

基于SSD与图像变换的镜下矿物光片智能识别

侯振隆, 申晋容, 魏继康, 赵文天

(东北大学 资源与土木工程学院, 辽宁 沈阳 110819)

摘要: 在矿物识别中,当识别伴生矿物时,有时会产生漏判、误判.为了解决上述问题,开展了显微镜下矿物的智能化识别方法研究.首先,改进了SSD网络并结合图像变换构建了一种智能识别方法;其次,将该方法应用于中国辽宁省某铁矿光片的显微镜下矿物图像,通过试验证明了方法的准确性;最后,确定了学习率、批量尺寸对损失函数的影响,使用梯度下降法进一步提高了识别精度.在试验中,识别精度超过90%,最高可达100%,损失函数值最小值约为0.008.结果表明,提出的方法具有较强的矿物识别能力.

关键词: 矿物识别;深度学习;SSD;图像变换;矿物含量估算

中图分类号: P 575

文献标志码: A

文章编号: 1005-3026(2025)06-0131-08

Intelligent Identification of Minerals Polished Section Under Microscope Based on SSD and Image Transformation

HOU Zhen-long, SHEN Jin-rong, WEI Ji-kang, ZHAO Wen-tian

(School of Resources & Civil Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: HOU Zhen-long, E-mail: houzhenlong@mail.neu.edu.cn)

Abstract: In mineral identification, there may sometimes be cases of missed identification or incorrect identification when identifying associated minerals. To address these issues, the intelligent mineral identification methods under the microscope were developed. Firstly, an intelligent identification method was constructed by improving the SSD network and incorporating image transformation. Secondly, the method was applied to microscope images of minerals in polished sections from the iron ore in Liaoning Province, China, and its accuracy was demonstrated through tests. Finally, the effects of learning rate and batch size on the loss function were determined, and the accuracy was further improved by using the gradient descent method. In the tests, the identification accuracy exceeds 90%, reaching up to 100%, with the minimum value of the loss function was approximately 0.008. The results indicate that the proposed method has strong mineral identification capabilities.

Key words: mineral identification; deep learning; single shot multibox detector (SSD); image transformation(IT); mineral content estimation

识别矿物是地质研究中不可或缺的部分.目前,由于人工智能高效、免受人为干扰等优点,该技术已经应用于矿物识别的研究中.智能识别是当前地球科学领域中的热点之一,许多研究都是利用图像数据开展的.例如,基于显微镜下图像^[1-3]和岩石图像^[4-6]进行的岩性识别与分类.在这些研究中,机器学习、图像处理和人工神经网络技术发挥了重要作用.随着深度学习的快速发展,不同类型的卷积

神经网络越来越多地被应用于地学中.在找矿预测^[7]、地质填图^[8]、遥感影像的识别与提取^[9-10]以及掌子面岩石结构分类^[11]等领域中,基于卷积神经网络算法的准确率分别可达88.9%,90.0%,95.3%和98.1%,体现了其具有较高的性能.近年来,这类智能算法也被应用于矿石图像分类^[12-14],取得了良好的识别结果,对于矿山高效生产具有重要意义.除了岩性识别,还需要辨别出显微镜下图像中的不同

收稿日期: 2023-12-27

基金项目: 国家自然科学基金资助项目(42204140); 辽宁省自然科学基金资助项目(2022-MS-107).

作者简介: 侯振隆(1988—),男,辽宁沈阳人,东北大学副教授.

矿物.然而,受限于矿物图像复杂程度,使用现有的方法易造成错误判断和种类的遗漏.

在图像识别领域,目标检测是重要的方法之一.相比于其他方法,它具有准确率高、速度快等优点.常用的网络包括:Faster R-CNN(faster region-based CNN)^[15],YOLO(you only look once)^[16]和SSD(single shot multibox detector)^[17]等.SSD是在Faster R-CNN和YOLO的基础上发展得来的,在它们中具有最强的检测能力^[17],被认为是当前较为高效的目标检测方法,将SSD用于矿物识别有望获取更准确的矿物种类信息.

因此,本文开展了一项基于SSD与图像变换的智能识别方法研究,以解决伴生矿物低精度识别等问题.将提出的方法应用于中国辽宁省某铁矿的显微镜下图像,分析和讨论了方法性能以及影响结果的因素.

1 方 法

1.1 改进的SSD网络

SSD是一种单阶段的目标检测方法^[18],具有

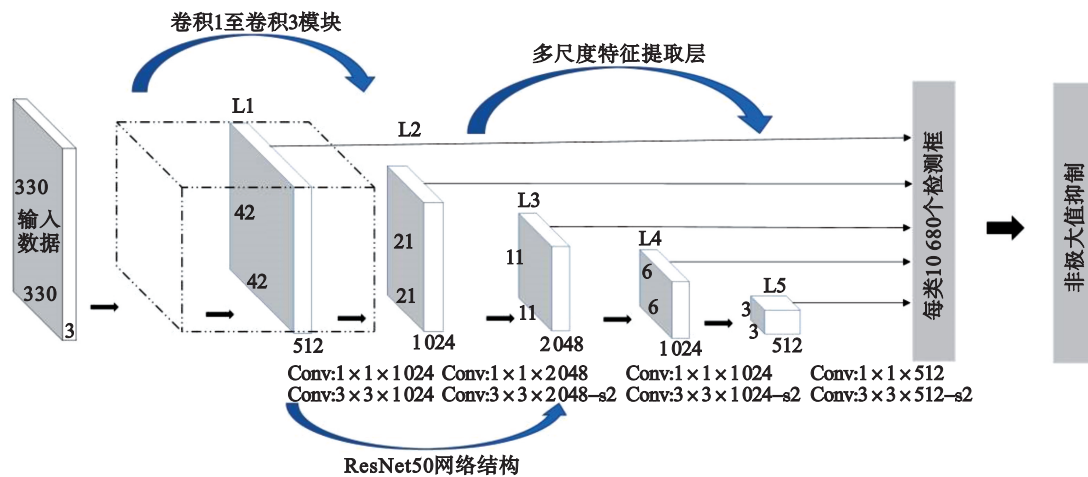


图1 改进的SSD网络结构

Fig. 1 Network structure of improved SSD

1.2 图像变换和评价指标

当存在矿物伴生情况时,显微镜下图像特征会比单一矿物更复杂,仅使用改进的SSD也不能非常准确地分辨矿物,故提出一种结合SSD和图像变换(image transformation, IT)的识别方法(简称SSD-IT)(图2)来解决这一问题.该方法以矿物在反射偏光显微镜下反射色、亮度的差别为基础,先通过SSD对图像中的矿物种类进行识别,再通过图像变换对矿物含量进行估算.具体地,当应用SSD-IT时,伴生体将被视为一种新的矿物种类,也要被加入训练集.将建立

高效率的优点.通过金字塔特征结构^[19],SSD能充分利用所有卷积层的特征信息.不同于YOLO网络只使用一个特征层进行预测,SSD方法利用多个特征层进行分类和回归,并执行非极大值抑制筛选边界框.因此,SSD具有比YOLO更强的小目标检测能力.

为了进一步提升SSD的识别能力与训练速度,将其基础网络由VGG16(visual geometry group network with 16 weight layers)替换为ResNet50(residual net with 50 layers).这是因为ResNet50在精度、计算速度上均优于VGG16;通过增加网络层数,可进一步提高ResNet50用于识别的效果^[20].因此,从SSD的第1层至第5层(L1~L5)提取特征图,其大小依次为42×42,21×21,11×11,6×6,3×3.其中,前3层沿用ResNet50的网络结构;第4层与第5层各包含两个部分,均由卷积层与BN(batch normalization)层组成:前半部分的卷积层步长为1,卷积核大小为1×1,后半部分的卷积层步长为2,卷积核大小为3×3.改进的SSD网络结构如图1所示.本文通过边界框识别出不同矿物的种类,从而确定整个视域中总的种类.

的数据集输入SSD进行训练并获取最优的模型后,利用经测试后的SSD进行初步识别,给出含伴生体的矿物种类.再按照式(1)^[21]将彩色图像变换为灰度图像.

$$L = R \times 0.299 + G \times 0.587 + B \times 0.114. \quad (1)$$

其中: L 表示灰度值; R, G, B 分别表示彩色图像的红、绿、蓝颜色通道值.

矿物在反射偏光显微镜下亮度是不同的,例如,赤铁矿的亮度高于磁铁矿.依据矿物的灰度值范围可将伴生体中矿物进行有效区分.统计同一灰度值的像素在视域中的面积占比,该比值即矿物含

量.通过变换,可以有效避免识别不准、不全等问题,并且该方法能够同时进行矿物识别与含量估算.

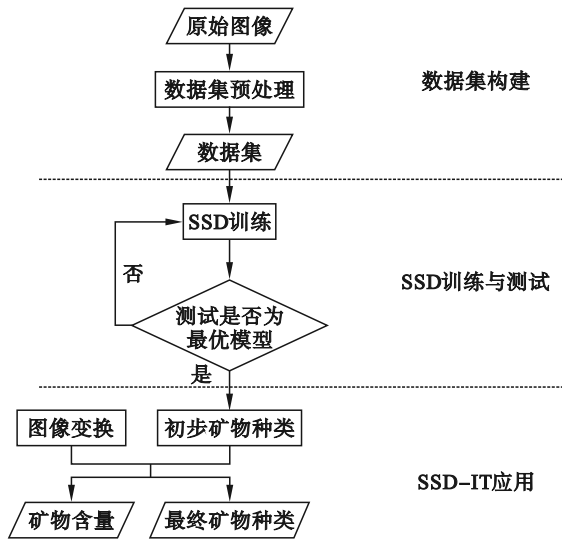


图 2 SSD-IT 流程示意图
Fig. 2 Flow chart of SSD-IT

在镜下图像中,目标矿物的含量可能很少,而非目标矿物很多,所以导致真负例(true negative, TN)值会过高,即在非目标范围内没有预测出矿物的情况增加.因此,使用 TN 计算的评价指标不宜被用来评价准确率.识别精度(P)、召回率(R)和 F -score(F)将作为本研究的评价指标,计算公式为

$$P = \frac{TP}{TP + FP}, \quad (2)$$

$$R = \frac{TP}{TP + FN}, \quad (3)$$

$$F = \frac{(\beta^2 + 1.0) \times P \times R}{\beta^2 \times P + R}. \quad (4)$$

其中:TP(true positive)是真正例,表示算法在目标范围内预测出了该矿物;FP(false positive)是假正例,表

示算法预测出一个目标范围中不存在的矿物;FN(false negative)是假负例,表示算法没有在目标范围内预测出该矿物; β 表示召回率相对于识别精度的重要性,当 $\beta=1$ 时,识别精度与召回率的权重相同,当 $\beta=0.5$ 或 2 时,召回率的权重是识别精度的一半或 2 倍^[22]. $\beta=0.5, 1$ 和 2 分别对应 $F0.5$ -score, $F1$ -score 和 $F2$ -score.

对伴生体的识别必须是精确的,否则伴生体中的矿物不能被准确地区分.首先需要模型已预测出某一矿物,再判断其预测正确与否.因此,利用识别精度评价方法更为合适.通常地,随着识别精度增高,召回率就会下降;反之亦然^[23].综上所述,识别精度被选作主要的评价指标,同时也使用召回率和 F -score 用于评价,避免只使用单一指标无法全面客观评价模型的问题.当识别精度 $P > 90\%$ 时,可以认为获得了期望的结果.

1.3 数据预处理

对于模型的训练,图像的质量会影响识别精度,且图像数量必须保持较大的量级才有助于增强模型泛化能力.当构建训练集与测试集时,采集的图像数据必须进行预处理,从而提高图像质量并扩充图像数量.预处理包括图像的标注与数量扩增,主要步骤如下:①根据显微镜图像中矿物特征,为边界框选择比例和长宽比,将其以 PASCAL VOC 数据集格式标注,并保存为 XML 文件;②对 XML 文件和图像数据进行数据增强,包括缩放、翻转、旋转、镜像、随机裁剪等(图 3);在预处理过程中,要保持边界框和图像同步变化;③对步骤②中生成的图像数据进行筛选,删除预处理后图像内无矿物可识别的数据;④挑选预处理图像的 70% 作为训练集,30% 作为测试集.

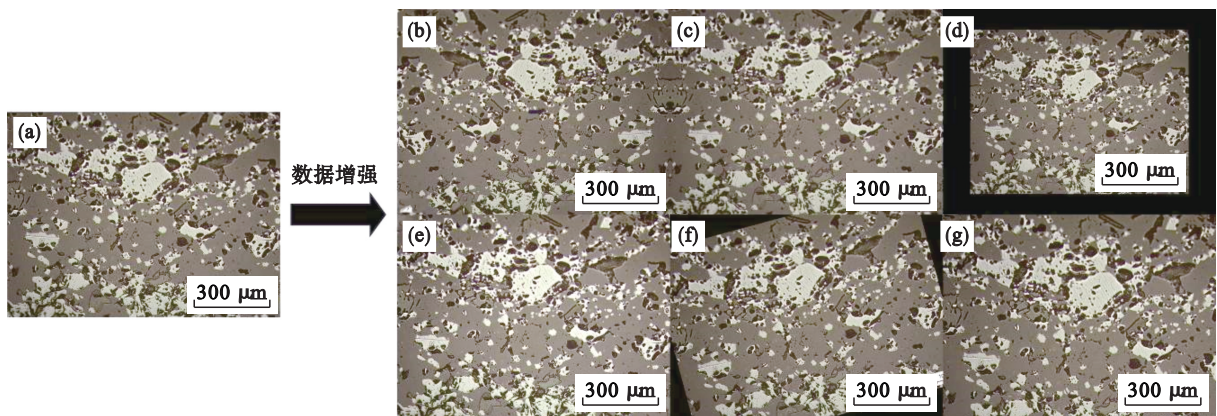


图 3 试验数据增强示意图

Fig. 3 Schematic diagram of test data enhancement

(a)—待增强图像数据; (b)—镜像处理后的图像数据; (c)—模糊处理后的图像数据; (d)—缩放处理后的图像数据; (e)—亮度变化后的图像数据; (f)—旋转处理后的图像数据; (g)—随机裁剪后的图像数据.

2 试验和讨论

为了检验 SSD-IT 方法的效果与可行性,将其应用于辽宁省某铁矿的显微镜下矿物图像.采集的岩石样本包括磁铁石英岩、赤铁石英岩、含硅酸盐赤铁石英岩和含碳酸盐赤铁石英岩.在用于获取图像的光片中,矿物主要包含硅酸盐矿物、褐铁矿、磁铁矿、赤铁矿.其中,磁铁矿和赤铁矿是目标矿物,赤铁矿在显微镜下具有上述矿物中最高的亮度,其 R, G, B 值均大于 210. 根据式(1),可得 $L > 210$; 磁铁矿较赤铁矿亮度低,其 R, G, B 值和灰度值分布于区间 $[120, 210]$ 内; 同理,可得其他矿物的 L 值. 使用了卡尔蔡司 Primotech 单偏光显微镜(20×10 倍镜),通过 12 个光片采集了 160 张图像数据.在数据预处理后,扩增了图像数量:数据集包含 4 184 张图像,其中,训练集图

像数量为 3 075 张,测试集图像为 1 109 张.对于每一个模型,需进行 200 次迭代,训练约 6.6 h. 试验是在天河 HPC4 超级计算机的 GPU(graphics processing unit)节点(配置:处理器 Intel Xeon Gold 6354 @ 3.00 GHz, 18 核×2; 256 GB 内存; GPU 为 NVIDIA HGX A100×2; 系统为 RedHat Enterprise Linux 8.4)上进行的.代码是利用 Python 语言编写的,调用了 LabelImg, imgaug, torch, PIL 等工具箱用于图像标注、图像增强、网络构建、图像处理等.图像数据的标注方式和识别结果如图 4 所示,其中,红色实线框为预测的目标矿物边界框, Mag 表示磁铁矿, Hem 表示赤铁矿, Mag+Hem 表示磁铁矿与赤铁矿伴生体.图像变换能够进一步区分伴生体中的矿物,使 SSD-IT 可以准确地识别出目标矿物,且能够计算出图像中矿物的含量.在图 4 中,磁铁矿和赤铁矿质量分数分别为 9.5%, 8.4%.

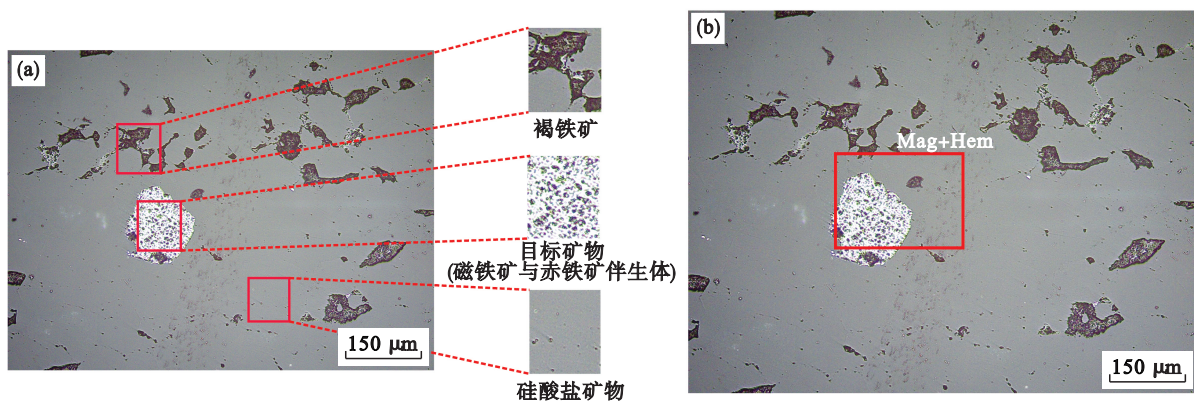


图 4 SSD-IT 图像数据标注方式与识别结果示意图

Fig. 4 Data labeling and identification results of SSD-IT image

(a)—标注的矿物图像; (b)—识别结果.

下面通过几组试验来讨论和分析 SSD-IT 的识别能力.

将 SSD-IT 与其他基于神经网络的识别方法作对比,以验证方法的有效性,评价结果见表 1. 在所有训练中,初始学习率均为 0.01,批量尺寸均为 32. 评价结果中,SSD-IT 的识别精度、召回率值最高, Faster-RCNN 和 YOLOv5s 的识别结果则较差; 通过图像变换,识别精度均获得了较大的提升,说明引入图像变换是有效的,且 SSD-IT 可提供精确的识别. 此外,SSD-IT 的 F-scores 值最高,也反映出 SSD-IT 的优越性. 由表 1 可知,当识别精度作为主要评价指标且令 $\beta=0.5$ 时,SSD-IT 方法可以获得最高的 F-score 值,因为此时识别精度具有高权重.

通过不同光片的图像进一步测试 SSD-IT 的性能. 本试验中主要利用识别精度进行评价. 从

测试集中对 12 个光片分别选取 5 张显微镜下图像. 初始学习率为 0.01,记录不同迭代次数下的平均识别精度(图 5). 由结果可知,平均识别精度总体上均随迭代次数增加而增加,但呈现不同程度的波动. 最终的平均识别精度最低为 83.3%,最高可达 100%,平均值为 92.3%. 结果表明,对于大多数样本,SSD-IT 都能取得超过 90% 的识别精度,甚至达到 100%,只有极少数样本的识别精度较低. 原因是这些图像中目标矿物含量少,此时的准确率会较高,如果以准确率为主要评价指标,那么评价将不合理. 这也证明了使用识别精度作为 SSD-IT 的评价指标更为客观、准确. 上述试验还表明,足够多的迭代次数是识别结果良好的基础,而较少的迭代次数可能无法获得精准的识别.

引入损失函数以确定损失值和迭代次数、学习率

之间的关系,从而确定最优的训练参数.在识别过程中,损失值 L_t 用来表示预测值与真实值之间的差距.损失值越高表示识别精度越低.损失值的计算公式^[17]:

$$L_t = \frac{1}{N} (L_{\text{conf}} + \alpha L_{\text{loc}}). \quad (5)$$

其中: N 表示正样本数量; L_{conf} 表示置信损失; L_{loc} 表示定位损失; α 表示加权因子,设为 1.此外,不同的学习率会导致损失值的变化.在不同学习率下,损失值随迭代次数变化的曲线如图 6 所示.

表 1 不同识别模型的评价结果

Table 1 Evaluation results of different identification models

模型	识别精度	召回率	F0.5-score	F1-score	F2-score
SSD-IT	91.19	80.91	88.93	85.74	82.78
SSD	24.46	40.00	26.33	30.36	35.40
Faster-RCNN-IT	40.99	52.62	42.88	46.08	49.79
Faster-RCNN	9.63	17.78	10.60	12.49	15.21
YOLOv5s-IT	40.30	1.45	6.33	2.80	1.80
YOLOv5s	11.80	4.35	8.79	6.36	4.98

注:后缀“-IT”表示不同神经网络分别与图像变换相结合.

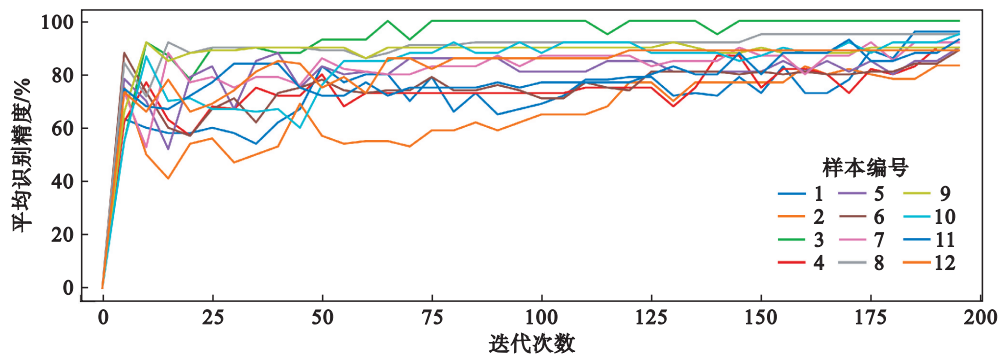


图 5 不同光片的镜下图像识别精度

Fig. 5 Identification precision for the microscopic images from polished sections

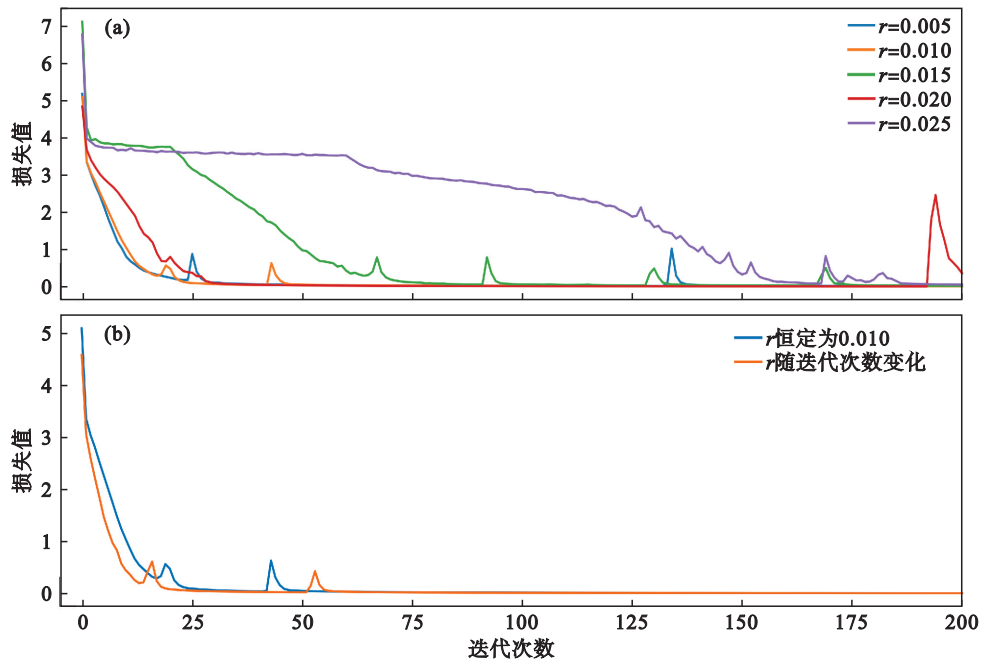


图 6 不同学习率情况下的损失值

Fig. 6 Loss values for different learning rates

(a)—不同学习率下损失值曲线图; (b)—学习率恒定为 0.010 以及学习率随迭代次数变化的损失值曲线图.

在图 6a 中,随着迭代次数的增加,除了学习率 $r=0.020$,所有的损失值均下降且最终趋于稳

定.当 r 分别为 0.005, 0.010, 0.015, 0.020, 0.025 时,相应损失值分别为 0.015, 0.008, 0.022, 0.362,

0.053. 因此, 当 $r=0.010$ 时损失值最小, 识别效果最好; 当 $r>0.010$ 时, 损失值总体上随着学习率的增加而增加. 尽管当 $r=0.005$ 时前 20 次迭代内收敛最快, 但其最终损失值不是最小值, 这是由于学习率过小, 导致后续训练中缓慢收敛. 当损失值趋近于一个平稳且较小的值时, 即可得到合适的模型. 试验证明学习率能够影响数据拟合和训练时间.

为了进一步验证学习率对结果的影响, 还对比了在学习率保持不变和变化情况下训练的损失值(图 6b). 在本试验中, 变化的学习率指在训练过程中采用梯度下降法, 即随着迭代次数的增加, 学习率的值将按式(6)更新:

$$r_k = r_0 \times d^t \quad (6)$$

其中: r_0, r_k 表示初始学习率和第 k 次迭代中的学习率; d 表示下降率; t 表示迭代次数除以整数 n 所得的整数值(向下取整). 试验中 $d=0.8, n=40$, 即每经过 40 次迭代, 学习率将下降 20%. 该方法使训练从较大的学习率开始, 但随着迭代次数的增加, 学习率将下降. 由于学习率过大造成的过拟合问题

将得以解决, 使拟合程度得到改善, 提升了训练效率, 从而减少了训练时长. 需要注意的是, 当迭代次数过大时, 学习率将过低, 可能对识别产生负面影响.

试验中, 初始的学习率设为 0.010, 则当使用固定的和变化的学习率时, 最终损失值分别为 0.008 4, 0.007 8. 这表明当使用相同初始学习率时, 梯度下降法可以获得更小的损失值. 由图 6b 还发现, 当学习率随着迭代次数增加而减小时, 损失值下降得更快.

在上述试验中, 发现批量尺寸对损失值也有影响. 记录不同批量尺寸情况下的损失值(图 7), 当批量尺寸 b 为 16, 32, 64, 128, 256 时, 所对应的最终的损失值分别为 0.036, 0.008, 0.015, 0.016, 0.041. 其中, 当 $b=32$ 时, 收敛速率最快; 当 $b=16$ 时, 损失值下降缓慢, 这是由于取值过小导致单次训练不充分, 故收敛速率慢, 且最终损失值较大. 随着批量尺寸的增大, 最终损失值也随之增加, 且收敛速率减慢. 批量尺寸对训练时间和模型精度都会产生影响, 故批量尺寸取值应适中.

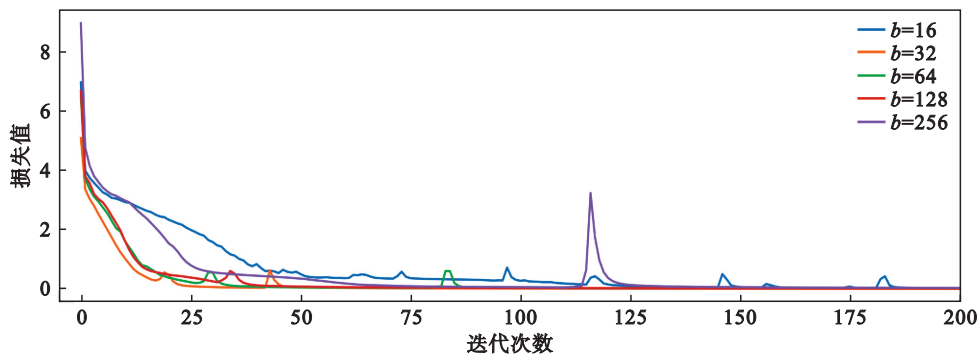


图 7 不同批量大小情况下的损失值

Fig. 7 Loss values for different batch sizes

对于部分变换后的图像出现目标矿物含量计算偏高的情况, 是由于显微镜下光线强度高, 变换为灰度图像后使图像中心的亮度值过高(图 8),

从而导致灰度值升高, 使目标矿物含量估算不准确. 由于该灰度值是 SSD-IT 识别的依据, 采集图像数据时应该尽量保持白天室内自然光照下的强度.

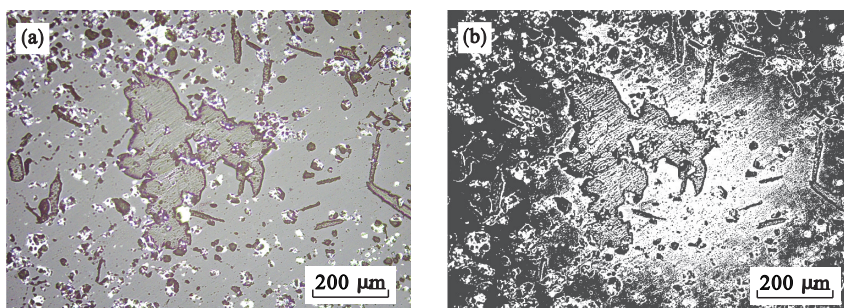


图 8 图像变换前后的对比

Fig. 8 Comparison of the images before and after the transformation

(a)—变换前; (b)—变换后.

3 结 语

本文提出了一种基于SSD与图像变换的矿物识别方法SSD-IT,并将其应用于辽宁省某铁矿的矿物图像数据.与基于YOLOv5s和Faster-RCNN的方法相比,SSD-IT大幅提高了伴生矿物的识别精度;同时,还有效估算出了图像中的矿物含量.研究发现,使用识别精度评估SSD-IT更为客观,特别是对于目标矿物含量少的图像,且识别精度作为主要指标能够获得更高的F-score.对于不同光片,SSD-IT的平均识别精度为92.3%,最大值甚至可达100%.学习率、批量尺寸对于训练过程中的收敛速度有一定影响.一般地,这些参数的值越小,收敛越快,最终的损失值越小;但是,当参数值过小时,收敛速度会变慢,损失值将变大.因此,为了获得最优模型,参数取值要适中.试验还证明了梯度下降法有助于提高识别精度,但迭代次数不宜过多.综上,SSD-IT能够提供期望的结果,并为智能化地质研究提供重要依据.

参考文献:

- [1] Młynarczyk M, Górszczyk A, Bartłomiej Ś. The application of pattern recognition in the automatic classification of microscopic rock images [J]. *Computers & Geosciences*, 2013, 60: 126-133.
- [2] Xu Z H, Ma W, Lin P, et al. Deep learning of rock microscopic images for intelligent lithology identification: neural network comparison and selection [J]. *Journal of Rock Mechanics and Geotechnical Engineering*, 2022, 14 (4): 1140-1152.
- [3] Polat Ö, Polat A, Ekici T. Automatic classification of volcanic rocks from thin section images using transfer learning networks [J]. *Neural Computing and Applications*, 2021, 33(18): 11531-11540.
- [4] Zhou Y M, Wong L N Y, Tse K K C. Novel rock image classification: the proposal and implementation of HKUDES_Net [J]. *Rock Mechanics & Rock Engineering*, 2023, 56(5): 3825-3841.
- [5] Singh V, Rao S M. Application of image processing and radial basis neural network techniques for ore sorting and ore classification [J]. *Minerals Engineering*, 2005, 18 (15): 1412-1420.
- [6] Tessier J, Duchesne C, Bartolacci G. A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts [J]. *Minerals Engineering*, 2007, 20(12): 1129-1144.
- [7] Ding K, Xue L F, Ran X J, et al. Siamese network based prospecting prediction method: a case study from the Au deposit in the Chongli mineral concentrate area in Zhangjiakou, Hebei Province, China [J]. *Ore Geology Reviews*, 2022, 148: 105024.
- [8] Wang X P, Zuo R G, Wang Z Y. Lithological mapping using a convolutional neural network based on stream sediment geochemical survey data [J]. *Natural Resources Research*, 2022, 31(5): 2397-2412.
- [9] 谢洪途,陈佳兴,张琳,等.基于脉冲神经网络的轻量化SAR图像舰船识别算法[J].东北大学学报(自然科学版), 2024,45(4):474-482.
(Xie Hong-tu, Chen Jia-xing, Zhang Lin, et al. Lightweight ship recognition algorithm based on SNN in SAR images [J]. *Journal of Northeastern University (Natural Science)*, 2024,45(4):474-482.)
- [10] 武锦龙,吴虹麒,李浩,等.基于改进DeepLabV³⁺的荞麦苗期无人机遥感图像分割识别方法研究[J].农业机械学报,2024,55(5):186-195.
(Wu Jin-long, Wu Hong-qi, Li Hao, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV³⁺ at seedling stage [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(5): 186-195.)
- [11] Chen J Y, Yang T J, Zhang D M, et al. Deep learning based classification of rock structure of tunnel face [J]. *Geoscience Frontiers*, 2021, 12(1): 395-404.
- [12] Deo A J, Sahoo A, Behera S K, et al. Vision-based size classification of iron ore pellets using ensemble convolutional neural network [J]. *Neural Computing and Applications*, 2022, 34(21): 18629-18641.
- [13] Zhou W Y, Wang H, Wan Z B. Ore image classification based on improved CNN [J]. *Computers and Electrical Engineering*, 2022, 99: 107819.
- [14] 王伟,李擎,张德政,等.基于深度学习的矿石图像处理研究综述[J].工程科学学报,2023,45(4):621-631.
(Wang Wei, Li Qing, Zhang De-zheng, et al. A survey of ore image processing based on deep learning [J]. *Chinese Journal of Engineering*, 2023, 45(4): 621-631.)
- [15] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [16] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016: 779-788.
- [17] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [C]// ECCV 2016: 14th European Conference. Amsterdam, 2016: 21-37.
- [18] Zhang S F, Wen L Y, Bian X, et al. Single-shot refinement neural network for object detection [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, 2018: 4203-4212.