

文章编号: 2095-2163(2023)01-0104-08

中图分类号: TP391.4

文献标志码: A

# 基于改进的 SKnet 和 Bi-GRU 的岩石薄片图像矿物识别

刘勇<sup>1</sup>, 吴晓红<sup>1</sup>, 滕奇志<sup>1</sup>, 何海波<sup>2</sup>

(1 四川大学电子信息学院, 成都 610065; 2 成都西图科技有限公司, 成都 610065)

**摘要:** 通过分析岩石薄片矿物成分, 研究储集层空间结构, 对后续的油气勘探开发具有重要意义。基于正交偏光序列图像的矿物识别研究已经取得了一些成果, 但多数方法未利用矿物颗粒在序列图中的变化信息, 本文借鉴视频分类的思想, 针对岩石矿物颗粒正交偏光序列图像, 结合岩石矿物颗粒消光性特点, 构建了卷积神经网络和循环神经网络相结合的识别模型。卷积神经网络选用 SKnet 并在此基础上增加了空间特征融合机制, 循环神经网络采用双向门控循环单元 (Bidirectional Gated Recurrent Unit, Bi-GRU) 来提取序列图像的前后关联特征。选取石英、碱性长石、斜长石、岩屑 4 类矿物颗粒序列图像构建数据集进行验证, 结果表明本文提出的矿物颗粒识别方法识别效果良好。

**关键词:** 矿物颗粒识别; 偏光序列图像; 消光性; SKnet; 双向门控循环单元

## Mineral recognition of rock thin section images based on improved SKnet and Bi-GRU

LIU Yong<sup>1</sup>, WU Xiaohong<sup>1</sup>, TENG Qizhi<sup>1</sup>, HE Haibo<sup>2</sup>

(1 College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2 Chengdu Xitu Technology Co. Ltd, Chengdu 610065, China)

**[Abstract]** By analyzing the mineral composition of rock slices, we can study the structure of reservoir space, which is of great significance to the subsequent oil and gas exploration and development. Mineral identification research based on orthogonal polarized light sequence images has achieved some results, but most methods do not use the change information of mineral particles in sequence images. In this paper, inspired by video classification, a mineral recognition model combining the convolutional neural network and recurrent neural network is constructed based on the orthogonal polarizing image sequence, which makes use of the extinction characteristics of rock mineral grains. A spatial feature fusion mechanism is added to SKnet as the convolutional neural network and the recurrent neural network is the bidirectional gated recurrent unit (Bi-GRU) to extract the sequence features of mineral grains in the polarization image sequence. We evaluate our method on the dataset containing four types of mineral grains image sequences, including quartz, alkali feldspar, anorthosite, and rock debris. The experimental results demonstrate that the proposed rock mineral recognition method in this paper achieves superior identification performance.

**[Key words]** mineral grains identification; polarization image sequence; extinction characteristics; SKnet; bidirectional gated recurrent unit

## 0 引言

传统的岩石薄片图像分析鉴定依赖于专家在光学显微镜下对岩石的目视观察, 存在很多局限性。首先, 薄片鉴定需要在多个角度反复观察单偏光图像和正交偏光图像, 大量岩石样本的识别工作需要耗费一定的人力资源和时间成本; 其次, 人工鉴定因人而异,

不能保证准确、量化的结果; 第三, 传统的鉴定依赖实物送样和人工观察, 但薄片样品长时间存放可能会变黄、脱胶, 影响鉴定结果, 最终导致资产损失。

近年来, 在矿物颗粒分类识别方面, 国内外学者开展了大量的研究工作, 也取得了一定的应用成果。2013年, Mariusz Mlynarczuk<sup>[1]</sup>等人基于最近邻算法等 4 种模式识别方法, 在 4 种颜色特征空间中对白

**基金项目:** 国家自然科学基金项目资助(62071315)。

**作者简介:** 刘勇(1995-), 男, 硕士研究生, 主要研究方向: 图像处理; 吴晓红(1964-), 女, 博士, 副教授, 主要研究方向: 图像处理、模式识别; 滕奇志(1961-), 女, 博士, 教授, 主要研究方向: 图像处理、模式识别; 何海波(1968-), 男, 学士, 高级工程师, 主要研究方向: 图像处理、计算机应用。

**通讯作者:** 吴晓红 Email: wxh@scu.edu.cn

**收稿日期:** 2022-09-22

云石、石灰岩、花岗岩等矿物进行了颗粒识别;2014年,杨宗瑞<sup>[2]</sup>结合矿物颗粒的亮度均值、颜色、纹理特征以及灰度共生矩阵熵信息,采用随机森林分类器识别多种矿物颗粒;2015年,Akkaş E 等人<sup>[3]</sup>使用 SEM-EDS 光谱仪获取矿物颗粒的特征,再利用决策树算法对橄榄石等十种矿物颗粒进行分类;2016年,闵涛<sup>[4]</sup>针对矿物正交偏光序列图的颜色、纹理等特征,采用机器学习模型,再结合复杂的权值投票机制,完成矿物颗粒识别任务。

随着计算机硬件能力的迅速提升和深度学习技术的飞速发展,基于神经网络的深度学习分类方法也相继出现。2012年,Krizhevsky A 等人<sup>[5]</sup>提出 AlexNet,在图像分类领域取得突破性进展;自此,用于分类的各种卷积神经网络层出不穷,包括 VGGNet、ResNet、MobileNet、InceptionResnet 等。2018年,彭志伟<sup>[6]</sup>选用 CaffeNet 作为识别分类网络模型,并提出一种基于颜色和纹理的序列图像筛选算法,实现了矿物颗粒偏光序列图像下的分类识别;2019年,T.Nanjo, S.Tanaka<sup>[7]</sup>等人提出了一种基于深度神经网络的碳酸盐岩岩性识别图像分析技术;2020年 Y Xu<sup>[8]</sup>等人利用卷积神经网络 ResNet-18 对变质岩、火成岩、沉积岩(包括碎屑岩和碳酸盐)的薄片图像进行自动分类;2021年,朱磊<sup>[9]</sup>提出了一种基于序列图分类的矿物颗粒识别方法,针对矿物颗粒偏光序列图像分类效果相对较好,但未充分利用岩石颗粒在不同正交偏光角度下的消光性信息,斜长石矿物颗粒存在一定的误识别情况。

综上,充分利用矿物颗粒在多个正交偏光角度下的丰富信息是矿物颗粒分类工作中十分值得探究的内容。在不同正交偏光角度下,矿物颗粒的干涉、纹理及消光特征表现不同,如图 1 所示。石英颗粒表面整体较为光滑;碱性长石表面纹理特征丰富,没有规律,整体表现出一种“不干净”的状态;斜长石颗粒在特定角度会呈现聚片双晶现象;岩屑颗粒在偏光序列图中几乎没有消光变化,表面有零星分布的碎屑,没有规律。本文在对岩石矿物颗粒偏光序列图像进行消光位校正处理的前提下,提出了基于改进的 SKnet 和双向 GRU 模型的岩石矿物颗粒分类方法,将进行了消光位校正对齐处理后的矿物颗粒偏光序列图像送入改进的 SKnet 特征提取模块进行特征提取,再通过双向 GRU 门控单元捕获序列图的前后关联特性,最终完成对常见的石英、斜长石、碱性长石、岩屑 4 类矿物颗粒的分类识别。该网络在矿物颗粒偏光序列图识别分类上取得了较好的效果。

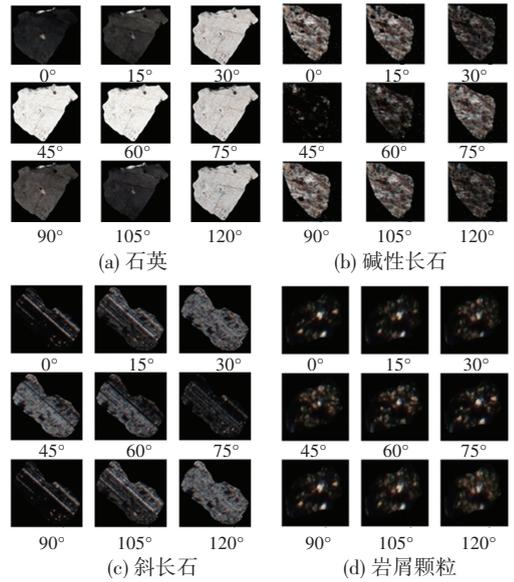


图 1 4 类岩石矿物偏光序列图

Fig. 1 Polarization sequence images of four types of rock minerals

## 1 岩石颗粒偏光序列图提取及消光位校正

### 1.1 岩石薄片正交偏光序列图像采集

在进行岩石矿物颗粒识别工作前需采集一定量的岩石薄片序列图,便于后续分割提取岩石颗粒序列图,制作实验所用训练集、测试集。将岩石样本磨制到足够“薄”,一般的标准薄片厚度为 30  $\mu\text{m}$ ,将薄片制作成玻片放置光学显微镜下,即可观察分析岩石薄片中的矿物。本文采用最新的薄片图像采集系统,在载物台不动的情况下,自动旋转上下偏振片来获取不同角度的正交偏光图像,从而实现序列图像的对齐分析,更有利于后续图像的智能识别,图像采集工作中设置的正交偏光旋转角度为 15°,每一组采集 9 张图片。在正交偏光角度下采集的岩石薄片序列图像如图 2 所示。

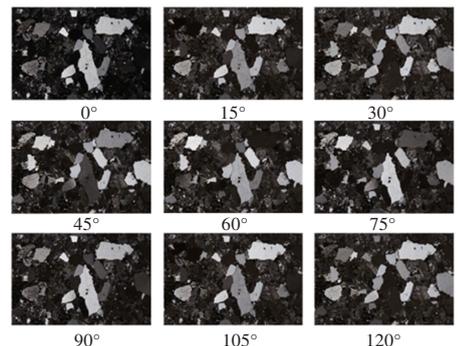


图 2 岩石薄片偏光序列图像

Fig. 2 Polarized light sequence images of rock slices

### 1.2 岩石矿物颗粒提取及消光位校正

在不同角度正交偏光镜下,由于非均质岩石薄

片除垂直光轴以外的切片,光率体为椭圆,在偏振片的旋转过程中,光率体椭圆长、短半径与上偏振片、下偏振片有4次平行的机会,平行下偏振片的偏光不能透过上偏振片,视域呈黑暗,说明该矿物颗粒处于消光位,并且岩石矿物颗粒的消光性在载物台旋转的过程中呈现以 $90^\circ$ 为周期的变化规律。在实际图像采集过程中,由于颗粒自身的一些光学特性,实际处于消光位的矿物颗粒表现为周期内的亮度最低值,并且不同矿物颗粒的消光位可能处于不同角度,分割提取的岩石颗粒偏光序列图第一张图像并不一定处于消光位,为了更好的利用后续双向GRU循环神经网络的特性,需要保证不同矿物颗粒正交偏光序列图像相同序号位置处的消光程序尽可能接近,因此需要从岩石薄片偏光序列图中提取出岩石颗粒序列图,然后对岩石颗粒序列图做消光位校正处理,步骤如下:

(1)对采集获取的岩石薄片偏光序列图像,一次读入以 $15^\circ$ 为间隔的9张正交偏光序列图,利用实例分割或者交互式分割方法,提取其中的岩石矿物颗粒偏光序列图像,如图3所示。

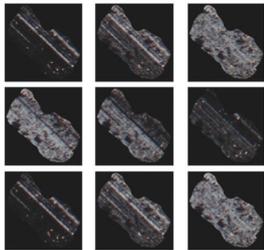


图3 利用分割方法得到的矿物颗粒序列图

Fig. 3 Mineral grain sequence map obtained by segmentation method

(2)对每一组提取出来偏光序列图像,找到序列图中处于消光位的颗粒图像,并以该图像为基准对序列图重新排序。

①把原来的9张矿物颗粒RGB序列图记为 $rgb = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9\}$ ,根据RGB值与灰度值之间的转换关系,将其转换为9张灰度序列图,记为 $gray = \{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8, g_9\}$ ;

②在9张灰度序列图中,计算每张图像的灰度值总和,并记录灰度值总和最低的图像序号,如第四张图 $g_4$ ;

③将原矿物颗粒RGB序列图中第四张(序号 $r_4$ )之前的图添加到序列图末端,结果为 $rgb = \{r_4, r_5, r_6, r_7, r_8, r_9, r_1, r_2, r_3\}$ ,此时偏光序列图消光位校正工作已完成。

## 2 结合改进的SKnet与Bi-GRU的矿物颗粒识别算法

针对采集处理后的岩石矿物颗粒偏光序列图像,岩石颗粒成分的识别不仅需要考虑单个图像中的纹理、亮度等特征,还需要结合不同正交偏光角度下的图像之间的关联信息。本文将二维序列图像类比为视频分类的方式进行岩石矿物颗粒成分识别,提出一种基于卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN)的岩石矿物颗粒识别网络, CNN采用改进的SKnet进行特征提取,同时利用双向门控循环单元(Bidirectional Gated Recurrent Unit, Bi-GRU)提取岩石矿物颗粒偏光序列图像的前后关联特征,完成识别任务,其框架如图4所示。

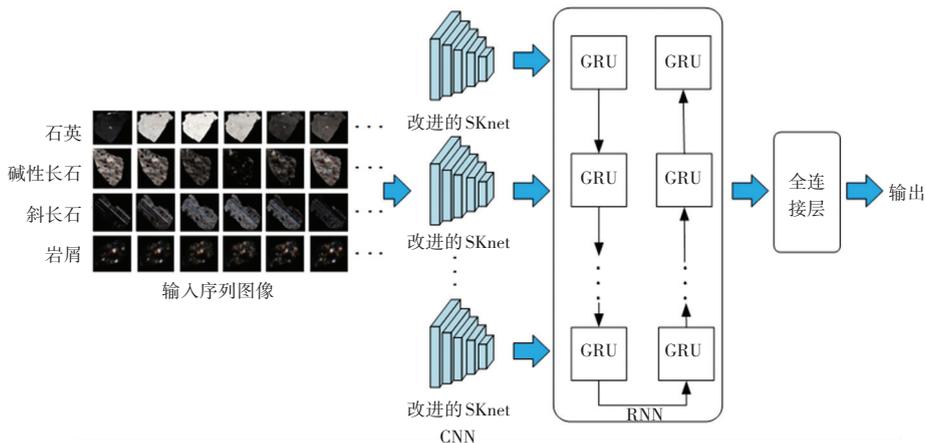


图4 结合改进的SKnet与Bi-GRU的矿物颗粒识别框架

Fig. 4 Combining improved SKnet with Bi-GRU for mineral particle identification framework

整体输入为岩石矿物颗粒正交偏光序列图像, 通过改进的 SKnet 特征提取模块对每一张偏光序列图像提取特征; 将提取出来的特征送入后续双向 GRU 循环神经网络模块进行序列维度的信息建模; 最后, 利用全连接层得到最终的分类识别结果。本文采用交叉熵作为损失函数, 计算公式为

$$loss = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n p(x_{ij}) \log(q(x_{ij})) \quad (1)$$

其中,  $p(x_{ij})$  表示第  $i$  个样本的真实标签为  $j$ ,  $q(x_{ij})$  表示第  $i$  个样本预测为第  $j$  个标签值的概率, 共有  $n$  个标签值,  $m$  个样本。

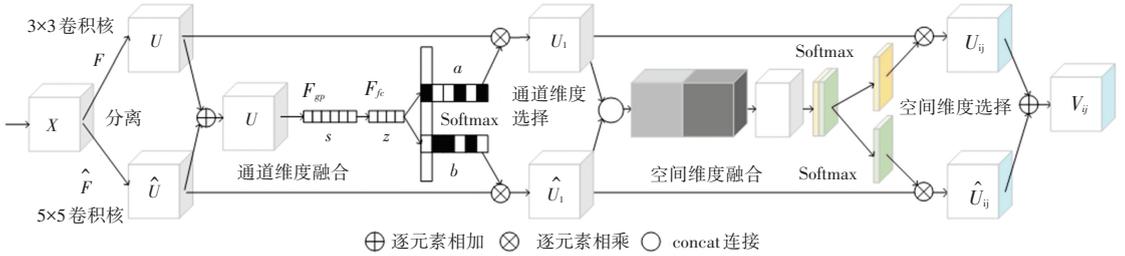


图 5 改进的 SK 卷积内部结构图

Fig. 5 Improved SK convolution internal structure diagram

首先是分离操作, 对于输入的特征图, 默认情况下进行两次转换  $\tilde{F} = X \rightarrow \tilde{U} \in R^{H \times W \times C}$  和  $\hat{F} = X \rightarrow \hat{U} \in R^{H \times W \times C}$ , 两个变换对应的卷积核的尺寸分别为  $3 \times 3$  和  $5 \times 5$ , 并且都是由卷积、Relu 激活函数、BatchNorm 批处理等操作组成。通道特征融合操作, 融合的基本思路是设计一个门控装置控制流入下一个卷积层中不同分支的信息流。 $s_c$  表示  $s$  的第  $c$  个元素内容, 通过在  $U$  的  $H \times W$  维度上进行压缩计算得到, 其计算过程为

$$s_c = F_{gp}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (2)$$

其中,  $U_c$  为  $\tilde{U}_c$  和  $\hat{U}_c$  相加融合得到。

通过全连接层产生一个紧凑的特征图像用于精确及调整的选择, 这部分也进行了降维处理, 式(3):

$$z = F_{fc}(s) = \delta(B(Ws)) \quad (3)$$

其中,  $\delta$  表示 ReLU 激活函数;  $B$  表示批规范化;  $W \in R^{d \times C}$ 。

为了验证  $W$  中  $d$  的作用, 采用了一个衰减比  $r$  控制  $d$  的值, 来探究  $d$  对于模型效率的影响为

$$d = \max\left(\frac{C}{r}, L\right) \quad (4)$$

其中,  $C$  表示通道数,  $L$  为  $d$  的最小值。

## 2.1 加入空间特征融合的 SKnet 特征提取模块

SKnet 网络由多个 SK (Selective Kernel) 单元堆叠而成, 通过 SK 卷积以非线性的方法聚合来自不同卷积核大小的特征, 进而实现动态调整感受野的尺寸, 便于对岩石颗粒多尺度信息的提取<sup>[10]</sup>。基础的 SK 卷积由分离、融合和选择等操作组成。本文在基础的 SK 卷积中添加了空间维度上的特征融合和选择, 利用软注意力机制使模型关注到更多重要的区域, 从而提升岩石薄片图像矿物识别分类的性能。本文改进的 SK 卷积的内部结构图如图 5 所示。

在压缩特征描述符  $z$  的引导下, 采用跨通道的软注意力机制自适应选择不同空间尺度的信息,  $a_c$  和  $b_c$  分别为  $\tilde{U}$  和  $\hat{U}$  软注意力矢量, 通过对各个核的注意力值分别得到, 式(5):

$$a_c = \frac{e^{A_c z}}{e^{C_c z} + e^{B_c z}}, b_c = \frac{e^{B_c z}}{e^{A_c z} + e^{B_c z}} \quad (5)$$

其中,  $A, B \in R^{C \times d}$ ,  $A_c \in R^{1 \times d}$  表示  $A$  的第  $c$  行,  $B_c \in R^{1 \times d}$  表示  $B$  的第  $c$  行。

空间特征融合的输入图  $\tilde{U}_1$  和  $\hat{U}_1$ , 为

$$\tilde{U}_1 = a_c \cdot \tilde{U}, \hat{U}_1 = b_c \cdot \hat{U} \quad (6)$$

使用空间注意力融合操作来关注图像哪些区域特征更有意义。不同于简单的像素级相加来整合多分支特征, 引入自适应分支特征融合, 实现不同特征的高效融合, 具体为不同分支的特征乘上相应的可学习权重并相加得到新的融合特征。自适应权重的学习过程为通过连接不同分支的特征图  $\tilde{U}_1$  和  $\hat{U}_1$ , 进行  $1 \times 1$  大小的卷积操作, 改变特征图的通道数, 得到与分支特征图相同尺寸、通道数为 2 的特征图, 2 个通道上的特征图分别表示为  $A_{ij}$  和  $B_{ij}$ , 然后通过 softmax 函数对其进行处理, 分别得到不同分支特征的融合比例矩阵为

$$\alpha_{ij} = \frac{e^{A_{ij}}}{e^{A_{ij}} + e^{B_{ij}}}, \beta_{ij} = \frac{e^{B_{ij}}}{e^{A_{ij}} + e^{B_{ij}}} \quad (7)$$

针对于两个分支特征图中同一位置的像素点,经过特征融合策略得到最终输出特征图像  $V$ , 为

$$V = \alpha_{ij} \cdot \tilde{U}_1 + \beta_{ij} \cdot \hat{U}_1, \alpha_{ij} + \beta_{ij} = 1 \quad (8)$$

本文使用的特征提取模块总体结构如图6所示,包含多个层级,每个层级内部由多个SK单元连接而成,每个层级内包含的SK单元的数量分别为3、4、6、3。通常情况下,在每一个层级中多个SK单元的连接方式是简单的相加组合,不同的SK单元

之间并没有共享信息,简单堆叠的方式无法有效地提取想要的特征。因此,本文借鉴 Xu Ma<sup>[11]</sup> 等人在2021年提出的深度连接注意力网络模式,在堆叠的相邻SK单元之间建立连接关系,使多个单元间能够进行信息交流,提高了网络的学习能力。具体的多个SK单元的连接结构如图7所示,由于相互连接的两个特征在空间、通道等方面维度有所不同,所以分别采用最大池化层和全连接层来实现特征间空间维度和通道维度的匹配。



图6 特征提取模块整体架构

Fig. 6 The overall architecture of the feature extraction module

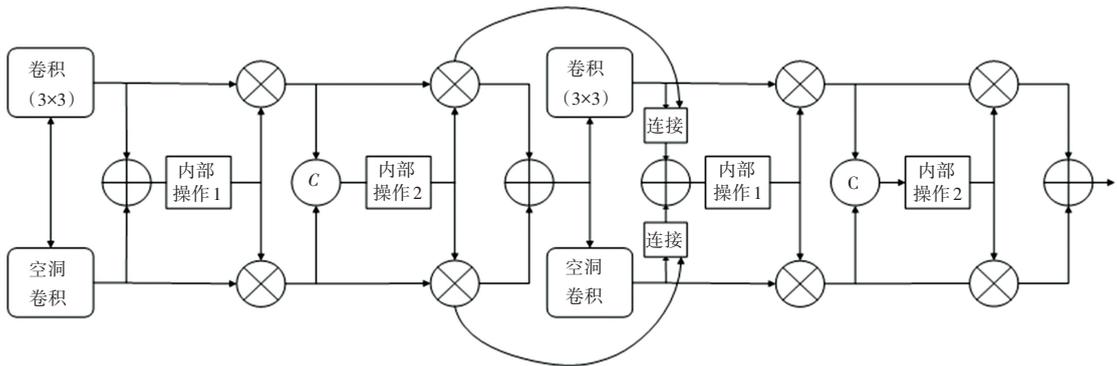


图7 改进的SK单元互连结构图

Fig. 7 Improved SK cell interconnect structure diagram

## 2.2 用于时间序列预测的双向GRU门控循环单元

岩石矿物颗粒成分的识别需要结合正交偏光序列图像中每一张图像的特征以及序列图像的变化信息才能准确的识别矿物颗粒种类。本文采用双向门控循环单元(Gated Recurrent Unit, GRU),利用SKnet提取的序列图特征信息送入GRU循环卷积神经网络,充分的利用岩石矿物颗粒序列图像的相关信息。GRU是基于门机制的,原理类似于长短期记忆(LSTM),依靠历史信息在当前点做出相应的预测。LSTM的门机制主要通过输入门、遗忘门、输出门来控制记忆单元的信息,解决信息的长依赖问题。GRU模型改进了LSTM的门机制结构,比LSTM更加简洁,训练效率相对较高。GRU也是典型的RNN架构,通过结合序列样本中当前样本的信息和上一个样本的隐状态,计算当前样本的隐状态并向下游继续传递。GRU的内部结构如图8所示。

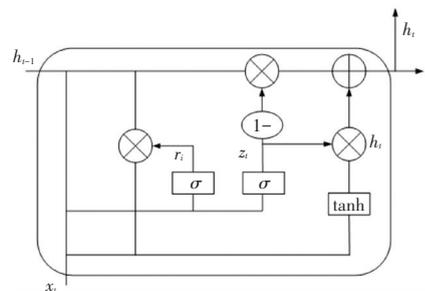


图8 GRU内部结构图

Fig. 8 GRU internal structure diagram

GRU有两个门,一个是重置门,另一个是更新门。更新门控的计算公式为

$$z_t = \text{sigmoid}(\omega_{(z)}x_t + u_z h_{t-1}) \quad (9)$$

其中,  $\omega_{(z)}$  和  $u_z$  为权重矩阵;  $x_t$  是经过线性变换后时间步长的向量输入;  $h_{t-1}$  保存了上一个经过线性变换后时间步长的信息。

将变换后的信息汇总并输入激活函数, 最终的结果在 0 到 1 之间变换。更新门决定了历史信息在传递时的比例, 复制过去的信息, 以减少梯度消失的风险。重置门控的计算公式为

$$r_t = \text{sigmoid}(\omega_{(r)}x_t + u_r h_{t-1}) \quad (10)$$

重置门的计算方法与更新门相同, 只是改变了用于线性变换的参数  $\omega_{(r)}$  和  $u_r$ 。复位门计算非聚焦信息的比例, 这些信息稍后会被过滤掉。GRU 结构中还包含一个候选的隐藏层  $h_t$ , 计算公式为

$$h_t' = \tanh(\omega x_t + u(r_t \odot h_{t-1})) \quad (11)$$

其中,  $\omega$  和  $u$  为权重矩阵。

候选隐藏层可以理解为当前的记忆内容。控制过滤后的信息在隐藏层的比例, 将这个信息与隐藏层当前时刻的信息相加, 得到隐藏层的最终输出信息  $h_t$ , 也是当前时刻的最终记忆状态, 计算公式为

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h_t' \quad (12)$$

可以发现, 该模块通过应用注意力机制使得  $z_t$  与  $1 - z_t$  是互相制约与影响的, 对于前一个序列样本的输出遗忘的越多, 当前得到的候选状态过滤掉的信息就越少, 反之同理。直接丢弃与模型分类无关的信息, 可以在一定程度上解决模型接收到的信息过多的问题, 提高模型分类的效率。

单向 GRU 的状态是从前到后单向传输的, 也就是序列图前面的样本无法考虑后面样本序列对前面样本的影响, 很容易忽略后面的序列图的影响, 双向 GRU 是单向 GRU 的一种变体, 其输出依赖于向前和向后状态的双重作用, 克服了单向 GRU 的局限, 从而提升基于正交偏光序列图像的岩石薄片矿物识别网络的性能。

## 3 实验结果与分析

### 3.1 数据集

本文利用课题组铸体粒度图像处理软件中的交互式分割方法, 从大量岩石薄片正交偏光序列图像中提取出不同类别矿物颗粒偏光序列图像, 以一组偏光序列图像对应一类矿物颗粒标签的方式, 构建矿物颗粒分类识别的原始数据集。该数据集包含 4 类矿物颗粒偏光序列图, 分别为石英、碱性长石、斜长石、岩屑。在实际数据采集过程中, 通常情况下, 岩石薄片图像中斜长石颗粒含量相对较少, 其余 3 种矿物颗粒数量相对较多。为了防止样本数量的不均衡影响最终识别效果, 本文根据 4 种岩石矿物颗粒样本数量占比, 使用图像几何变换、仿射变换等方

法对数据集进行了进一步的扩充, 建立最终数据集。通过数据增广, 石英、碱性长石、斜长石、岩屑 4 类矿物颗粒的偏光序列图像的数量均达到每类 2 000 组, 按照 8 : 1 : 1 的比例划分为训练集、验证集和测试集, 最终构成的数据集分布情况见表 1。

表 1 矿物颗粒识别数据集

Tab. 1 Mineral particle identification dataset			
种类	训练集	验证集	测试集
石英	1 600	200	200
碱性长石	1 600	200	200
斜长石	1 600	200	200
岩屑	1 600	200	200
总计	6 400	800	800

### 3.2 实验环境

实验中模型训练使用 NVIDIA RTX2080Ti GPU, 显存为 32 G, CPU 为 Intel (R) Core (TM) i7-9700。实验环境基于 Ubuntu20.04 系统, 使用 CUDA 9.0 加速运算平台, 深度学习的框架使用 Pytorch-1.2.0。

在送入网络训练之前对岩石矿物颗粒正交偏光序列图像进行消光位校正处理, 然后将图像数据调整为  $224 \times 224$ , 再进行归一化操作, 批处理大小设置为 8, 初始学习率设置为 0.001, 训练迭代次数为 400, 使用随机梯度下降优化器, 并且采用 poly 策略调整学习率。

### 3.3 评价指标

本文使用混淆矩阵展示所提出的基于改进的 SKnet 和 Bi-GRU 的岩石薄片图像矿物识别方法对 4 类岩石矿物颗粒的预测结果, 同时反映影响模型预测准确性的样本分布情况。此外, 为了评估本文提出的方法面向不同岩石矿物偏光序列图像在识别性能上的差异, 采用识别率作为分类网络的评价指标。识别率反映了每一类矿物颗粒正确分类的数量占该类矿物颗粒实际总数量的比例, 识别率越高, 表明正确分类的数量越多, 公式为

$$acc = \frac{n_i}{N_i} \quad (13)$$

其中,  $n_i$  为第  $i$  类矿物中正确分类的数量,  $N_i$  为第  $i$  类矿物颗粒实际总数量。

### 3.4 结果分析

根据本文提出的方法, 利用矿物颗粒序列图像预测和真实类别的识别结果, 建立混淆矩阵见表 2。表 2 中每一行代表该类别的判别结果的数量, 识别正确的数量除以该类别实际总数即为该类别的召回

率。每一列为其他类别识别为当前类别的数量,识别正确的数量除以列内数量总和即为该类别的精确率。可以看出本文提出的分类模型准确率和召回率都达到90%以上,说明利用本文提出的分类方法,石英、碱性长石、斜长石、岩屑这4类矿物可以被有效的区分开。

为表明本文所采用的岩石矿物颗粒识别方法的有效性,将本文提出的方法与现有的一些矿物颗粒识别算法进行了对比,对比结果见表3。

表3 本文提出的方法与其他方法的实验结果对比

Tab. 3 The method proposed in this paper is compared with the experimental results of other methods

分类方法	石英识别率	碱性长石识别率	斜长石识别率	岩屑识别率	合计识别率
k近邻机器学习模型以及权值投票机制分类方法 <sup>[4]</sup>	92.4	83.2	77.8	93.0	87.8
基于CaffeNet以及颜色和纹理的序列图像筛选方法 <sup>[6]</sup>	93.3	87.7	78.7	93.9	89.6
基于SKnet及亮度排序图像筛选方法 <sup>[9]</sup>	94.0	86.0	77.4	94.4	89.3
基于SKnet和单向GRU的方法 <sup>[12]</sup>	95.5	95.1	92.9	95.7	94.8
基于ResNet+LSTM的分类方法 <sup>[13]</sup>	94.5	93.5	90.7	94.2	93.2
本文方法	<b>98.0</b>	<b>98.0</b>	<b>95.5</b>	<b>97.5</b>	<b>97.2</b>

由表3可以发现,利用传统机器学习方法实现矿物颗粒分类,总体识别率较低;使用CaffeNet或SKnet再结合一定的图像筛选方法进行识别分类,针对石英、碱性长石、岩屑颗粒识别率相对较好,但是斜长石颗粒识别率较低,由于矿物颗粒自身特征的复杂性,单张图像很难将矿物颗粒特征完全表征出来,在某些情况下斜长石颗粒容易被误识别为碱性长石或者岩屑颗粒;基于SKnet结合单向GRU的分类方法总体来说识别率较前几种方法已有明显提升;本文提出的分类方法,更充分的利用了岩石矿物颗粒正交偏光图像的消光特征,改进的SKnet以及

表4 是否进行矿物颗粒偏光序列图像消光位校正实验结果对比

Tab. 4 Whether to carry out the comparison of the experimental results of the extinction position correction of the polarized sequence images of mineral particles

分类方法	石英识别率	碱性长石识别率	斜长石识别率	岩屑识别率	合计识别率
不使用消光位校正	93.3	96.5	95.8	96.3	95.47
基础SKnet	95.3	96.2	96.2	97.5	96.30
单向GRU循环神经网络	98.0	96.5	95.3	96.3	96.53
本文方法	<b>98.3</b>	<b>97.2</b>	<b>96.5</b>	<b>98.3</b>	<b>97.25</b>

由表4可以发现,对矿物颗粒偏光序列图像进行消光位校正,4类岩石矿物颗粒的最终识别率都得到了提升,经过消光位校正后,输入到双向GRU门控单元的序列图像特征保持了很好的一致性,更有利于网络学习序列前后图像的自身特征信息以及关联信息。

表2 基于本文分类方法的混淆矩阵

Tab. 2 Confusion matrix based on the classification method in this paper

真实	预测				召回率/%
	石英	碱性长石	斜长石	岩屑	
石英	196	1	3	0	<b>98.0</b>
碱性长石	2	196	2	0	<b>98.0</b>
斜长石	1	4	191	4	<b>95.5</b>
岩屑	0	1	4	195	<b>97.5</b>
准确率/%	<b>98.4</b>	<b>97.0</b>	<b>96.4</b>	<b>97.5</b>	<b>97.2</b>

双向GRU门控单元更好的利用了序列图像的前后关联信息,对于4种岩石矿物颗粒的识别率都达到了95%以上,相比于已有的算法有明显的提升。

为了测试本文分类方法中岩石矿物颗粒偏光图像消光位校正处理以及网络模型各模块对最终的识别效果产生的影响,基于本文提出的网络结构,分别进行了使用消光位校正和不使用消光位校正操作的对比实验测试,使用改进的SKnet网络模型和基础SKnet网络模型的对比实验测试以及使用单向GRU门控单元和双向GRU门控单元对比实验测试,测试结果见表4。

使用基础的SKnet替换本文提出的改进的SKnet后,对4种岩石矿物颗粒的识别率都产生了一定的影响,说明本文利用特征融合优化后的SKnet特征提取模块,更有效地提取出想要的特征信息,进一步降低了误识别的概率。双向GRU门控单元可以更好的学习岩石矿物颗粒偏光序列图的序

列信息,碱性长石、斜长石、岩屑矿物颗粒的识别率都得到了提升。

#### 4 结束语

为了更有效地完成岩石薄片偏光序列图的颗粒成分的分析工作,本文提出了一种结合改进的 SKnet 和双向 GRU 的岩石薄片正交偏光序列图像矿物颗粒分类方法。首先,利用已有分割方法,从岩石薄片正交偏光序列图像中提取出石英、碱性长石、斜长石、岩屑 4 类岩石矿物颗粒图像,对提取出的矿物颗粒序列图做消光位校正处理,将消光位校正后的序列图作为输入数据进行训练;其次,利用改进的 SKnet 网络结合具有记忆的双向 GRU 循环神经网络,加强对矿物颗粒序列图中序列特征的提取。本文的方法在已有的数据集上表现良好,最终 4 类岩石矿物颗粒的识别率都达到 95% 以上,总体的识别率为 97.25%, 优于已有的岩石矿物颗粒识别分类方法。但本文只针对岩石薄片相对较多的 4 种岩石矿物颗粒进行了分类识别,在实际石油地质开采工作中,岩石矿物颗粒的种类十分丰富,在后续的工作中可以继续探索更多岩石矿物颗粒类别下的识别分类方法。

#### 参考文献

[1] MIYNARCZUK M, GÓRSZCZYK A, ŚLIPEK B. The application of pattern recognition in the automatic classification of microscopic rock images[J]. *Computers & Geosciences*, 2013, 60: 126–133.

[2] 杨宗瑞. 基于偏光序列图像的岩石颗粒分割及矿物分类[D]. 成都: 四川大学, 2014.

[3] AKKAŞ E, AKIN L, ÇUBUKÇU H E, et al. Application of Decision Tree Algorithm for classification and identification of natural minerals using SEM-EDS[J]. *Computers & Geosciences*, 2015, 80: 38–48.

[4] 闵涛. 基于薄片图像的岩矿成分识别研究[D]. 成都: 四川大学, 2016.

[5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Advances in neural information processing systems*, 2012, 25: 1097–1105.

[6] 彭志伟. 岩矿薄片偏光序列图矿物颗粒分割与识别[D]. 成都: 四川大学, 2018.

[7] NANJO T, TANAKA S. Carbonate lithology identification with machine learning [C]//Abu Dhabi International Petroleum Exhibition & Conference. OnePetro, 2019.

[8] XU Y, DAI Z, LUO Y. Research on Application of Image Enhancement Technology in Automatic Recognition of Rock Thin Section [C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020, 605(1): 012024.

[9] 朱磊. 基于深度学习的岩石薄片偏光图像颗粒提取与识别[D]. 成都: 四川大学, 2021.

[10] 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.

[11] MA X, GUO J, TANG S, et al. Learning Connected Attention for Convolutional Neural Networks [C]//2021 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2021: 1–6.

[12] 钟逸. 结合岩矿消光特性的薄片图像颗粒提取与识别研究[D]. 成都: 四川大学, 2020.

[13] LIU H, REN Y L, LI X, et al. Rock thin-section analysis and identification based on artificial intelligent technique [J]. *Petroleum Science*, 2022, 19(4): 1605–1621.

(上接第 103 页)

#### 参考文献

[1] 李新叶, 龙慎鹏. 基于深度神经网络的少样本学习综述[J]. *计算机应用研究*, 2020, 37(8): 2241–2247.

[2] YAO L, MAO C, LUO Y. Graph Convolutional Networks for Text Classification [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019: 7370–7377.

[3] HUANG L, MA D, LI S, et al. Text Level Graph Neural Network for Text Classification [J]. *arXiv preprint arXiv: 1910. 02356*,

2019.

[4] ZHANG Y, YU X, CUI Z, et al. Every document owns its structure: Inductive text classification via graph neural networks [J]. *arXiv preprint arXiv:2004.13826*, 2020.

[5] VELIKOVI P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks [J]// *arXiv: 1710.10903*. 2017.

[6] WEI J, ZOU K. EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks [J]. *arXiv preprint arXiv:1901.11196*, 2019.