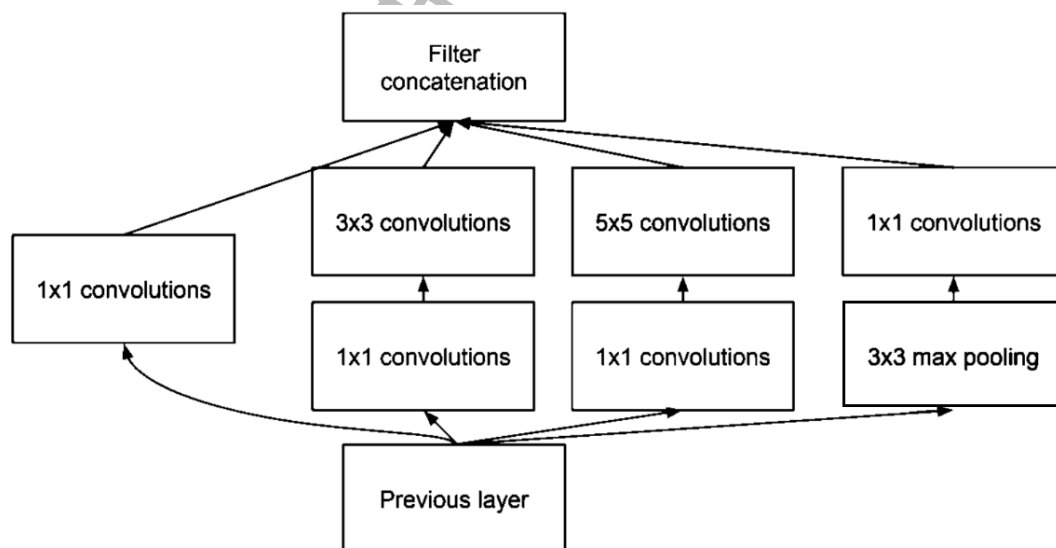


图 2 Inception 模块结构

2.1.3 组卷积

随着残差学习与 Inception 模块的加入，研究者只需加深网络即可达到提高精度的效果。但是，加深网络会带来超参数与计算量的增加，这会大大增加网络设计的难度与训练时间。谢赛宁等[13]提出的 ResNeXt 可以有效避免这一问题：在不增加网络参数数量的条件下提高网络精度，同时减少超参数的数量。

ResNeXt 由图 3 所示的残差学习模块堆叠而成。如 2.1.2 所述，Inception 使用了 split-transform-merge 策略，即先将输入分成几部分，然后分别做不同的运算，最后再合并到一起。这样可以有效减少参数的数量，但是对于不同尺寸的输入，Inception 模块需要设置不同的超参数来适应，这使得模型的设计过于复杂且难以泛用。ResNeXt 则可以指定分组数量，按分组数均分输入，不同分组的卷积层结构一致，然后再进行融合，从而在不增加网络复杂度的同时减少所需设置超参数的数量。ResNeXt 结构如下表 1 所示。

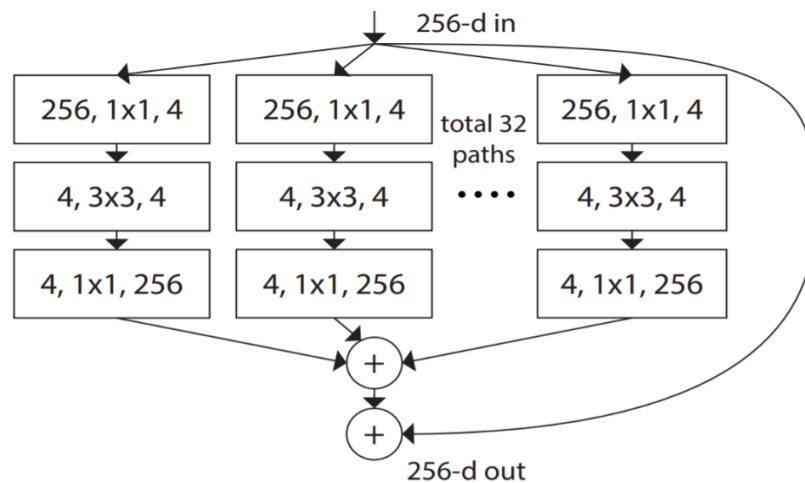


图 3 ResNeXt Block

表 1 ResNeXt 结构

结构	输出大小	ResNeXt-50 ($32 \times 4d$)
conv1	112×112	7×7 , 输出通道 64, 步幅 2
conv2	56×56	3×3 最大池化, 步幅 2 $\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128, C = 32 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3	28×28	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256, C = 32 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4	14×14	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512, C = 32 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5	7×7	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024, C = 32 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	全局平均池化, 全连接层 1000-d, softmax 激活函数
参数量		25.5×10^6
每秒浮点运算次数		4.2×10^9

2.1.4 注意力机制

注意力机制的灵感源于人类自身的实践认识。在观察图像时，我们的视觉系统更倾向于观察图像中关键部分而忽略掉图像中的无关信息[14]。这种机制也可以运用到神经网络的设计里：可否通过调整图像关键部分的权重或者调整卷积中不同通道的权重来提高分类的精度？胡杰等[15]提出的 SE(Squeeze-and-Excitation)模块，通过自适应的重新校准通道响应来引入通道注意力机制，李翔等[16]提出的 SK(Selective-kernel)模块，通过两个网络分别引入特征图注意。

SE 模块（图 4）首先对卷积得到的特征图做全局平均池化，得到通道级的全局特征，然后通过训练判断不同通道的重要性，得出不同通道的权重，最后用得到的权重乘以最初的特征图得到引入注意力机制后的最终特征。SK 模块（图 5）

首先使用不同大小的卷积分别对输入图像进行卷积操作，然后通过训练判断不同大小卷积核卷积结果的重要性，得到不同卷积核卷积图像的权重，用得到的权重乘以对应的特征图得到加权后的特征图，最后进行融合得到最终的特征图。

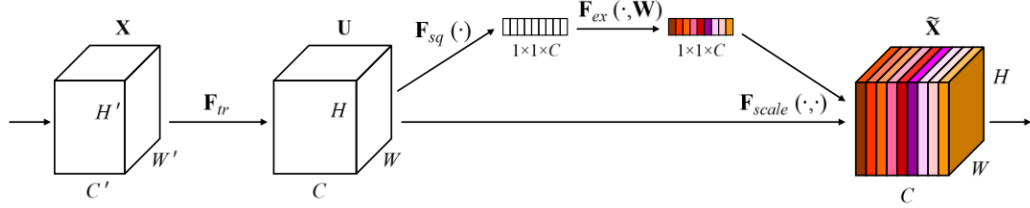


图 4 SE 模块

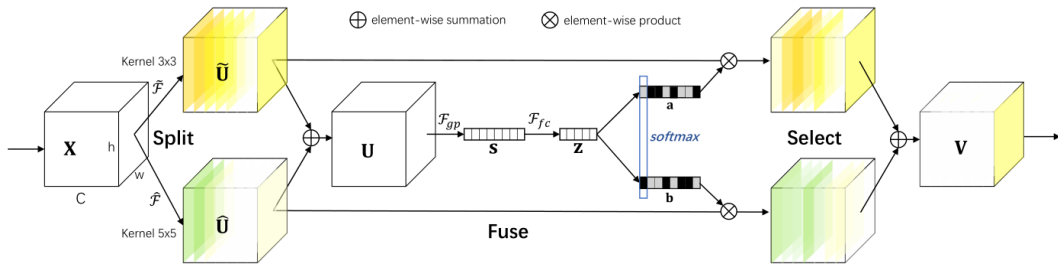


图 5 SK 模块

2.2 实验过程

2.2.1 数据预处理

原始数据集图像共 315 张，本文首先按 9: 1 的比例划分训练集与测试集，测试集 32 张图像，训练集 283 张图像。283 张图像对于神经网络的训练来说是不够的，因此需要进行数据增强。我们首先对原始数据集的图像做了旋转、翻转与镜像处理，得到 1132 张图像。数据集共有 7 类岩石，各类岩石图像数并不相同，其中浅灰色细砂岩图像数是灰色细砂岩的近 5 倍，即数据集存在数据失衡的现象。所以，我们进一步对其中占比较少的岩石种类的图像进行小幅度旋转处理来达到数据增强的目的。最终，使得数据集各岩石种类的图像占比大致相等。数据增强前后效果对比如表 2 所示。

表 2 数据增强效果

	浅灰色 细砂岩	深灰色 泥岩	深灰色 粉砂质 泥岩	灰色泥 质粉砂 岩	灰色细 砂岩	灰黑色 泥岩	黑色煤	总计
数据增强前	85	75	40	46	18	30	21	315
数据增强后	313	279	306	340	322	305	306	2171

由于训练集中部分图片既包括岩石，也包括地板背景，因此需要对这一部分图像进行裁剪处理，只保留图像中岩石的部分，如图 6 所示。对于测试集，由于测试图像仅有 32 张，故在训练前需预先对图像进行切割。对于不包含地板的图像，我们用大小为 320×320 的滑动窗口进行切割；对于包含地板的图像，我们首先裁剪出图像中心 960×960 大小的部分以去除无用的地板部分图像，再从中用大小为 320×320 的滑动窗口进行切割，最终得到 3,051 张测试图像。图像裁剪前后对比如图 7 所示。训练集、测试集详情如表 3 所示。

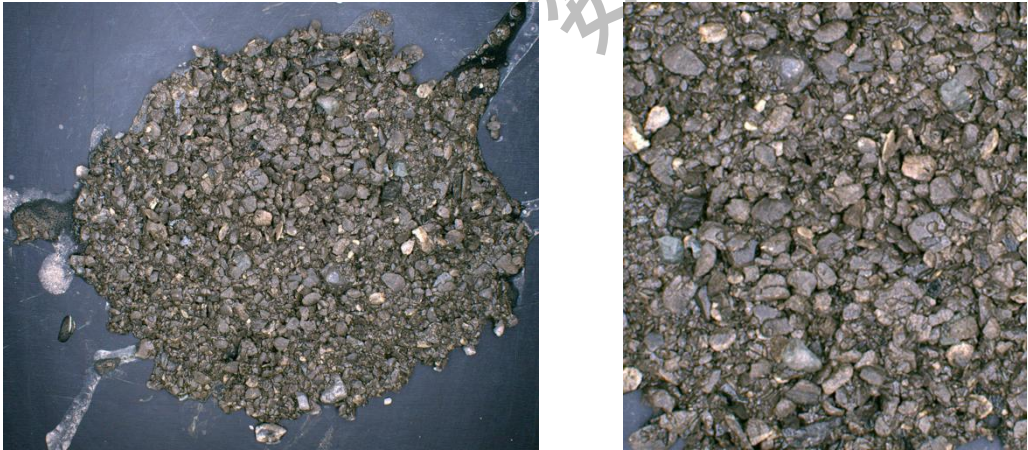


图 6 图像裁剪效果

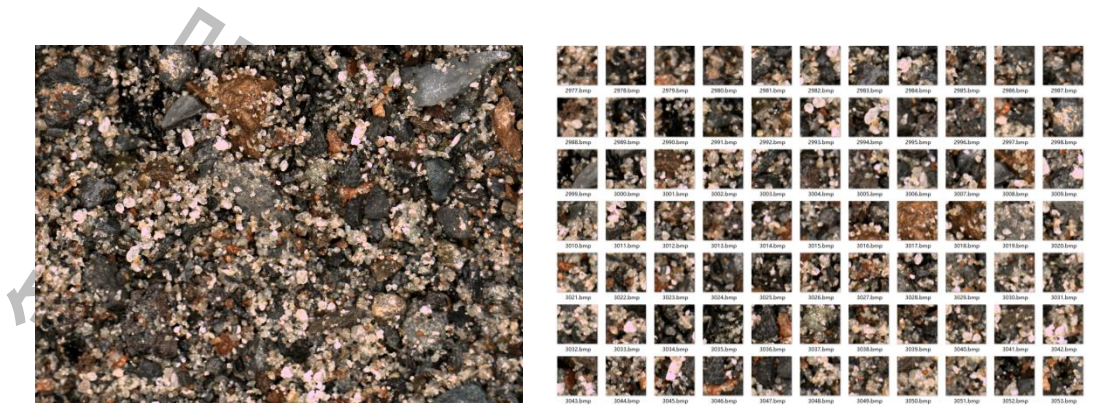


图 7 图像裁剪前后对比

表 3 训练集、测试集详情

	图像数量	图像大小
训练集	2,140	4096×3000 或 2448×2048
测试集	3,051	320×320

2.2.2 实验过程

2.2.2.1 实验环境

实验采用的硬件环境与软件环境如表 4 所示。

表 4 实验环境

硬件项目	内容	软件项目	内容
CPU	Intel(R) Core(TM) i5-10600KF CPU @ 4.10GHz 4.10 GHz	python 版本	3.8.5
GPU	NVIDIA GeForce RTX 3060	Conda 版本	4.9.2
内存	16G	Cuda 版本	11.0
显存	12G	深度学习框架	Pytorch 1.7.1

2.2.2.2 模型建立

张航等人基于 ResNet 提出的一种变体 ResNeSt，结合了残差学习、组卷积、注意力机制等思想，性能显著提升而参数量没有显著增加。ResNeSt 将注意力机制与 ResNeXt 模块进行结合，得到分割注意力模块，ResNeSt Block 结构如图 8 所示。ResNeSt Block 首先将输入分为 k 个，每个记为 Cardinal- k ，然后将每个 Cardinal 拆分为 r 个，记为 Split- r 。Cardinal 块结构如图 9 所示。Cardinal 模块借鉴了 SE 模块与 SK 模块的思想，将划分而得的 Split 块经过池化、激活等操作，计算不同 Split 块的注意力权重。相较于 ResNeXt 在 block 中将输入划分为 k 个切片，ResNeSt 进一步对切片再划分为 r 个切片，对这 $k \times r$ 个切片计算得出其注意力参数，在 Split attention 计算得出 Cardinal 模块中 r 个切片的总体特征。通过堆叠几个分割注意力模块创建了一个类似 ResNet 的网络，即 ResNeSt（S 代表分

割) [17]。

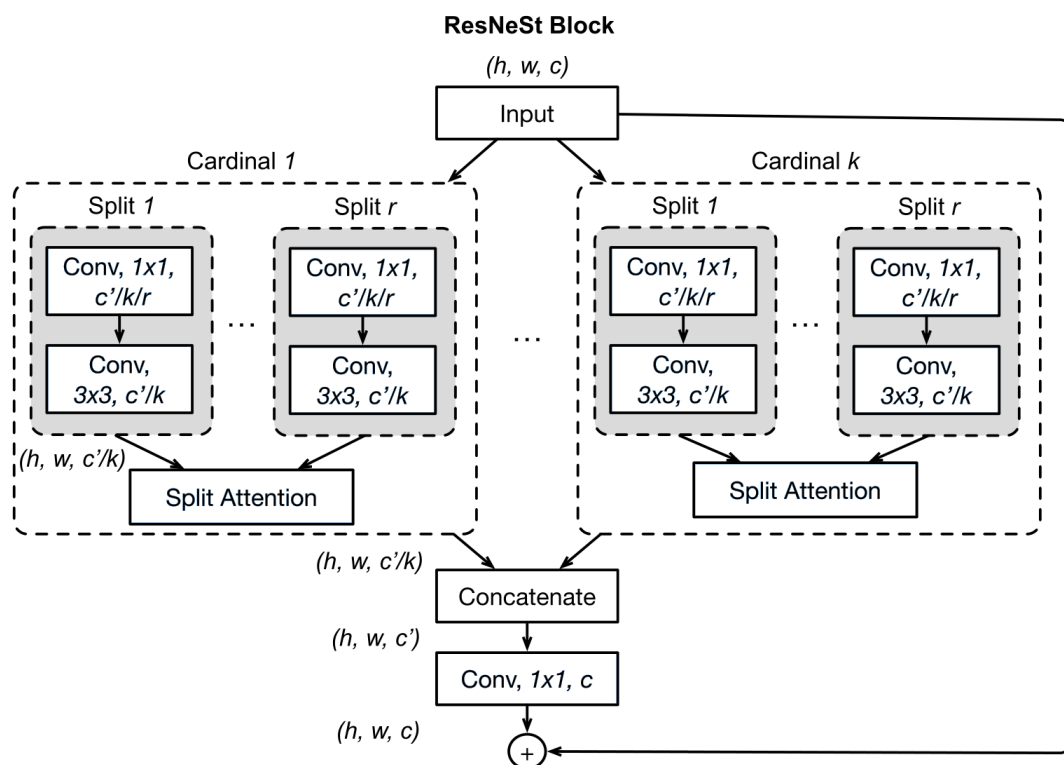


图 8 分割注意力模块

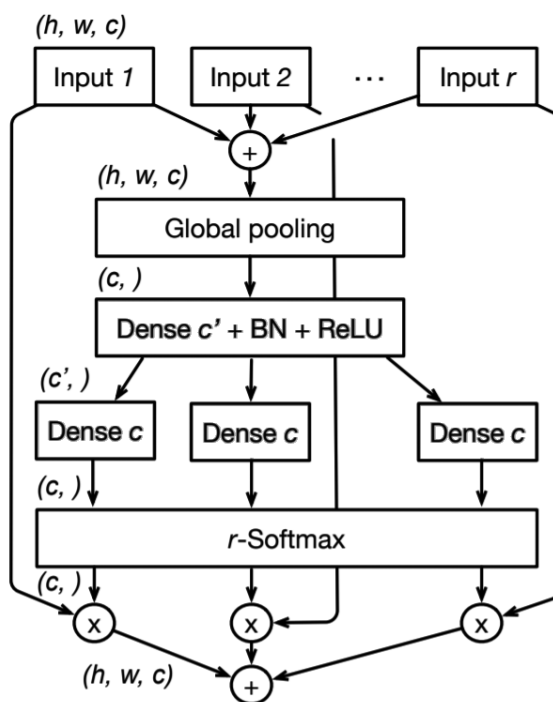


图 9 Cardinal 块结构

我们搭建 ResNeSt50 进行岩石图像分类。由于岩石的种类数量为 7，我们在 ResNeSt50 的结尾添加一个输出为 7 的全连接层，导入预训练的 ResNeSt50 参数

后进行训练。

2.3 实验结果

在训练集上，基于 2.2.2 中建立的模型进行训练。我们设置训练迭代次数为 500 次，使用 Adam 优化器，学习率为 0.0001，损失函数为交叉熵 (CrossEntropyLoss)，批训练大小为 16。Adam 一种有效的随机优化方法，该方法通过第一、第二梯度的估计，计算不同参数的自适应学习速率，其具有实现简单、计算高效、对内存需求少和梯度对角缩放的不变性等优点。交叉熵经常在神经网络中作为损失函数， p 表示真实标记的分布， q 则为训练后的模型的预测标记分布，交叉熵损失函数可以衡量 p 与 q 的相似性，交叉熵作为损失函数的另一个优点是使用 sigmoid 函数在梯度下降时能避免均方误差损失函数学习速率降低的问题。由于图像较大，我们在预测岩石种类时首先以 960×960 裁剪出图像中心，然后再裁剪出 9 张 320×320 大小的图像进行预测，取其中占比较大的岩石种类作为预测并与只随机裁剪一张 320×320 大小的图像用于预测进行对比。实验结果如表 5 所示。

表 5 实验结果

方法	准确率
裁剪 1 张图像用于预测	92.8%
裁剪 9 张图像用于预测	97.2%

由表 5 可知，单独裁剪 1 张图像用于预测时准确率达到 92.8%，裁剪 9 张图像用于预测时准确率达到 97.2%。

2.4 本章小结

本章提出并实现了一种岩石样本岩性智能识别与分类的方法。我们采用 ResNeSt 技术作为模型的主干，搭建了 ResNeSt50 神经网络，在完成参数调整和迭代训练后，得到了一个基于岩石样本图像的岩性自动分类的模型。实验表明，我们的网络在裁剪一张图像用于训练的情况下准确率可达 92.8%，在裁剪九张图像用于训练时准确率可达 97.2%，表现良好，相比于之前的相关研究，我们的方

法基于数据特点引入了几种先进的深度学习技术,综合考虑了岩石的全局图像信息和局部纹理信息,分类准确率高,模型大小适中,计算速度快。这一研究为地质调查中岩石图像的岩性快速分类问题提供了新的解决途径。

3 基于 U-Net 构建的岩石含油面积百分含量计算模型

通过第二章的方法，我们可以得到岩石样本的岩性类别，在油气勘探中和矿产资源勘探中，尤其是在固体金属矿产资源勘探中，还需要对其岩石样本的含油面积百分含量进行计算，这样才能进一步确认岩石样本的类别。在给定的数据图像中，存在一部分含有地板背景的特殊图像，我们使用语义分割 U-Net 技术构建模型分割出岩石，随后运用图像处理技术计算含油面积百分比。

3.1 U-Net 技术

3.1.1 语义分割

语义分割是计算机视觉研究中的一个经典难题，已经成为图像理解领域关注的一个热点，语义分割是图像分析的第一步，是计算机视觉的基础，是图像理解的重要组成部分，同时也是图像处理中最困难的问题之一。所谓图像分割是指根据灰度、彩色、空间纹理、几何形状等特征把图像划分成若干个互不相交的区域，使得这些特征在同一区域内表现出一致性或相似性，而在不同区域间表现出明显的不同。简单的说就是在一副图像中，把目标从背景中分离出来。比如图 10 中左图的语义就是三个人骑着三辆自行车；分割的意思是从像素的角度分割出图片中的不同对象，对原图中的每个像素都进行标注，图 10 右图中粉红色代表人，绿色代表自行车。



图 10 语义分割

语义分割运用广泛，主要应用领域有地理信息系统、无人车驾驶、医疗影像分析和机器人等。虽然到目前为止，还不存在一个通用的完美的语义分割的方法，但是对于语义分割的一般性规律则基本上达成了共识，已经产生了相当多的研究成果和方法。其中基于深度学习的语义分割方法主要包括基于特征编码的方法、基于区域选择的方法和基于上采样的方法，基于上采样的方法具有很多优点，它能对图像进行像素级的分类，从而解决语义级别的图像分割问题；它能较好解决卷积神经网络不可逆导致的图像分辨率低、细节丢失等问题，从而得到更加准确的分割边界。所以，本章使用基于上采样的方法中具有代表性的 U-Net 模型对于含有地板背景的特殊图像进行处理。

3.1.2 U-Net 模型

2015 年，Olaf Ronneberger 等[18]提出了 U-Net 模型，U-Net 模型基于 FCN 模型，对 FCN 模型的基本结构进行了更精细的设计，更为高效。U-Net 属于编码器-解码器模型，这类模型将任务分为两个部分：第一部分为编码器，它将图像输入压缩成特征向量；第二部分为解码器，它将编码器的输出即特征向量解码为掩膜图像。

U-Net 模型呈现为 U 型结构，编码器为左半部分，由两个 3×3 的卷积层（ReLU）加上 2×2 的最大池化层反复堆叠构成，每经过一次下采样，通道数翻倍；解码器为右半部分，由一个 2×2 的上采样卷积层（ReLU）加上 Concatenation，再加上 2 个 3×3 的卷积层（ReLU）反复堆叠构成；最后通过一个 1×1 卷积层将通道数变成期望的类别数，即转化为掩膜图像输出。

由于 U-Net 的 U 型结构，使得它能够结合底层和高层的信息。底层(深层)信息即经过多次下采样后的低分辨率信息，能够提供分割目标在整个图像中上下文语义信息，可理解为反应目标和它的环境之间关系的特征。这个特征有助于物体的类别判断。高层(浅层)信息经过 concatenation 操作从编码器直接传递到同高度解码器上的高分辨率信息，能够为分割提供更加精细的特征，如梯度等。U-Net 结合了底层信息与高层信息，所以在图像分割领域表现优秀。U-Net 模型结构如下图 11 所示。

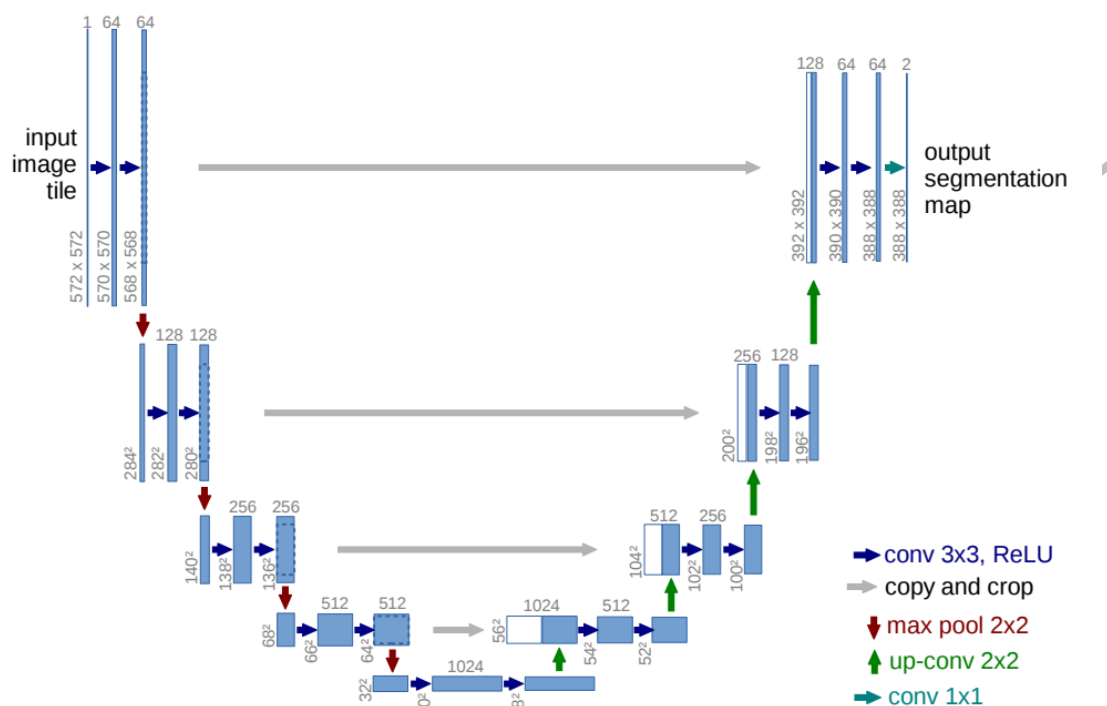


图 11 U-Net 模型结构

3.2 实验流程

3.2.1 数据处理

原始数据集没有标注油气的量，所以首先需从图像中提取出油气部分用于计算面积。我们首先将原始图像转化为 HSV 颜色空间的图像。HSV 对于人来说是一种直观的颜色模型，它由色调（H）、饱和度（S）、明度（V）三个参数组成。色调表示色彩信息，即所处的光谱颜色的位置。纯度 S 为一比例值，范围从 0 到 1，它表示成所选颜色的纯度和该颜色最大的纯度之间的比率。S=0 时，只有灰度。明度表示色彩的明亮程度，范围从 0 到 1。将 RGB 图像转化为 HSV 图像后，我们可以用色调（H）来指定所需的颜色，用饱和度（S）来指定颜色的深浅，明度（V）指定颜色的明暗。对于油气面积提取任务，将原始图像转化为 HSV 图像后，将其中绿色与黄色部分转为白色，其余部分转为黑色，得到掩膜图像。最后，我们利用掩膜图像提取原始图像中的油气部分并计算其面积。原始图像、HSV 图像、掩膜图像、提取结果图像如图 12 所示。

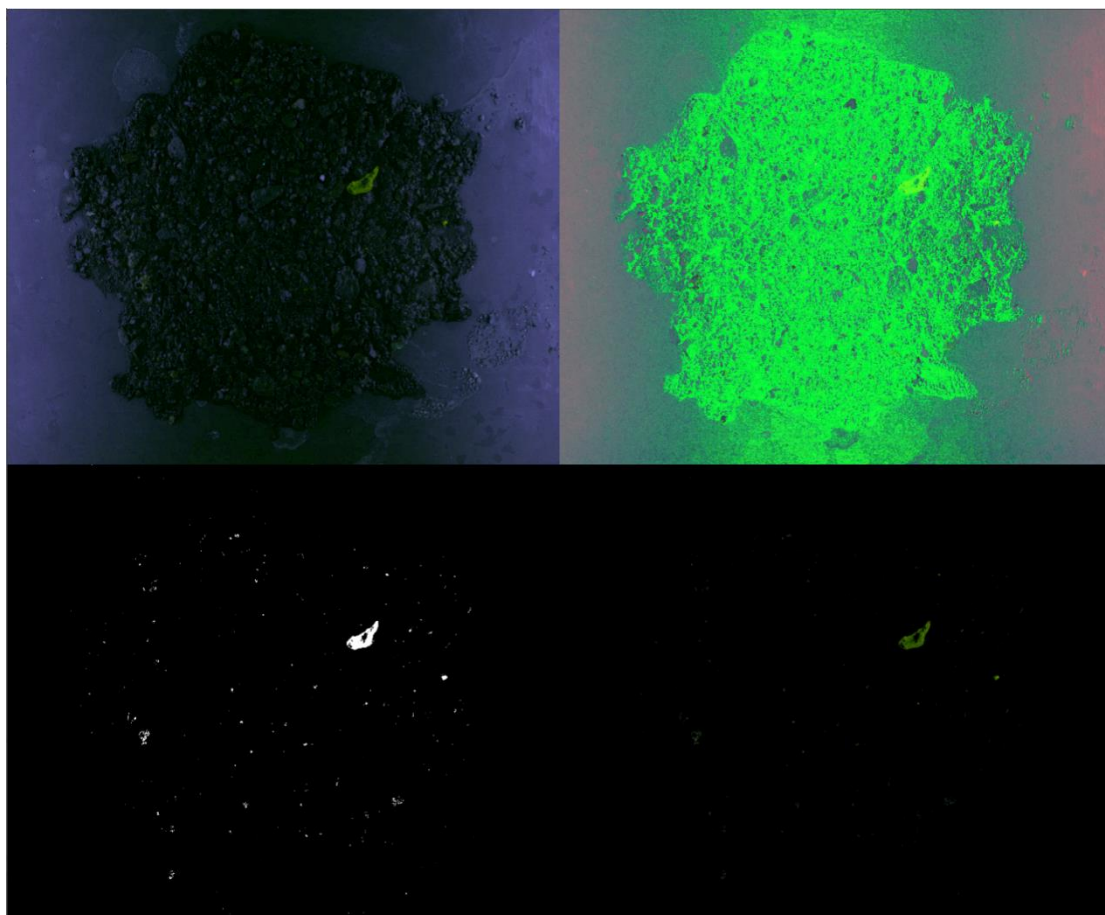


图 12 原始图像、HSV 图像、掩膜图像、提取结果图像

由于数据集中部分图像不仅包含岩石，还包括地板，所以我们需要从原始图像中提取出岩石部分的面积。本文采用 U-Net 对该类岩石图像进行分割。由于初始图像尺寸为 2048×2448 ，若将完整图像送入模型进行训练所需占用的显存超过 12G，所以我们将图像按比例缩小至 320×384 。由于数据集不包含岩石图像的掩膜图像，所以我们使用 labelme 人工标注了图像。标注效果如图 13 所示。

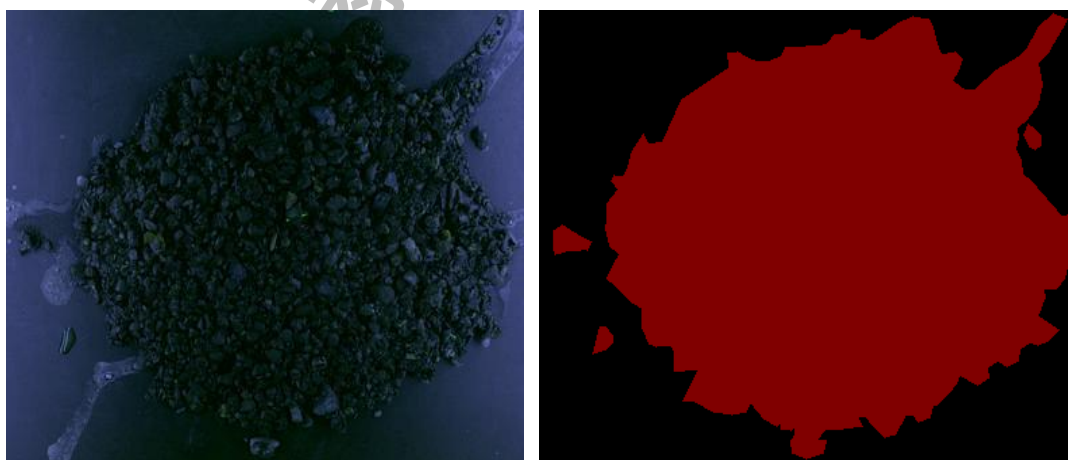


图 13 岩石图像标注效果

3.2.2 实验环境

实验采用的硬件环境与软件环境如上文中表 4 所示，与第二章一致。

3.2.3 评价指标

在图像分割领域中，Dice 系数(Dice coefficient)是一个常用的评价指标，所以本章使用 Dice 系数(Dice coefficient)评价实验结果。Dice 系数(Dice coefficient)根据 Lee Raymond Dice 命名，是一种集合相似度度量函数，通常用于计算两个样本的相似度(值范围为 $[0, 1]$)。其中 DiceCoefficient 计算公式如下：

$$DiceCoefficient = \frac{2|X \cap Y|}{|X| + |Y|}$$

其中 $|X \cap Y|$ 表示 X 和 Y 集合的交集， $|X|$ 和 $|Y|$ 表示其元素个数，对于分割任务而言， $|X|$ 和 $|Y|$ 表示分割的正确掩膜图像与预测掩膜图像。此外，Dice Loss 计算的公式如下：

$$DiceLoss = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

3.3 实验结果

下面对这些实验结果进行分析：

(1)提取岩石图像中油气部分图像并计算面积

对于不包含地板的图像，使用 3.2.1 方法得到提取结果后，对提取结果图像按像素点进行遍历，计算得油气部分与原图像的像素点个数，按下述公式即可得油气部分的面积占比。

$$\text{油气面积占比} = \frac{\text{油气部分像素点个数}}{\text{原图像像素点个数}}$$

(2)使用 U-Net 分割岩石图像

对于包含地板的图像，还需额外计算岩石部分占图像面积的比例。首先将按比例缩放后的图像与人工标注的掩膜图像按 9: 1 划分训练集与测试集，训练集送入 U-Net 模型进行训练建模，采用交叉熵作为损失函数，优化器为 Adam，进

行迭代 20 次的模型训练后,得到岩石图像分割模型,测试集 DiceCoeff 达到 94.3%;其次将所有图像送入训练好的 U-Net 模型,得到对应的掩膜图像;最后,遍历掩膜图像像素点,计算得岩石部分占图像面积的比例。该类图像的油气面积计算公式如下所示:

$$\text{油气面积占比} = \frac{\text{油气部分像素点个数}}{\text{岩石部分像素点个数}}$$

根据上述方法,可得出所有图像油气面积百分含量,部分计算结果如下表 6 所示,平均含油面积百分含量为 0.2976%。

表 6 油气面积计算结果

样本编号	含油面积百分含量
1	0.9381%
2	0.0505%
3	0.0262%
4	1.3769%
5	0.0300%
.....	
346	6.6744%
347	2.1584%
348	0.9904%
349	0.2860%
350	2.4046%
平均值:	0.2976%

3.4 本章小结

本章提出并实现了一种岩石含油面积百分含量计算的方法。我们采用 HSV 颜色模型方便的提取出图像中的绿色与黄色部分,并用 U-Net 对数据集中包含地板的图像进行分割,提取出其中岩石部分的图像,效果图像如下图 14 所示。我们采用交叉熵作为损失函数,优化器为 Adam,进行迭代 20 次的模型训练,得到了

一个岩石图像分割的模型。实验表明，我们的模型的 DiceCoeff 达到 94.3%，这一研究为地质调查中岩石图像的油气面积计算提供了新的方法。

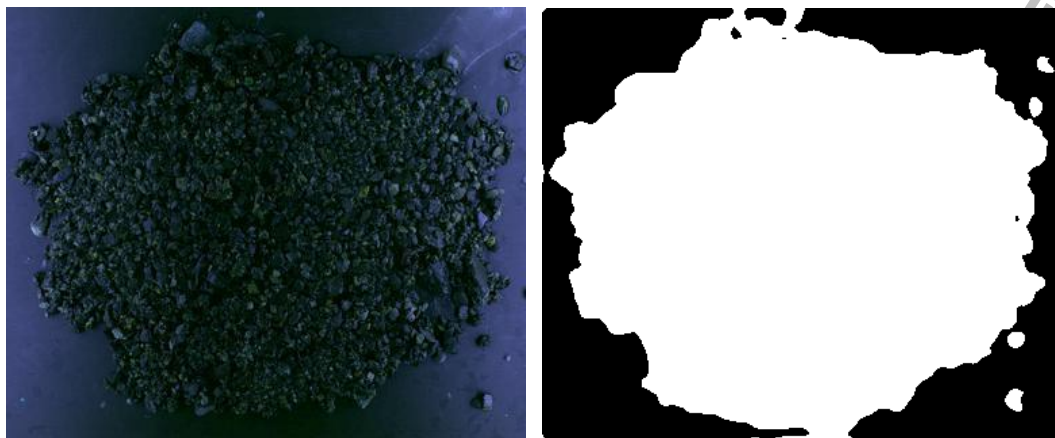


图 14 分割效果

4 总结

本文主要围绕岩石样本识别中两个问题开展：岩石样本岩性识别与含油面积百分含量计算。在岩石样本岩性识别问题上，第二章利用 ResNeSt 技术构建了岩石岩性分类模型，通过数据分析，我们首先对数据进行了旋转、翻转与镜像等数据增强处理，构建 ResNeSt50 模型进行训练。测试结果表明，本文模型在裁剪一张图像用于训练的情况下准确率可达 92.8%，在裁剪九张图像用于训练时准确率可达 97.2%，完成了岩石样本岩性的智能识别与分类。在含油面积百分含量计算的问题上，第三章首先构建 U-Net 模型对含油地板背景的图像进行分割，分离岩石与背景，实验结果表明，本文模型的 DiceCoeff 达到 94.3%，分割效果优秀。最后运用图像转换、掩膜处理等图像处理技术计算含有面积百分比，顺利完成了含油面积计算的工作。本文对于岩石样本岩性识别与含油面积计算这两个问题的提出、思考、处理、解决均有比较完整的描述，尽可能利用现有方法进行对比，阐述出该方案的效果与不足之处，最终得出较为满意的结论。

本文针对岩石图像的特征进行分析，构建了 ResNeSt 模型并将该模型应用于岩石图像数据集上进行训练与预测，虽然取得了较好的识别精度，但鉴于时间和精力有限，还有许多研究工作值得进一步的扩展，本文后续的工作将着力于继续提高模型的精度。首先我们观察到原始数据集的部分标签可能标注错误，一方面可以邀请专家对数据集进行重新标注，另一方面可以基于深度学习对数据集进行数据清洗[19]，挑选出模型容易分错的图像即可能标签有误的图像进行重新标注，以此提升数据集的正确性；其次可以尝试集成学习的方法，构建新的模型，将图像采用直接缩放的方式输入模型进行训练，从而学习图像的全局特征等方法来提高模型的精度。

参考文献

- [1] 李昌年.简明岩石学[M].中国地质大学出版社,2010.
- [2] 张儒,马炯.榆林气田石英砂岩气层的测井方法识别评价研究[J].辽宁石油化工大学学报,2019,39(04):65-71.
- [3] 胡祺.融合多维信息的岩石薄片图像深度学习分类方法[D].浙江大学,2019.
- [4] 周永章,左仁广,刘刚,袁峰,毛先成,郭艳军,肖凡,廖杰,刘艳鹏.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/OL].矿物岩石地球化学通报:1-18[2021-05-03].<https://doi.org/10.19658/j.issn.1007-2802.2021.40.038>.
- [5] 程国建,郭文惠,范鹏召.基于卷积神经网络的岩石图像分类[J].西安石油大学学报(自然科学版),2017, 32(4):116-122.
- [6] 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(2).
- [7] 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(07):178-182.
- [8] 徐述腾,周永章.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, 2018,34(11):3244-3252.
- [9] 冯雅兴,龚希,徐永洋,谢忠,蔡惠慧,吕霞.基于岩石新鲜面图像与孪生卷积神经网络的岩性识别方法研究[J].地理与地理信息科学,2019,35(05):89-94.
- [10] 李燕.基于深度学习的岩石图像识别[D].北京林业大学,2020.
- [11] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [12] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [13] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1492-1500.
- [14] Zhu X, Cheng D, Zhang Z, et al. An empirical study of spatial attention mechanisms in deep networks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 6688-6697.
- [15] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [16] Li X, Wang W, Hu X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 510-519.
- [17] Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks[J]. arXiv preprint arXiv:2004.08955, 2020.

- [18] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28
- [19] 余华擎. 基于深度学习的图像数据清洗方法研究[D].北京工业大学,2018.