

基于深度学习的岩石样本智能识别分类

摘要

本文从设计了岩石样本智能识别分类深度学习算法，构建了岩石样本岩性智能识别模型，针对数据集 rock 实现岩石样本岩性智能识别与分类。并利用传统的数字图像处理方法，设计出了计算岩石的含油面积百分含量的算法。

由于岩石样本图像像素过大且图片的数据集过小，故本文首先对岩石样本图像进行的数据平衡处理和数据增强处理，以扩充数据集。

针对问题 1，本文提出了两种基于深度学习神经网络的岩石样本识别分类模型。一是基于改进VGG的岩石样本识别分类模型，二是基于改进 ResNet-50 的岩石样本识别分类模型，并分别利用两种模型对岩石样本图片进行训练。经过 50 轮迭代后，训练集在两种模型上的准确率均为 100%，验证集在基于改进VGG的岩石样本识别分类模型上的准确率为 82.456%，在基于改进 ResNet-50 的岩石样本识别分类模型上的准确率为 86.842%。

针对问题 2，本文将问题拆分成含油面积计算与岩石面积计算两个子问题。基于HSV颜色空间变换与典型颜色统计最后结合阈值分割，实现了符合预期的含油像素分割。然后基于移动平均阈值分割与高斯滤波分区域阈值校正，解决了荧光图片光照不均的问题，然后基于形态学闭运算、开运算、迭代高斯滤波与阈值分割实现岩石样本分割与面积计算，最终为每一个荧光照射图片计算出了合理的含油面积百分含量。

关键词： 岩石识别 深度学习 图像分割 高斯滤波 形态学运算

目录

1 引言	1
1.1 研究背景	1
1.2 研究任务	1
1.2.1 岩石样本识别分类	1
1.2.2 岩石含油面积百分量	1
2 岩石样本识别分类	2
2.1 数据准备	2
2.1.1 数据描述	2
2.1.2 数据预处理	2
2.1.2.1 样本平衡处理	2
2.1.2.2 数据增强	2
2.1.2.3 图像放缩	3
2.2 基本理论	3
2.2.1 卷积神经网络	3
2.2.2 全卷积神经网络	4
2.2.3 残差连接(skip connect)	4
2.3 建立模型	5
2.3.1 基于改进VGG的岩石样本识别分类模型	5
2.3.1.1 实验过程及结果	7
2.3.2 基于改进ResNet-50的岩石样本识别分类模型	7
2.3.2.1 实验过程及结果	8
2.4 模型效果评价与优化	9
2.4.1 模型优点	9
2.4.2 模型尚存缺陷	9
2.4.3 优化方向	9
3 含油量分析	10
3.1 建立模型	10
3.1.1 含油像素分割映射 \mathcal{M}_{oil}	10
3.1.1.1 基于灰度的含油像素分割	10
3.1.1.2 基于HSV颜色空间的含油像素分割	11
3.1.2 岩石分割映射 \mathcal{M}_{rock}	12
3.1.2.1 移动平均阈值分割	13

3.1.2.2 结合高斯滤波的形态学处理	14
3.1.2.2.1 数学形态学	14
3.1.2.2.2 高斯滤波	16
3.2 算法设计与实验	17
3.2.1 含油像素分割	17
3.2.2 岩石分割	19
3.2.3 含油百分比计算	20
3.3 模型效果评价与优化	22
3.3.1 模型优点	22
3.3.2 模型尚存缺陷	22
3.3.3 优化方向	22
4 总结	23
参考文献	26

1 引言

1.1 研究背景

岩石岩性及矿物的识别与分类是地质学研究中十分重要的内容，近年来诸多学者运用不同的方法对其进行了研究。随着人工智能时代的到来，运用智能学习分析方法，通过机器学习等智能算法对岩石图像特征进行分析处理，减少对于专业知识和设备的依赖，将图像识别技术运用在识别岩石岩性的跨界结合越来越多。例如，Singh et al. (2010)构建神经网络对玄武岩矿物图像进行处理分析，实现了对玄武岩矿物纹理的有效识别；张嘉凡等(2016)提出了基于聚类分析算法的岩石 CT 图像分割及量化方法；张翠芬等(2017)利用岩性单元的特征向量进行图像的彩色合成，使得岩性单元可识别性显著增加；Li et al. (2017)采用迁移学习方法对砂岩显微图像进行了训练，最终获得了精度较高的砂岩显微图像分类模型。

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。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，应用图像深度学习等智能算法可以通过分析岩石样本图像的特征，从而建立岩石识别分类的数学模型，实现岩石样本自动识别分类，并且使识别过程更加智能化、自动化。

1.2 研究任务

在数据集 rock 中，“白光/荧光”标签为“1”的数据是相同白光环境下拍摄的岩石样本图像数据，“白光/荧光”标签为“2”的数据是相同荧光环境下拍摄的岩石样本图像数据，利用数据集 rock 的数据，完成以下两部分内容。

1.2.1 岩石样本识别分类

设计合适的机器学习或深度学习算法，构建岩石样本岩性智能识别模型，针对数据集 rock 实现岩石样本岩性智能识别与分类。

1.2.2 岩石含油面积百分量

石油在紫外线照射下具有发光特征，荧光灯下拍摄的相片中绿色或黄色部分是含油的，岩石的含油面积百分含量为绿色和黄色部分的面积占总岩石面积的百分比。本部分需要设计出合适的算法计算岩石的含油面积百分含量。

2 岩石样本识别分类

2.1 数据准备

2.1.1 数据描述

在数据集 rock 中，岩石样本分为 7 类，白光下的岩石样本图片共计 315 张。其中，黑色煤 21 张，灰黑色泥岩 30 张，灰色泥质粉砂岩 46 张，灰色细砂岩 18 张，浅灰色细砂岩 85 张，深灰色粉砂质岩 40 张，深灰色泥岩 75 张。各类占比如图1所示：

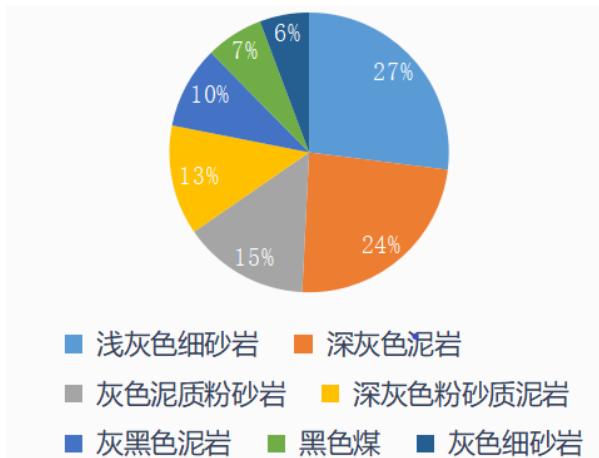


图 1: 样本类别统计图

2.1.2 数据预处理

2.1.2.1 样本平衡处理

通过上述数据观察可知，样本不平衡严重，为了使数据集平衡，提高训练效果。由于样本最多的一类岩石有 85 张图片，我们对每一个类别进行样本复制扩充，对一个类别现有图片随机抽取图片进行复制，当该类别图片数量复制至 85 张时，停止复制。

2.1.2.2 数据增强

由于本文数据量较少，因此对数据进行增强以增加在后续模型学习过程中的样本多样性，防止陷入局部最优形成过拟合现象。

本文采用的数据增强方式有：

1. 以 50% 的概率对图片上下翻转；以 50% 的概率对图片左右翻转；
2. 从区间 $[0, 0.2]$ 随机采样一个数，对图片的两条边按该比例裁剪丢弃，后放缩至裁剪前；

- 3.从区间 $[0.75, 1.25]$ 随机采样一个数，对图片按该数值改变图像对比度；
- 4.从区间 $[0.75, 1.5]$ 随机采样一个数，对图片按该数值改变图像亮度；
- 5.从区间 $[0.8, 1.2]$ 随机采样一个数，对图片的宽按该数值进行放缩；同样的方法对图像的高也进行一次变换；
- 6.从区间 $[-0.2, 0.2]$ 随机采样一个数，按该数值比例对图片进行水平平移，负数表示向左平移，正数表示向右；同理对图像进行上下平移；
- 7.从区间 $[-180, 180]$ 随机采样一个数，按该数值作为旋转角度对图像进行旋转。

2.1.2.3 图像放缩

岩石数据集中 bmp 图片的宽高为 4096x3000，而后续模型设计时出于模型训练时内存空间的限制，在输入数据给模型时将图片放缩至 224x224，但发现在训练过程中实时进行该放缩操作，由于放缩比例太大，放缩缓慢，因此我们将图片提前离线放缩至 224x224 后保存至本地文件夹以加快后续网络训练。

2.2 基本理论

2.2.1 卷积神经网络

卷积神经网络(convolutional Neural Networks, CNN)是一种带有卷积结构的深度神经网络，卷积结构可以减少深层网络占用的内存量，有三个关键的操作，其一是局部感受野，其二是权值共享，其三是pooling层。权值共享的模式有效的减少了网络的训练参数，降低了复杂度，使其变得简单且适应性强，缓解了模型的过拟合问题。

卷积神经网络直接以图像作为输入，避免了传统方法提取特征的过程，在图像处理上可保留像素的空间位置关系。其网络结构对图像的平移、比例缩放、旋转等其他形式的形变具有高度不变性。卷积神经网络的强大之处在于它的多层次网络结构能自动学习输入数据的深层特征，不同层次的网络可以学习到不同层次的特征。浅层网络层感知区域较小，可以学习到输入数据的局部域特征（如图像物体的颜色、几何形状等）；深层网络层具有较大的感知域，能够学习到输入数据中更加抽象一些特征（如图像物体的属性、轮廓特点、位置信息等高维性质）。深层次的抽象特征对图像中物体的大小、位置和方向等敏感度较低，从而大大提高了物体的识别率。

在卷积层中，卷积核的作用等同于滤波器，由于 RGB 色彩模式的图像每个通道等同于一个二维矩阵，所以卷积层通过滑动窗口的方式将卷积核与输入图像的每个通道进行卷积操作，提取出不同类型的特征，称为特征图(feature map, FM)，特征图的个数和卷积核的个数相同；池化层又称为子采样层(sub-sampling

layer), 通过池化操作对数据进行降维, 缩小输入数据规模, 减少计算量, 通常有最大池化、平均池化和随机池化等计算方式; 由于卷积和下采样操作造成特征图分辨率降低, 所以利用反卷积层通过插值运算将特征图恢复到输入图像大小, 然后输出为二维矩阵, 矩阵的值表示某个像素归属为某一类的概率。卷积网络通过误差反向传播算法周期性地更新卷积核的权重实现进行求解优化, 通过卷积网络模型在数据集上的误差和精确度确定网络优化程度。

作为深度学习的代表算法之一, 卷积神经网络常应用于计算机视觉、分类等领域上。

2.2.2 全卷积神经网络

传统的卷积神经网络只能整幅图像进行分类, 也就是说只能解决图像识别问题。全卷积神经网络(FCN)与经典的 CNN 在卷积层之后使用全连接层得到固定长度的特征向量进行分类(全连接层 + softmax 输出)不同, FCN 可以接受任意尺寸的输入图像, 采用反卷积层对最后一个卷积层的 feature map 进行上采样, 使它恢复到输入图像相同的尺寸, 从而可以对每个像素都产生了一个预测, 同时保留了原始输入图像中的空间信息, 最后在上采样的特征图上进行逐像素分类, 最后逐个像素计算 softmax 分类的损失, 相当于每一个像素对应一个训练样本, 从而解决了图像分割问题。

2.2.3 残差连接(skip connect)

深度网络模型的容量和特征的判别能力会随着网络深度的加深而不断增强. 然而梯度弥散问题往往会成为训练深层次网络的障碍, 即过深的网络容易导致训练无法收敛. 针对这个问题, 深度残差网络提出了残差连接, 在保证网络之间的数据流通顺畅的前提下, 避免了因梯度消失导致的欠拟合问题, 从而在加深网络层次的同时有效提高了模型的表达能力。

假设 $H(x)$ 表示多层神经网络在输入样本 x 后的对应输出。根据神经网络理论, $H(x)$ 可以拟合任意函数。假设输入和输出维度一致, 则可令 $F(x) = H(x) - x$ 为残差函数。由于 x 是输入图像, 可以验证 $F(x)$ 拟合与 $H(x)$ 拟合的目标是等价的。在此条件下, 原始函数可表述为

$$H(x) = F(x) + x. \quad (2.1)$$

如图2所示, 对于输入样本 x , 通过若干残差层后的输出与原始 x 相加, 可以得到网络的输出, 具体定义如下:

$$y = F(x, \{W_i\}) + x. \quad (2.2)$$

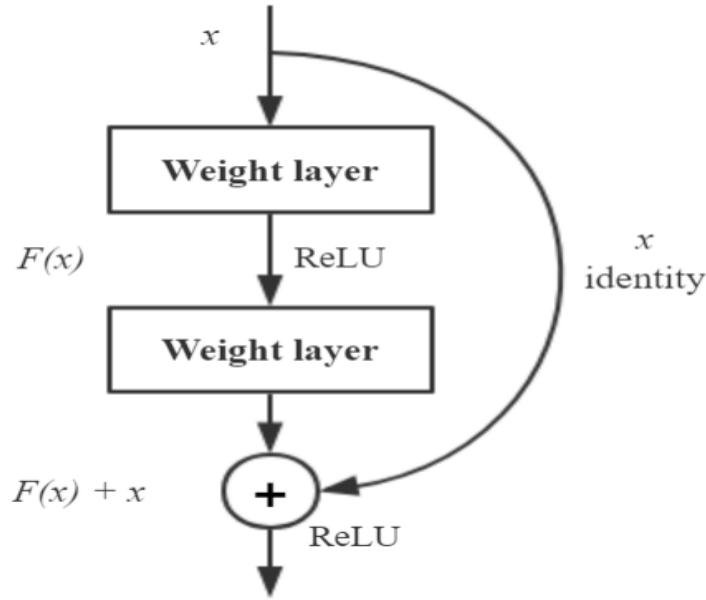


图 2: 残差连接

其中 y 是网络的输出向量; $F(x, \{W_i\})$ 是待学习的残差映射。在图2 的双层网络问题中, 输出层可表示为 $F = W_2\sigma(W_1x)$, 其中 σ 为 ReLU激活函数。与常规深度网络最大的差别是, 残差网络利用残差连接, 实现了输入与输出之间的数据叠加。在上述的连接方式中, 网络的参数数量和计算复杂度并没有增加。换言之, 残差结构可以应用到各种已有的模型中, 而不改变网络已有的架构。

由于网络输入和输出向量维度一致的约束难以满足大多数应用场景. 因此, 可以在残差的计算过程中增加一个线性映射 W_s , 以满足维度变化的需求, 具体如下:

$$y = F(x, \{W_i\}) + W_s x. \quad (2.3)$$

不难验证, 当 W_s 是方阵时, 式2.3可以退化成输入输出维度相同的条件。残差映射 F 有多种结构可以选择, 研究表明可在 2 到 3 层网络后, 通过残差连接叠加输入样本可以取得可靠的特征提取效果。

2.3 建立模型

在岩石样本识别分类过程中, 我们基于经典的VGG16网络和ResNet-50残差神经网络分别建立了相应的识别分类模型。

2.3.1 基于改进VGG的岩石样本识别分类模型

在图像识别领域广泛运用的深度学习模型为卷积神经网络, 卷积神经网络利用图像的空间相关性提取图像的轮廓信息, 提高了网络的学习能力。卷积神经网

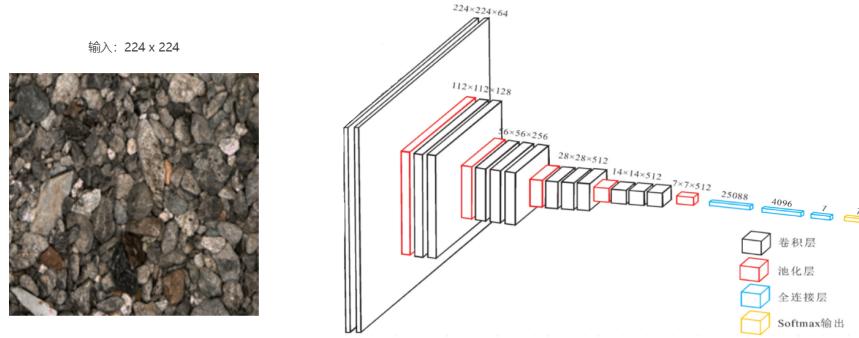


图 3: 基于改进VGG的岩石样本识别模型

络经典模型有LeNet、VGG [3,4] 等

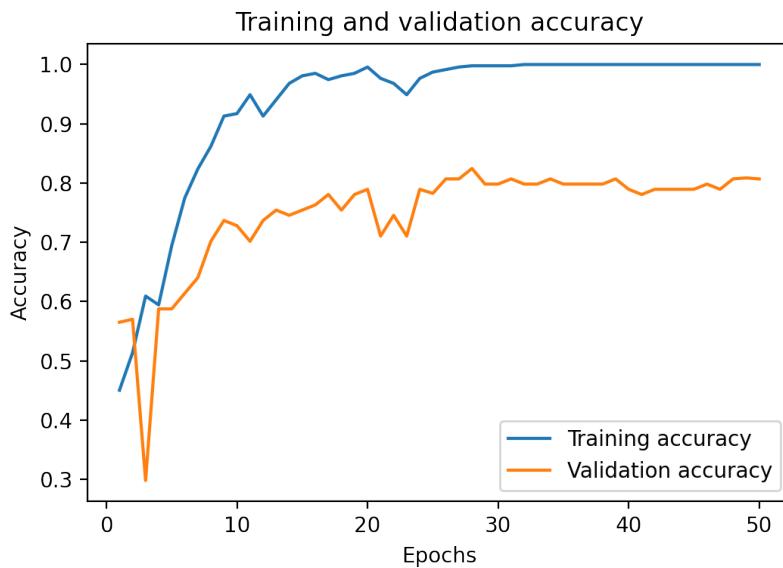


图 4: 岩石样本分类训练过程准确率图(VGG)

考虑到岩石样本图像较大，本文岩石样本分类实验采用稳定且高性能的VGG模型。岩石样本分类的VGG模型结构如图3所示。输入岩石样本图像大小为 224×224 像素，岩石图像对应R、G、B三个颜色通道，因此把一个样本图像图像数字化为一个 $224 \times 224 \times 3$ 的矩阵。VGG模型共有5组卷积和池化的迭代过程，5次卷积组的参数不尽相同。在进行完卷积和池化的迭代操作过后，原 $224 \times 224 \times 3$ 的数字图像矩阵已经缩小至 $7 \times 7 \times 512$ 的数字矩阵。将缩小后的矩阵中的数据依次取出排成一个向量，为了减少训练测试，使模型训练更快速，我们将原本进行3次全连接操作的改为只进行2次，第一个全连接层的隐含节点数为4096，第二个全连接层的隐含节点数为7，最后利用softmax函数输出分类概率，确定对应的

岩石薄片类别。

2.3.1.1 实验过程及结果

将数据预处理操作后的岩石样本图片，输入到建立的基于改进VGG的岩石样本识别模型进行训练，设置初始的学习率为 0.0001，随机种子为 520，模型通过随机参数初始化开始训练。图4为岩石样本分类训练过程准确率图。

经过 50 轮迭代后，训练集在模型上的准确率为 100%，验证集在模型上的准确率为 82.456%。

为了得到岩石样本识别精度更高的模型，我们采用训练参数更少的ResNet网络，建立了基于改进 ResNet-50 的岩石样本识别分类模型。

2.3.2 基于改进ResNet-50的岩石样本识别分类模型

为了得到岩石样本识别精度更高的模型，我们采用训练参数更少学习能力更强的ResNet网络，建立了基于改进 ResNet-50 的岩石样本识别分类模型。

layer name	output size	18-layer	34-layer	50-layer	101-layer
conv1	112×112			7×7, 64, stride 2	
				3×3 max pool, stride 2	
conv2_x	56×56	$\left[\begin{array}{c} 3 \times 3, 64 \\ 3 \times 3, 64 \end{array} \right] \times 2$	$\left[\begin{array}{c} 3 \times 3, 64 \\ 3 \times 3, 64 \end{array} \right] \times 3$	$\left[\begin{array}{c} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{array} \right] \times 3$	$\left[\begin{array}{c} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{array} \right] \times 3$
conv3_x	28×28	$\left[\begin{array}{c} 3 \times 3, 128 \\ 3 \times 3, 128 \end{array} \right] \times 2$	$\left[\begin{array}{c} 3 \times 3, 128 \\ 3 \times 3, 128 \end{array} \right] \times 4$	$\left[\begin{array}{c} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{array} \right] \times 4$	$\left[\begin{array}{c} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{array} \right] \times 4$
conv4_x	14×14	$\left[\begin{array}{c} 3 \times 3, 256 \\ 3 \times 3, 256 \end{array} \right] \times 2$	$\left[\begin{array}{c} 3 \times 3, 256 \\ 3 \times 3, 256 \end{array} \right] \times 6$	$\left[\begin{array}{c} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{array} \right] \times 6$	$\left[\begin{array}{c} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{array} \right] \times 23$
conv5_x	7×7	$\left[\begin{array}{c} 3 \times 3, 512 \\ 3 \times 3, 512 \end{array} \right] \times 2$	$\left[\begin{array}{c} 3 \times 3, 512 \\ 3 \times 3, 512 \end{array} \right] \times 3$	$\left[\begin{array}{c} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{array} \right] \times 3$	$\left[\begin{array}{c} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{array} \right] \times 3$
	1×1			average pool, 1000-d fc, softmax	
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9

图 5：常见的ResNet网络结构

图5为常见的几种ResNet的网络结构。本文采用其中最常用的ResNet50卷积部分conv1至conv5_x作为模型的骨干网络，其后接GlobalAveragePooling层然后连接一个含有7个神经元的隐层作为输出层，最后经过softmax函数输出各个类别对应的分类，最终结构如图6。

在搭建网络后，本文基于迁移学习的思想，载入ResNet50在ImageNet数据集上的模型权重，ImageNet是一个1000分类的大型开源数据集，该模型权重能够提取到1000分类的特征，因此具有良好的特征提取能力，有望在本文的岩石需求上同样提取到有效特征。但载入模型权重后，为了避免重新训练的过程打乱了模型权重的有效性，我们将载入权重后的ResNet网络中的BatchNormalization层和Dropout层以外的结构冻结，使BatchNormalization层学习到岩石数据集的通道均值，但不破坏具有特征提取作用的卷积结构。

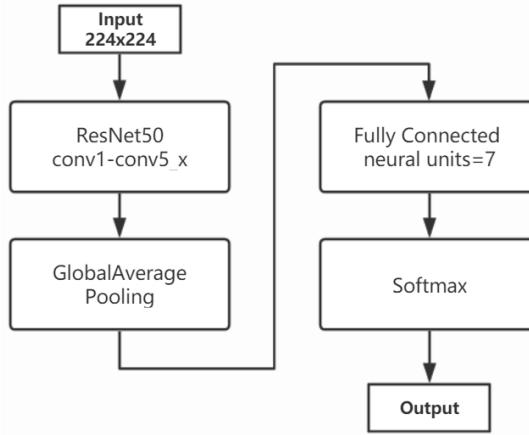


图 6: 基于ResNet50的岩石识别网络

2.3.2.1 实验过程及结果

将数据预处理操作后的岩石样本图片，输入到建立的基于ResNet50的岩石样本识别模型进行训练，设置初始的学习率为 0.001，随机种子为 520，载入ImageNet权重并按上述方法冻结卷积层，开始训练。图7为岩石样本分类训练过程准确率图。

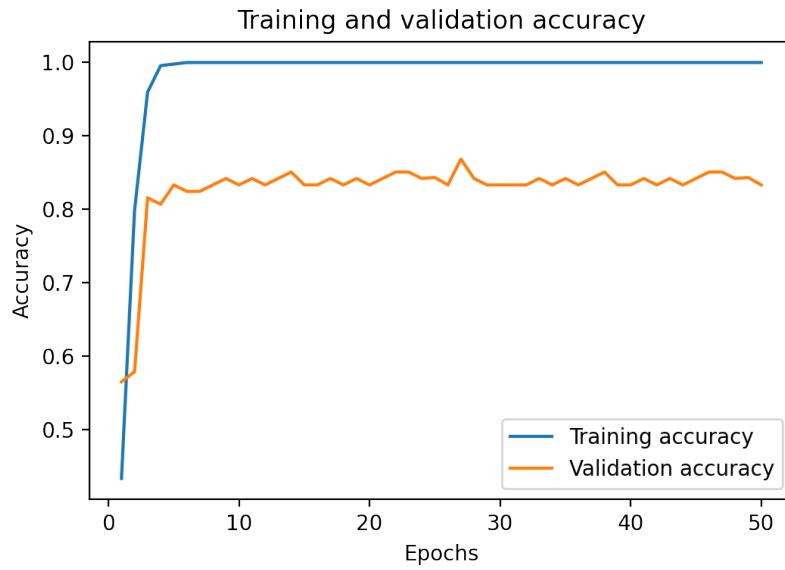


图 7: 岩石样本分类训练过程准确率图(ResNet-50)

经过 50 轮迭代后，训练集在模型上的准确率为 100%，验证集在模型上的准确率为 86.842%。与改进的VGG网络相比，准确率有所提升。部分识别结果可视化如图8所示。

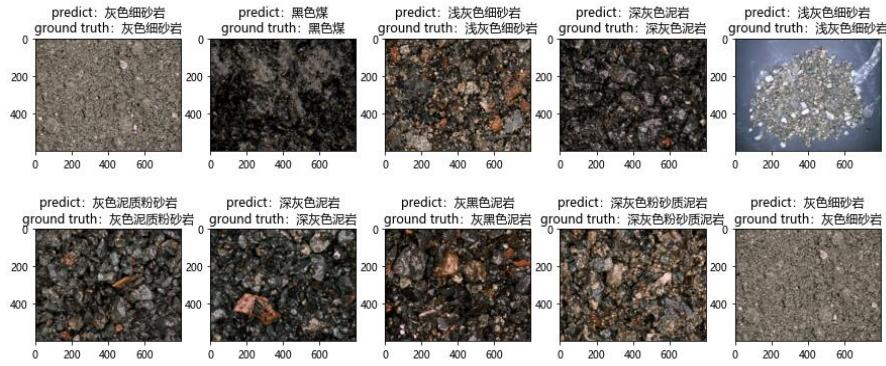


图 8: 部分识别结果可视化

2.4 模型效果评价与优化

2.4.1 模型优点

本文的模型拥有优秀的分类效果，先进性得采用迁移学习思想，并对网络结构进行冻结，能够通过最新方法寻找到最优的初始学习率，模型训练自动化程度高。在模型设计上具有很强的鲁棒性，对数据输入没有强约束，能够适应多种情景下的图像数据输入。

2.4.2 模型尚存缺陷

与深度学习其他任务情况而言，依然具有一定的进步空间，样本不平衡处进行简单图像复制增强，较难产生更多新的特征以保证泛化性能的进一步增强。图像简单放缩至 224×224 ，下采样幅度较大，容易损失岩石纹理信息等细节。

2.4.3 优化方向

为了更进一步提高模型分类能力与模型泛化能力，将做以下后续优化。

1. 通过随机裁剪图片进行小样本类别数据增强，以此实现样本平衡的同时保留图片细节；

2. 通过将 4096×3000 的图片随机裁剪拆分成 300×300 、 600×600 、 1000×1000 等多个规格的图片后放缩至 224×224 保存生成新的数据图片，数据增强的同时希望模型学习到多尺度的特征；

3. 将骨干网络输出映射后的特征向量归一化到超球面上，并运用ArcFace损失函数，通过增大类间夹角距离，减小类内夹角距离以提高分类准确率与模型泛化性能。

3 含油量分析

3.1 建立模型

岩石含油量百分比 P 可由式 3.4 计算。

$$P = Q_{\text{oil}} / Q_{\text{rock}} \quad (3.4)$$

其中, Q_{oil} 为含油量, Q_{rock} 为岩石总量, 在基于图像的含油量分析意义下, 对于给定高为 m , 宽为 n 的三通道图像矩阵 I ,

$$\begin{aligned} Q_{\text{oil}} &= \sum_{i=1}^m \sum_{j=1}^n M_{\text{oil},ij}, & M_{\text{oil}} &= \mathcal{M}_{\text{oil}}(I), \\ Q_{\text{rock}} &= \sum_{i=1}^m \sum_{j=1}^n M_{\text{rock},ij}, & M_{\text{rock}} &= \mathcal{M}_{\text{rock}}(I), \end{aligned}$$

其中, $M_{\text{oil},ij}, M_{\text{rock},ij} \in \{0, 1\}$, 含油掩码矩阵 M_{oil} 由图像 I 经过映射 \mathcal{M}_{oil} 得到, 映射 \mathcal{M}_{oil} 将图像 I 中含油像素在掩码矩阵 M_{oil} 中的对应元素置为 1, 其他元素置为 0; 岩石掩码矩阵 M_{rock} 由图像 I 经过映射 $\mathcal{M}_{\text{rock}}$ 得到, 映射 $\mathcal{M}_{\text{rock}}$ 将图像 I 中岩石像素在掩码矩阵 M_{rock} 中的对应元素置为 1, 其他元素置为 0. 因此对掩码矩阵所有元素求和可得图像中含油像素和岩石像素的数量, 即含油成分在图像中的面积和岩石在图像中的面积。

3.1.1 含油像素分割映射 \mathcal{M}_{oil}

根据含油像素的特点制定查找含油像素的规则, 即可实现对图像中含油像素的分割, 将该分割算法作为映射 M_{oil} 的实现, 即可将图像中的含油像素标记为 1。

3.1.1.1 基于灰度的含油像素分割

含油像素在荧光照射下表现为黄色至绿色, 在图像中较为明显, 在视觉感受下常常比其他像素更亮, 即灰度较大。因此可以对图像的灰度表示进行阈值分割, 将亮度大于某一阈值的像素判断为含油像素, 实现映射 \mathcal{M}_{oil} 。

根据人眼对不同颜色分量的敏感度不同, 颜色的 RGB 分量与灰度 z 的关系为

$$z = 0.299R + 0.587G + 0.114B,$$

其中, R, G, B 分别为像素的红色、绿色、蓝色分量。

3.1.1.2 基于HSV颜色空间的含油像素分割

在特定的颜色空间中，一种颜色由该空间中的一个点及其坐标表示，RGB颜色空间便是最经典的颜色空间，其颜色与坐标的关系如图9所示。可见，该模型下黄绿色不易简单得由立方体表示，不利于含油像素颜色坐标范围的表示。

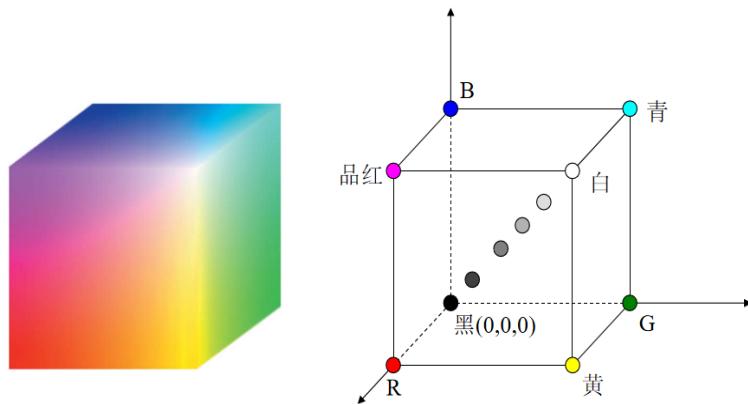


图 9: RGB颜色模型

HSV颜色模型是人们用来从颜色轮或调色板中挑选颜色(例如颜料或墨水)时使用的彩色模型之一，这个颜色系统比 RGB 系统更接近人们的经验和对彩色的感知。HSV颜色模型用色调(Hue)、饱和度(Saturation)、明度(Value)三个参数描述颜色特征 [6]。

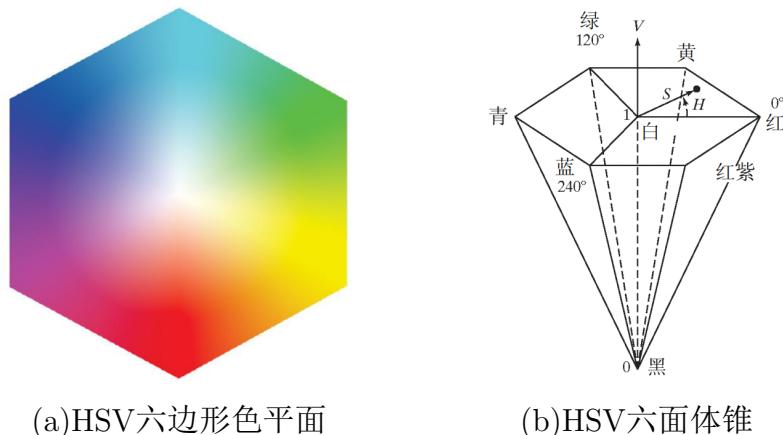


图 10: HSV颜色模型

HSV彩色空间可以通过RGB 彩色立方体的灰度轴(连接黑色顶点和白色顶点的轴)表达，垂直该轴得如图10(a)所示的六边形表示的彩色调色板，RGB立方体延伸可得如图10(b)所示的颜色六面体锥。当我们沿着图10(b)中的垂直轴(灰)轴移动时，色调分量为围绕彩色六边形的角度，通常将红轴作为参考(0°)轴。饱和

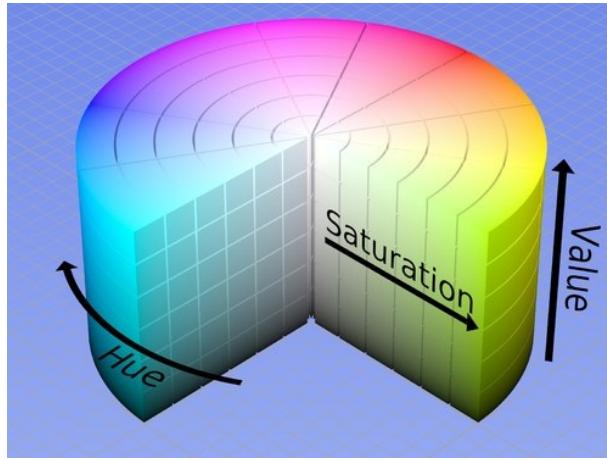


图 11: HSV圆柱模型 [8]

度(颜色的纯净度)分量由距灰度轴 V 轴的距离度量。明度分量由沿 V 轴距离锥体黑色端点的距离度量。

在应用中, 将上述的锥体HSV模型转换到如图11所示的圆柱坐标系下, 式3.5为RGB模型到圆柱HSV模型的计算式。

$$V = \max(R, G, B)$$

$$S = \begin{cases} \frac{V - \min(R, G, B)}{V}, & V \neq 0 \\ 0, & V = 0 \end{cases}, \quad (3.5)$$

$$H = \begin{cases} 60(G - B)/(V - \min(R, G, B)), & V = R \\ 120 + 60(B - R)/(V - \min(R, G, B)), & V = G \\ 240 + 60(R - G)/(V - \min(R, G, B)), & V = B \end{cases}$$

当上式计算得到的 $H < 0$ 时, 令 $H \leftarrow H + 360$, 其中 \leftarrow 表示将符号右侧赋值给符号左侧并更新取值, 此时 $0 \leq V \leq 1, 0 \leq S \leq 1, 0 \leq H \leq 360$ 。最终, 当所给图像位深为8比特时, 做如下变换使HSV模型与RGB模型定义域相适应 [7]。

$$V \leftarrow 255V, \quad S \leftarrow 255S, \quad H \leftarrow H/2. \quad (3.6)$$

此时分别给出 H, S, V 的上下限即可对HSV模型切割出一个规则的扇形体后, 然后根据图像中像素的HSV坐标是否在扇形体内判断该像素是否为含油像素。

3.1.2 岩石分割映射 $\mathcal{M}_{\text{rock}}$

由于bmp格式的图片整张图片都是岩石, 因此其掩码矩阵为与图片同大小的全 1 矩阵, 分割映射 $\mathcal{M}_{\text{rock}}$ 只需要简单得将整张图片映射为同大小的全 1 矩阵即可。

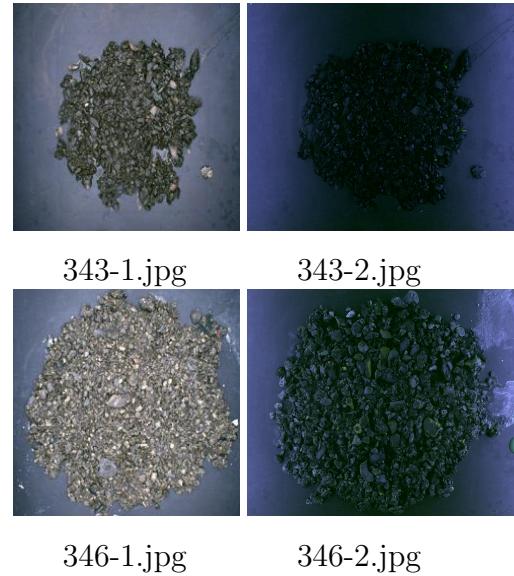


图 12: 正常光条件下的 jpg 图片(左)与荧光条件下的 jpg 图片(右)

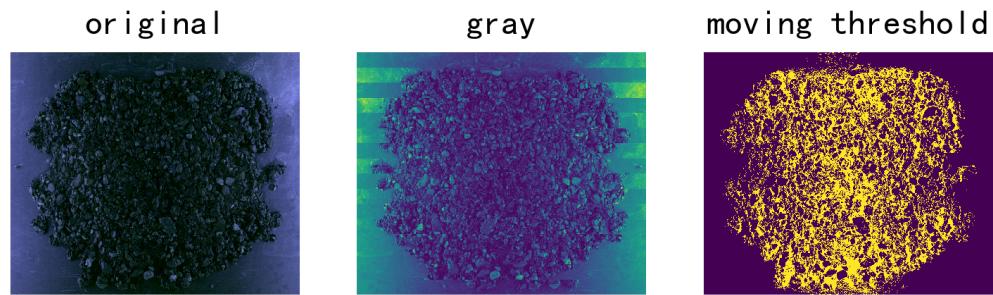


图 13: 349-2.jpg移动平均阈值分割效果图

图12为两对在荧光条件和正常光条件下的 jpg 格式图片，观察该图可知，对于 jpg 格式的图片，石头所在位置不统一，覆盖面积不统一，背景不统一，亮度不统一，且出现光照不均匀的情况，因此要对图片进行处理从而实现岩石部分的分割。

因为正常光下图像背景颜色不一、岩石样本颜色更为多样，即正常光下 jpg 图片更为复杂，因而岩石分割算法考虑基于荧光照射条件进行算法设计。

为了解决荧光照射下的光照不均，尤其是中间与图片上下边缘暗，两边明亮与中间明亮的问题，我们将图像映射成灰度图像后使用移动平均阈值分割来减小光照不均的影响。

3.1.2.1 移动平均阈值分割

移动平均阈值分割是以沿一幅图像的扫描行计算移动平均为基础的。为减少光照偏差，扫描以Z字形模式逐行执行。记 z_{k+1} 为图像展平后扫描序列中遇到的

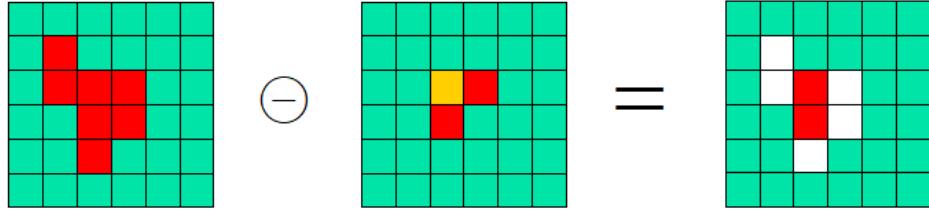


图 14: 腐蚀示意图

第 $k+1$ 个点的灰度。该点的移动平均定义为:

$$\eta(k+1) = \begin{cases} \frac{1}{l} \sum_{i=k+2-l}^{k+1} z_i, & k \geq n-1 \\ \eta(k) + \frac{1}{l}(z_{k+1} - z_{k-l}), & k \geq l+1 \end{cases} \quad (3.7)$$

式中, l 是用于计算平均的点数, $\eta(1) = z_1$. 此时用 $c\eta_{ij}$ 实现分割, η_{ij} 是点 (i, j) 处计算的移动平均, c 为参数。

图13为349-2.jpg经过移动平均阈值分割后的效果。从图13中可见, 分割出的像素点几乎都是岩石上的像素点, 效果良好, 但是图片上方边缘附近依然可见有错判的像素, 且像素之间并不连通, 而岩石在图像中本身是连续的, 因此我们需要对移动平均阈值分割的结果进行处理并打通岩石像素。

3.1.2.2 结合高斯滤波的形态学处理

3.1.2.2.1 数学形态学

数学形态学操作常常用于连通像素或断开较弱的连通域。因此本文也利用形态学操作处理移动平均阈值分割的结果来实现像素连通, 达到分割岩石的目的。

设有两幅二值图象 I, I_s 。若 I 是被处理的对象, 而 I_s 是用来处理 I 的, 则称 I_s 为结构元素(Structure element)。结构元素通常都是一些比较小的图象。数学形态学以几何学为基础对图像进行分析, 基本思想是用一个结构元素作为基本工具来探测和提取图像特征, 看这个结构元素是否能够适当有效的放入图像内部。基本操作是腐蚀(erode)与膨胀(dilate)。

I 用 I_s 腐蚀的结果是所有使 I_s 中心平移至点 (i, j) 后仍在 I 中的 (i, j) 的集合。换句话说, 用 I_s 来腐蚀 I 得到的集合是 I_s 完全包括在 I 中时 I_s 的中心位置的集合, 用公式表示为 [9]:

$$I \ominus I_s = \{(i, j) | I_{s,(i,j)} \subseteq I\},$$

其中, $I_{s,(i,j)}$ 表示 I_s 以点 (i, j) 为中心。该运算示意图如图14, 图中橙色像素表示结构元素的中心。

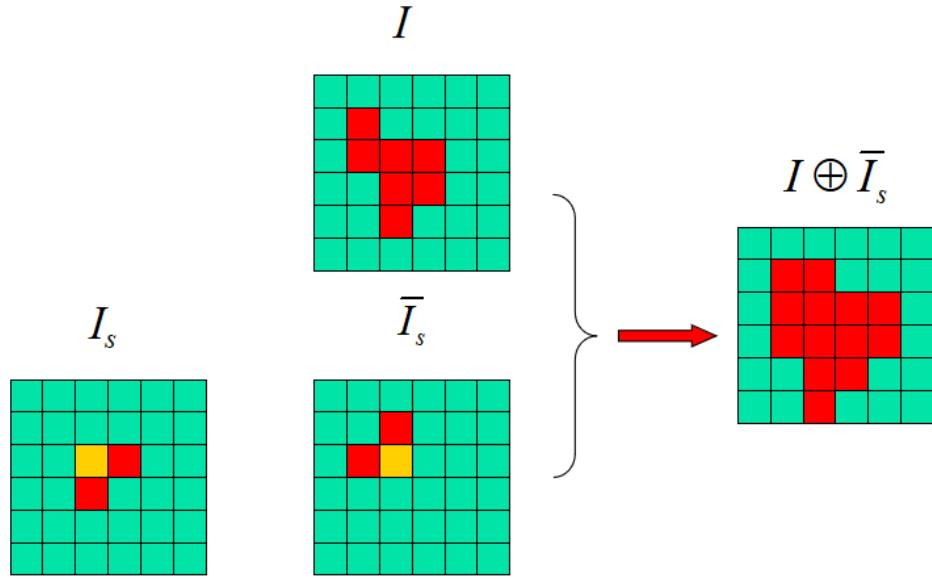


图 15: 膨胀示意图

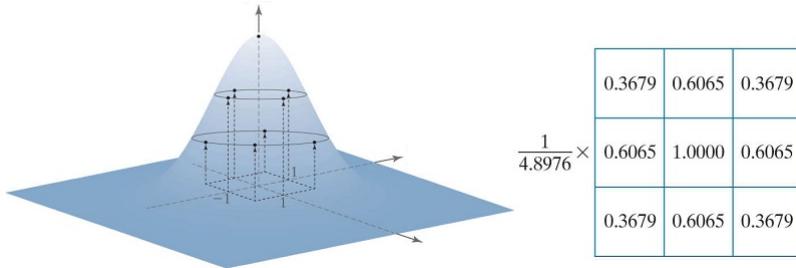


图 16: 高斯核示意图

膨胀可以看做是腐蚀的对偶运算，膨胀是结构元素 I_s 将中心平移至点 (i, j) 后，关于这个中心 (i, j) 翻转，使得翻转后的结构元素 $\bar{I}_{s,(i,j)}$ 与翻转前关于中心 (i, j) 对称，若 $\bar{I}_{s,(i,j)}$ 与待处理图像 I 仍有交集，则该点放入膨胀的结果集合里，运算示意图如图15，公式表示如下：

$$I \oplus I_s = \{(i, j) | \bar{I}_{s,(i,j)} \cap I \neq \emptyset\}.$$

选择合适形状与大小的结构元素下，腐蚀具有断开连通、减小边界的作用；膨胀则有连通像素、扩张边界的作用。先腐蚀后膨胀则能够除去孤立的小点，毛刺和图像中的连通小桥，而总的位置和形状不便，称为开运算；先膨胀后腐蚀能够填平二值图像的小孔，弥合小裂缝，而总的位置和形状不变，称为闭运算。

因此，我们对移动平均阈值分割的结果先进行了闭运算，后进行了开运算，

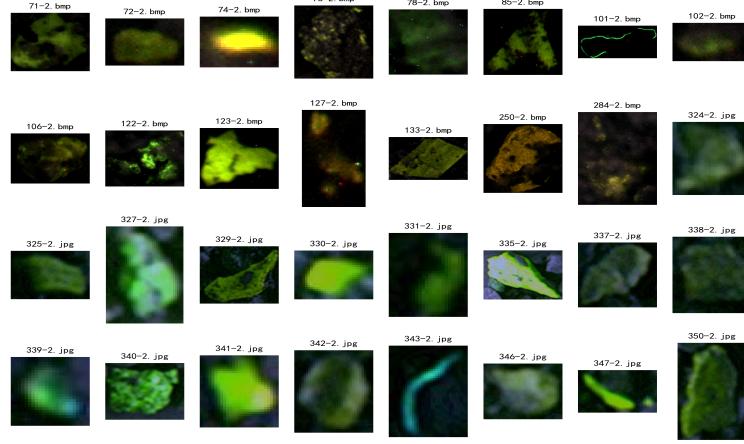


图 17: 含油碎片

以此达到连通孔洞，保证边界的效果。但是结构元素大小是不易于设置的一个参数，在闭运算期间，结构元素较大，则会边界扩张严重，往往将不是岩石像素的点也判断为连通的岩石成份；而结构元素较小时，在岩石像素点较分散的情况下，无法达到连通像素的目标。

3.1.2.2.2 高斯滤波

高斯滤波是平滑图像时非常重要的一种方法，该滤波器的卷积核的权重按式3.8所示的二维高斯分布生成，并乘以一个归一化系数 K 使卷积核模板求和为1，保证作用像素邻域卷积运算后的非递减。

$$K e^{-\frac{d^2}{2\sigma^2}}, \quad (3.8)$$

其中， d 表示卷积核内一点到卷积核中心的距离， σ 为二维高斯分布的标准差，控制核的中心权重集中程度。

图16展示了二维高斯分布(左)及一个高斯核的例子。

高斯滤波通过调整 σ ，实现不同的权重设置，设置较大的卷积核时，作用像素点能够考虑邻域内点的密度，并在作用像素点生成受其影响的一个连续数值，连续数值相比形态学的离散二值取值具有更多的灵活性与更高的精度，设定较好的阈值后，将卷积结果较大的像素点置为 1，其他置为 0，则可在原有分割基础上，进行扩张的进一步平滑与分割。这样通过将该高斯滤波分割作为移动平均阈值分割的后续步骤、形态学处理的前置步骤，增加了移动平均阈值分割所得像素点，能够有效提高岩石像素连通的效果。

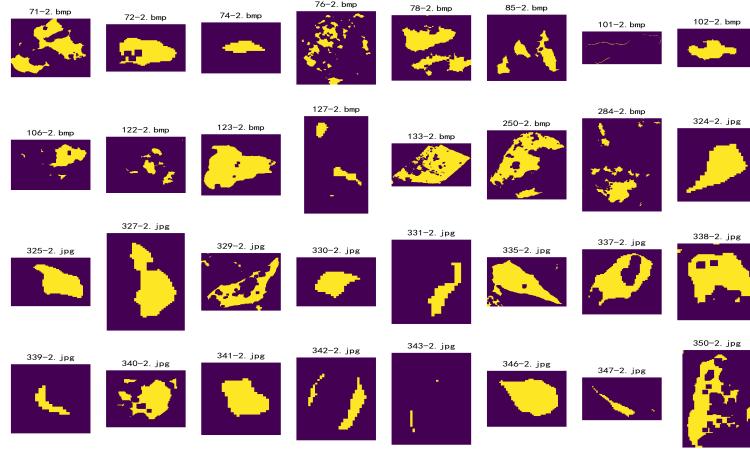


图 18: 含油碎片的分割结果

3.2 算法设计与实验

3.2.1 含油像素分割

通过找出岩石含油部分在荧光照射下的颜色像素，对像素进行统计得到含油部分此时应有的颜色范围，并作为后续的含油像素查找条件。

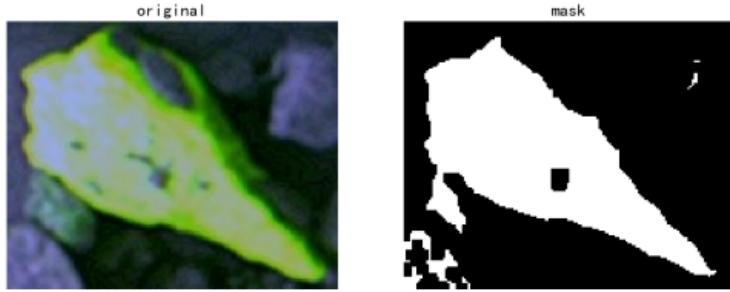


图 19: 样例图片

为了找出颜色范围，我们从岩石样本中筛选具有代表性颜色的含油部分，裁剪得到含油碎片。如图17所示。

实际上，裁剪碎片中或含有岩石底纹或含有非含油像素，还需要通过分割实现对含油部分的提取。从筛选可见颜色的亮度不一，显然不易通过灰度单一阈值实现二值化分割，但通过大津法分割能对大多数含油碎片进行分割，分割结果可能含有一定误差，但分割主体缺少部分含油像素是可以接受的，依然可以找出主要的颜色范围，而如果分割结果含有非含油像素，则可能对后续颜色范围的计算产生较大错误，因此在大津法分割后对分割结果进行核为 5×5 的腐蚀处理。

图18为对灰度图做大津法分割联合形态学腐蚀的结果。

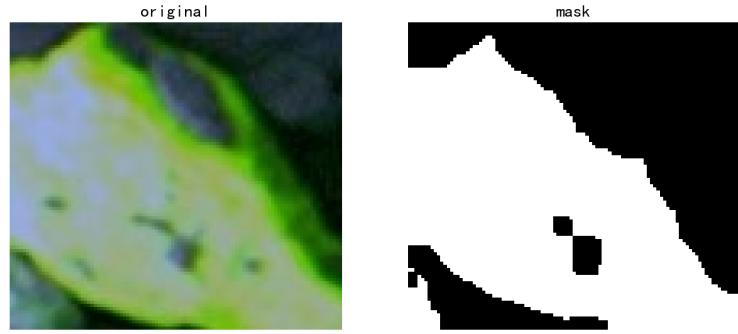


图 20: 裁剪后的样例图片

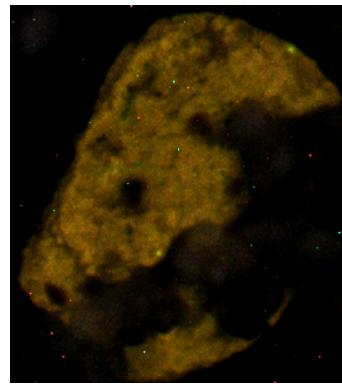


图 21: 250-2.bmp 噪声点

部分依然含有非含油像素的样本如图19所示，对这部分分割不佳的图像再次进行手动裁剪，去除非含油部分，得到图20，使含油像素筛选结果得到保证。

在计算含油像素的颜色范围时，除了非含油像素的影响外，图片还含有许多噪点，例如图21所示，由 250-2.bmp 裁剪得到的含油碎片中明显可见红绿蓝的噪点，采用对像素范围求最大值与最小值会受到噪点的影响，因此对含油碎片求均值作为该碎片的颜色特征是最好的方式。

计算各个碎片中含油部分的HSV颜色均值，最终得到各个碎片的均值颜色特征如表1。

计算得到HSV范围上下限，如表2所示。

然后读取所给荧光照射图片并转换至HSV颜色空间后，判断每个像素点是否符合含油像素范围并生成含油掩码矩阵 M_{oil} .

然后对掩码二值图像采用 5x5 的矩形作为结构元素做闭运算操作，得到如图22的含油像素分割结果。可以看到，分割结果良好，对目标颜色敏感，较亮与较暗的目标颜色均能分割出来。

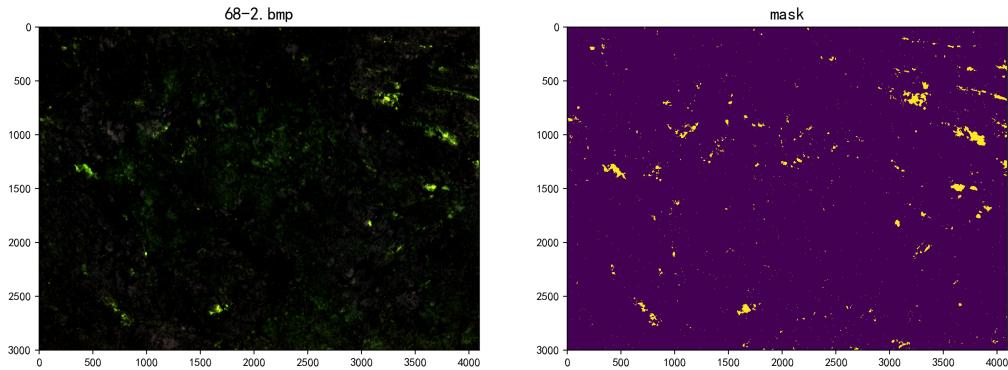


图 22: 含油像素分割结果

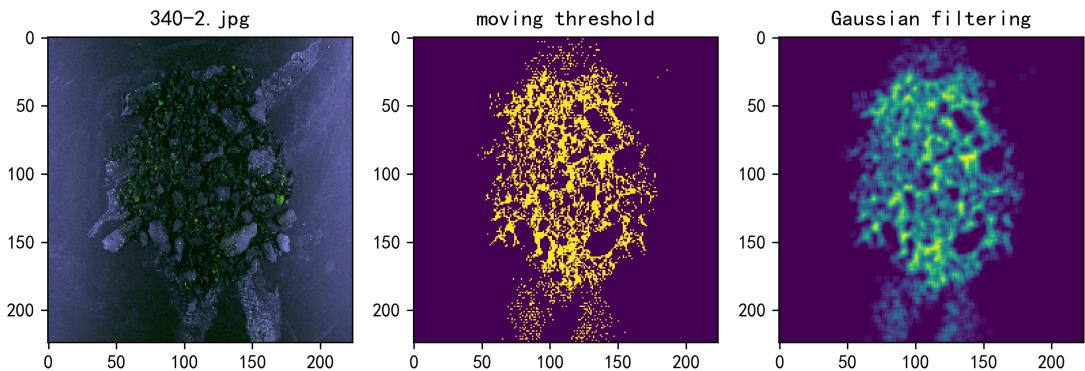


图 23: 移动平均阈值分割与高斯滤波

3.2.2 岩石分割

对于 bmp 格式图片，将岩石分割结果置为与图片同大小的全 1 矩阵。

对于 jpg 格式图片，读取高为 m 宽为 n 的荧光照射图片后，为了加快计算和便于参数调整，将图像放缩至 224x224。然后计算灰度图 Y ，逐行扫描得到 z 。

对图片进行移动平均阈值分割，其中， $c = 0.4, l = 224$ ，低于阈值 $c\eta$ 即为岩石候选像素点。为了增多岩石像素点方便形态学闭运算连通像素，对分割进行高斯滤波，卷积核大小为 5x5， $\sigma = 5$ ，记所得矩阵为 I_{g1} 。

但有时会得到如图23所示的不理想平滑结果。观察容易发现，在移动平均阈值分割时依然将一些较暗的背景像素置为 1，导致高斯滑动滤波的结果也不够理想。但这也正体现出高斯滤波的好处，高斯滤波的结果由作用像素点邻域内值为 1 的像素点的密度决定，而背景点时少数误判的点，在观察多张图像后我们提出如下解决方案缓解该问题：

1. 将 I_{g1} 的第 50 行第 60 至 170 列中小于 \bar{I}_{g1} 的像素值置为 0， \bar{I}_{g1} 为 I_{g1} 所有像素的均值。

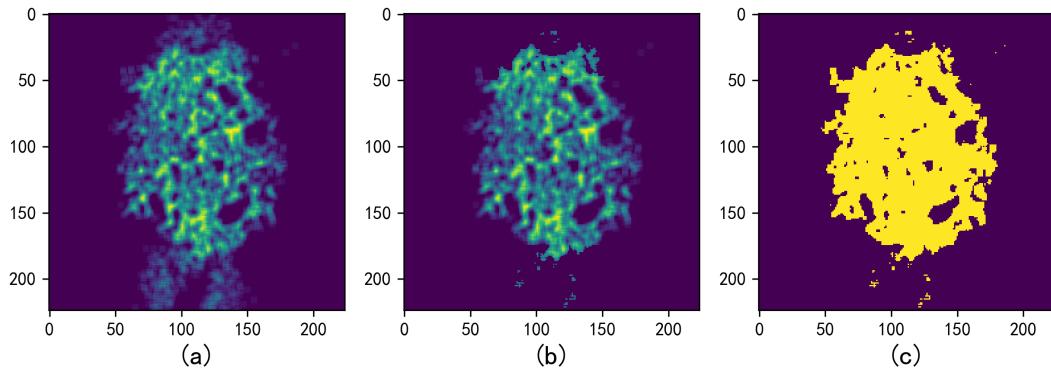


图 24: 校正高斯滤波 I_{g1} 示意图

2. 将 I_{g1} 第 175 至 224 行第 60 至 170 列中小于 \bar{I}_{g1} 的像素值置为 0。

此时解决了滤波后上下两侧分割点中含有背景的问题，而左右两侧的值虽然较小，但通常也是正确的岩石像素，此时令 I_{g1} 中大于 $0.4\bar{I}_{g1}$ 的像素点置为 1，得到校正后再次二值化的分割结果。图24为此解决方案的效果，其中，(a)是高斯滤波后的结果，(b)是步骤1-2的结果，(c)是再次二值化的分割结果。

然后经过如下组合操作实现连通像素，减少孔洞：

1. 闭运算，结构元素为 5×5 的矩形；
2. 开运算，结构元素为 3×3 的矩形；
3. 高斯滤波，卷积核大小为 3×3 , $\sigma = 1$ ，每次结果记为 I_{g2} ；
4. 将 I_{g2} 中大于 0.3 的像素值置为 1；
5. 循环 5 次步骤 3-4，提高效果；
6. 将 I_{g2} 中小于等于 0.3 的像素值置为 0。

最终得到岩石分割结果如图25，可以看见，分割效果理想，多数较易分割的样本不仅能够分割出岩石堆，还可以准确分割出孤立的岩石块，比如 329.jpg、337.jpg。

最后，将分割结果 I_{g2} 放缩回高为 m 宽为 n 的图像矩阵得到岩石掩码矩阵 M_{rock} ，与含油像素分割结果相统一。

3.2.3 含油百分比计算

对每一张图按上述算法计算岩石掩码矩阵 M_{rock} 与含油掩码矩阵 M_{oil} 后，计算含油百分比 $P = \frac{M_{\text{oil}}}{M_{\text{rock}}} \times 100\%$ 。表3为部分结果展示。

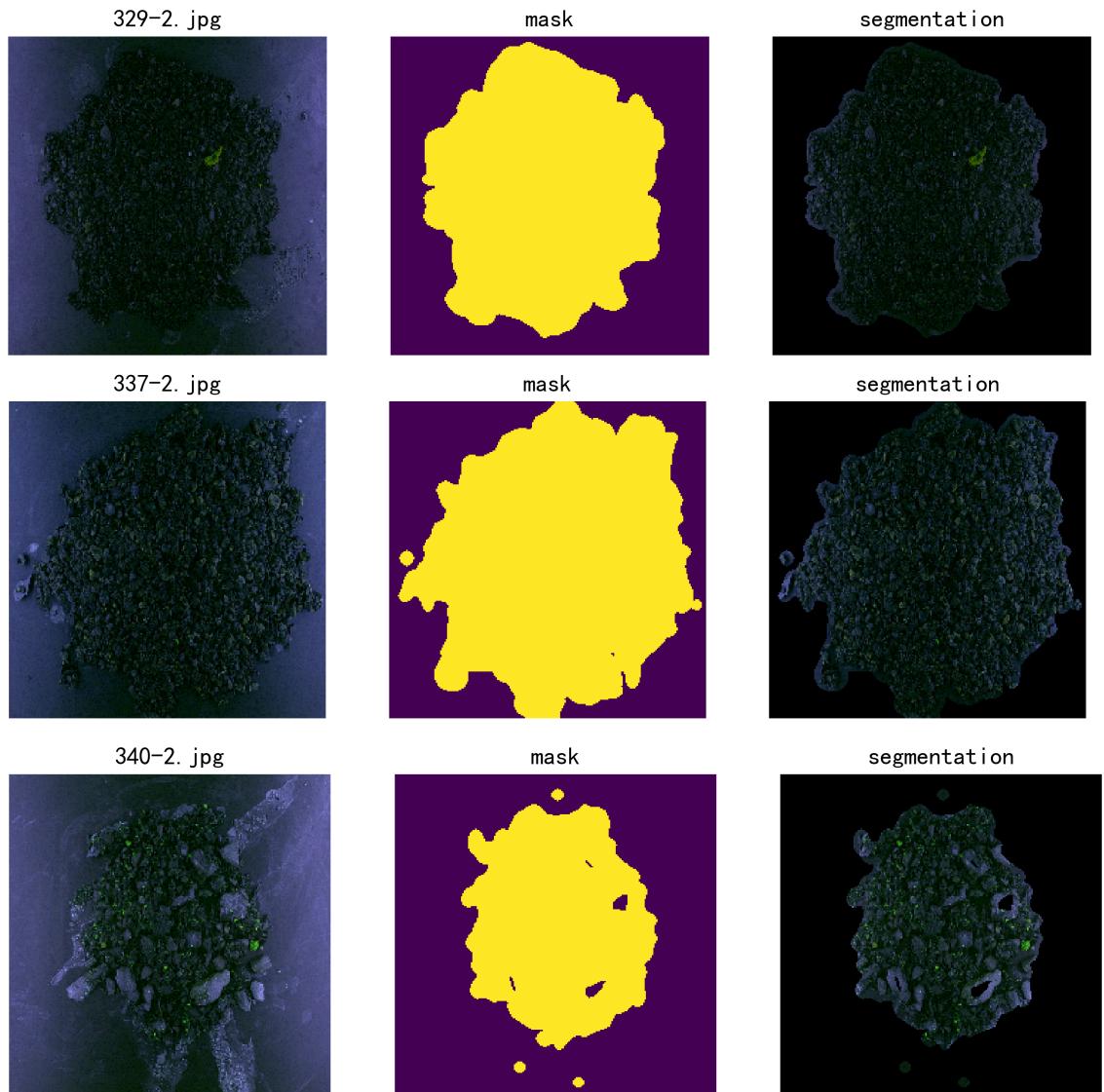


图 25: 岩石分割效果示意图

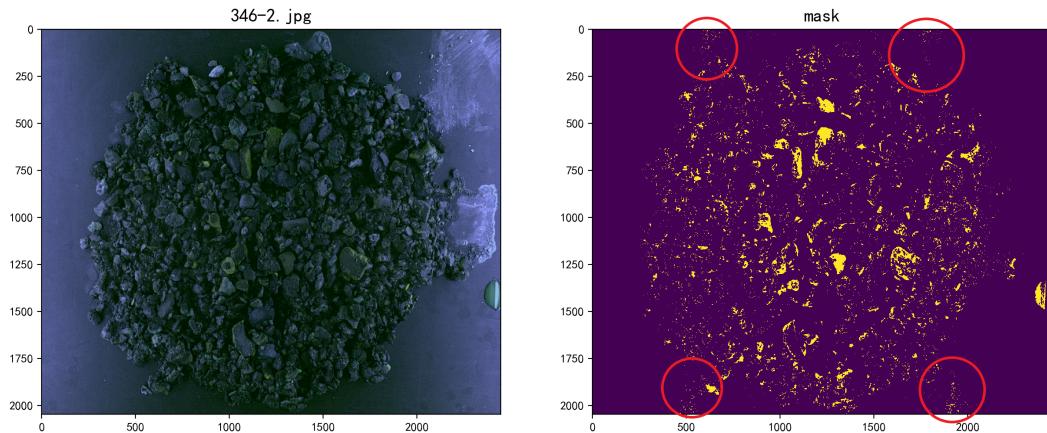


图 26: 含油像素分割反例

3.3 模型效果评价与优化

3.3.1 模型优点

- 1.此含油量模型与算法能够准确分割岩石样与含油像素； 2.自动化程度高；
- 3.考虑情况详细周到；

3.3.2 模型尚存缺陷

当条件较为复杂时依然会受到部分干扰，其中颜色范围设置较为关键，上述颜色设置不当，导致如图26红圈所示的上下两侧像素将部分背景判断为了含油像素。

经检查发现是由于从 78-2.bmp 裁剪的含油碎片的含油像素明度较低引起颜色范围的明度较低。于是删去 78-2.bmp 的含油碎片，重新利用 31 张碎片计算颜色范围，得到如表4所示的HSV颜色上下限，并将此HSV颜色范围作为最终范围。图27为优化后的含油像素分割。

3.3.3 优化方向

针对极少数样本，会出现图27中右侧边缘同样颜色但是不在岩石上的情况，后续优化方法是在岩石分割结果上进行含油掩码的计算，而非直接在图片数据上进行，从而避开岩石外像素的干扰。

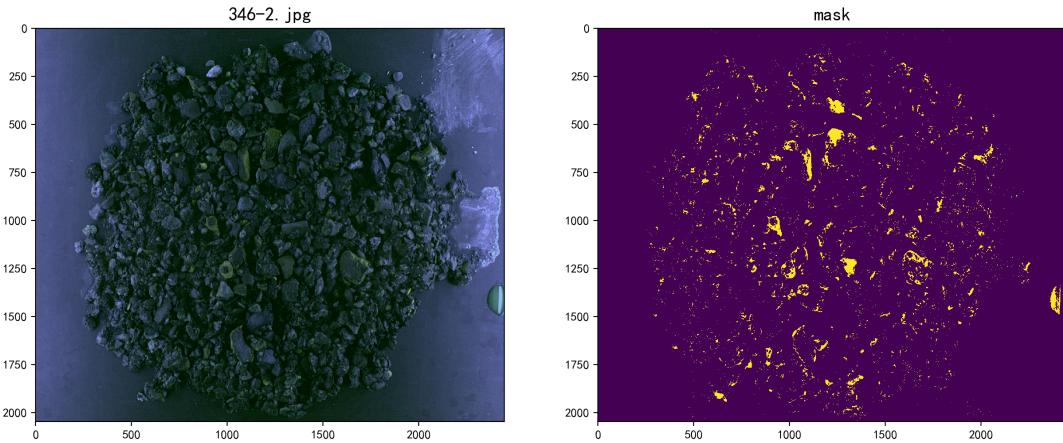


图 27: 优化后的含油像素分割反例

4 总结

本文在岩石识别问题上提出了以VGG、ResNet50为基础的岩石识别模型，对数据经过适当的数据预处理后训练模型得到了86%的分类正确率，分类效果优秀，具有实际生产意义。在含油百分比问题上，本文基于传统数字图像处理技术，对图像进行HSV空间变换、筛选典型目标颜色碎片、含油像素颜色范围计算，最终进行HSV阈值分割实现含油面积计算；通过对图像数据进行移动平均阈值分割、高斯滤波平滑、分区域阈值高斯滤波校正、形态学运算、迭代高斯滤波与阈值分割实现了岩石面积计算，最终计算得到合理的石油百分含量。

表 1: 各碎片中含油像素HSV均值

碎片样本来源	色度(Hue)	饱和度(Saturation)	明度(Value)
71-2.bmp	39.2	230.7	71.6
72-2.bmp	35.2	239.4	69.6
74-2.bmp	29.9	224.2	255.0
76-2.bmp	28.4	145.2	69.0
78-2.bmp	48.0	155.1	56.1
85-2.bmp	39.0	253.9	88.3
101-2.bmp	58.2	208.5	236.9
102-2.bmp	35.9	163.1	62.7
106-2.bmp	31.6	246.6	66.3
122-2.bmp	45.6	202.8	155.0
123-2.bmp	38.0	251.5	179.5
127-2.bmp	23.4	244.6	84.9
133-2.bmp	34.3	241.6	66.8
250-2.bmp	21.0	252.5	93.6
284-2.bmp	29.9	142.2	74.3
324-2.jpg	55.1	106.1	91.4
325-2.jpg	51.5	168.7	92.7
327-2.jpg	71.5	108.2	180.7
329-2.jpg	44.3	242.7	86.6
330-2.jpg	40.3	240.8	186.6
331-2.jpg	48.5	235.1	94.6
335-2.jpg	63.0	82.1	178.5
337-2.jpg	66.9	91.6	73.8
338-2.jpg	59.7	127.5	58.7
339-2.jpg	68.1	162.2	157.6
340-2.jpg	58.0	166.2	155.3
341-2.jpg	42.4	225.3	195.6
342-2.jpg	53.0	109.7	104.3
343-2.jpg	88.5	202.8	134.5
346-2.jpg	62.8	84.2	116.0
347-2.jpg	42.2	244.9	183.6
350-2.jpg	46.7	155.4	121.0

表 2: 含油像素HSV上下限

	色度(Hue)	饱和度(Saturation)	亮度(Value)
上限	21.0	82.1	56.1
下限	88.5	253.9	254.9

表 3: 含油百分比

样本编号	含油面积	岩石面积	含油百分比
1	76852	12288000	0.63%
2	5600	12288000	0.05%
3	2371	12288000	0.02%
4	128914	12288000	1.05%
5	2794	12288000	0.02%
6	3042	12288000	0.02%
8	2375	12288000	0.02%
9	4836	12288000	0.04%
10	3116	12288000	0.03%
335	11448	2291064	0.50%
336	41478	2480990	1.67%
337	42924	3191858	1.34%
338	83782	3518136	2.38%
339	679	2568005	0.03%
340	66136	1664822	3.97%

表 4: 修正后的含油像素HSV上下限

	色度(Hue)	饱和度(Saturation)	明度(Value)
上限	21.0	82.1	58.7
下限	88.5	253.9	254.9

参考文献

- [1] 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.
- [2] Shang W, Sohn K, Almeida D, et al. Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units[J]. 2016: 2217-2225.
- [3] Lecun Y, Bottou L, Bengio Y, et al. Gradient- based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [4] Zhu Y, Bai L, Peng W, et al. Depthwise Separable Convolution Feature Learning for Ithomogeneous Rock Image Classification[C]// ICCSIP, 2018, 1005: 165-176.
- [5] 白林,魏昕,刘禹,吴崇阳,陈立辉.基于VGG模型的岩石薄片图像识别[J].地质通报,2019,38(12):2053-2058.
- [6] Gonzalez R C , Woods R E , Eddins S L . Digital Image Processing Using MATLAB[J]. Digital Image Processing Using Matlab, 2010, 21(84):197-199.
- [7] OpenCV, Color conversions, https://docs.opencv.org/4.2.0/de/d25/imgproc_color_conversions.html#color_convert_rgb_hsv, 2021.5.5
- [8] 卡米雷特的小站, 关于HSV模型的探讨, <https://kamilet.cn/about-hsv/>, 2021.5.5
- [9] CSDN, 数学形态学之二值形态学（腐蚀、膨胀、开闭运算）基本原理, <https://blog.csdn.net/zxli3011/article/details/50522939>, 2021.5.6