选矿与原料预处理

基于改进 Mask R-CNN 的矿石类型检测算法

肖成勇[a]，李 擎[a]，栗 辉[a]，王 莉[a]，陈子一[a]，张德政[b]，车伟杰[a]
（北京科技大学 a. 自动化学院；b. 计算机与通信工程学院，北京 100083）

摘要：针对不同的矿石类型进行选矿工艺参数设置和操作有利于保障精矿质量，提高回收率和降低物资消耗。由于缺乏矿石类型实时检测有效手段，选矿过程的精准操作目前较难实现。为此，本文提出一种全面改进的 Mask R-CNN 矿石类型检测算法。算法步骤：① 使用 ResNetV1d-50 提取矿石图像各阶段的特征图，并在主干网络中加入可变形卷积以便增强异形矿石的特征；② 改进 FPN，通过在主干网络的 C5 特征层加入特征残差模块，并融合到 P5 特征层，得到具有更强语义信息的多尺度特征图；③ 改进 RPN，设计自适应的正样本 IOU 选取方案来匹配宽高比异常的矿石，进一步提高异形矿石的识别精度；④ 在 RoIAlign 网络基础上加入 Global-Context，以提高小矿石的检测能力；⑤ 在数据增强和训练技巧方面对模型进行改进。结果表明，本文算法的平均精度为 67.92%，平均交并比为 63.54%，分别比基准模型提高了 13.67% 和 9.71%。本文研究方法在矿石类型识别领域具有较好的应用价值。

关键词：Mask R-CNN；矿石类型识别；可变形卷积；训练技巧
中图分类号：TD92；TP391.4 **文献标识码**：A **文章编号**：1000－8764(2024)02－0065－09
doi：10.13403/j.sjqt.2024.02.026

Ore type detection algorithm based on improved Mask R-CNN

XIAO Chengyong[a], LI Qing[a], LI Hui[a], WANG Li[a], CHEN Ziyi[a], ZHANG Dezheng[b], CHE Weijie[a]
(University of Science and Technology Beijing a. School of Automation; b. School of Computer and Communication Engineering, Beijing 100083, China)

Abstract: The setting and operation of beneficiation process parameters for different ore types are conducive to ensuring the quality of concentrate, improving the recovery rate and reducing the material consumption. Due to the lack of effective means for real-time detection of ore types, it is difficult to achieve accurate operation in the beneficiation process. Therefore, a comprehensive improved Mask R-CNN ore type detection algorithm is proposed. Algorithm steps: 1) ResNetV1d-50 is used to extract the feature maps of each stage of the ore image, and deformable convolutions are added to the backbone network to enhance the characteristics of the special-shaped ore; 2) To improve FPN, a multi-scale feature map with stronger semantic information is obtained by adding a feature residual module to the C5 feature layer of the backbone network and fusing it into the P5 feature layer; 3) Improve the RPN, and design an adaptive positive sample IOU selection scheme to match the ore with abnormal aspect ratio, so as to further improve the identification accuracy of special-shaped ore; 4) Global-Context is added on the basis of the RoIAlign network to improve the detection ability of small ores; 5) The model is improved in terms of data augmentation and training skills. The results show that the average accuracy of this algorithm is 67.92%, and the average intersection ratio is 63.54%, which is 13.67% and 9.71% higher than the

收稿日期：2023－07－22；**修回日期**：2023－09－13
基金项目：科技创新 2030—重大项目(2020AAA0108702－01)；教育部第二批新工科研究与实践项目(E-ZDH20201602)；教育部高等学校自动化类专业教学指导委员会专业教育教学改革研究课题(202104&202149)；教学部产学合作协同育人新工科建设项目(202101320001&2021010460001)；北京科技大学教育教学改革与研究面上项目(JG2021M29&JG2021M28)。
作者简介：肖成勇(1978—)，男，博士，高级工程师，从事智能制造、矿山自动化、智能化方面研究。
通信作者：李 擎(1971—)，男，教授，博士生导师，从事智能信息处理、智能优化理论及其应用方面的研究。

benchmark model. The research method has good application value in the field of ore type identification.

Key words: Mask R-CNN; ore type identification; deformable convolution; training skills

在矿物加工领域，识别不同的矿石类型以及计算其混合比例对精矿质量的控制起着至关重要的作用。传统的岩性识别主要依靠专业人员制备岩石薄片对标本进行分析[1]，然而，这种方法工作量大、时间周期长，无法满足选矿生产实时检测的要求，容易导致精矿品位波动大、尾矿品位高。为了提高矿山低品位矿产资源的利用率，并确保选矿过程中精矿品位的稳定性，首先要解决矿石类型的自动识别问题。近年来，随着计算机图像处理技术，特别是深度学习的快速发展，岩石种类识别方法取得了较大进展。如：程国建等[2]利用卷积神经网络自动提取薄片显微图像特征，构建了岩石颗粒自动分类网络，实现了真实样品的自动识别，取得了良好的应用效果；此外，为减少岩石薄片图像鉴定的主观性，将卷积神经网络应用到岩石薄片图像检索中，避免了传统岩石薄片鉴定的主观性[3]；张野等[4]建立了基于 Inception v3 的岩石图像深度学习迁移模型，该模型能够有效识别和分类花岗岩、千枚岩和角砾岩；冯雅兴等[5]提出了一个能兼顾岩石全局图像信息和局部纹理信息的孪生卷积神经网络，实现了具有新鲜面的岩石图像自动识别；柳小波[6]等则利用 Faster R-CNN 深度学习框架，成功识别了橄榄岩、玄武岩、大理岩、片麻岩、砾岩、石灰岩等多种岩石图像，解决了多岩石混合场景下无法区分识别岩石种类的难题。

目前岩石识别方法主要以岩石薄片数字图像为分析对象，很少直接采用原始岩石图像作为数据来源，因此所得的识别模型很少能直接应用到生产实践当中。本文通过对 Mask R-CNN 网络的全面改进，采用改进的 ResNetV1d 和加入特征残差模块的特征金字塔(feature pyramid network，FPN)作为图像特征提取网络，得到更强语义信息的多尺度特征图；在区域提交网络(region proposal network，RPN)中，设计自适应的样本交并比(intersection over union，IOU)选取方案，以匹配宽高比异常的矿石，同时在 RoIAlign 中加入全局语境以提高对小矿石的检测能力，最后将模型应用到选矿实际生产场景中，为解决实际应用问题提供参考。

1 矿石类型自动识别检测算法流程

本文提出的皮带上矿石类型自动识别方法包括三个核心步骤：① 样本集制备：采集皮带上矿石影像数据，进行标注、双边滤波、Z-score 标准化、图像增广等处理，形成样本集；② 模型训练：进行全面改进 Mask R-CNN 网络训练，将样本集输入模型得到检测结果；③ 后优化处理：采用随机权重融合方法集成预测，提高矿石类型识别精度。皮带上矿石类型识别流程如图 1 所示。

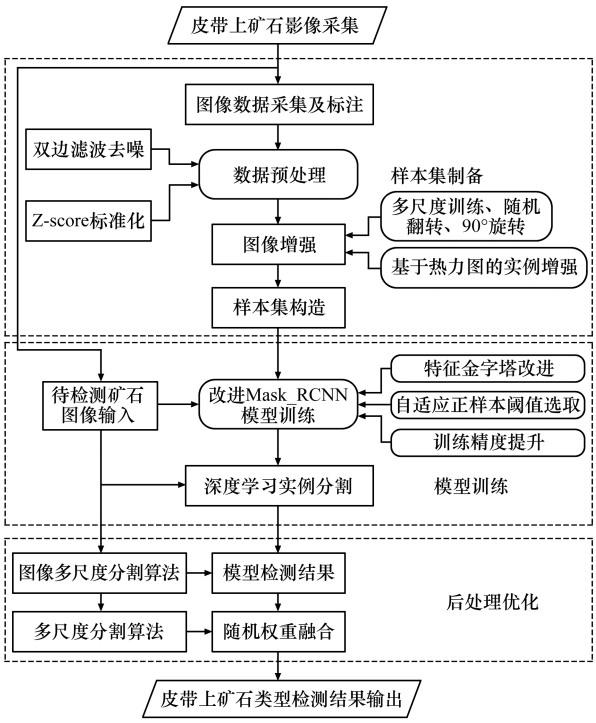


图 1 皮带上矿石类型检测流程

Fig. 1 Flow chart of ore type detection on belt

2 模型设计

2.1 Mask R-CNN 模型

目标检测是计算机视觉领域的一项重要任务，通常可以分为单阶段检测和二阶段检测两大类。其中，单阶段检测算法包括 YOLO、SSD、RetinaNet 等，具有速度快、结构简单的特点，但其精度较

低[7]；而二阶段检测网络则从 Fast R-CNN 到改进的 Faster R-CNN，旨在提升检测速度和精度[8-9]。2017 年，HE 等[10]提出了 Mask R-CNN，该算法是在 Faster R-CNN 基础上融入特征金字塔网络以提取图像的多尺度特征，进一步提升目标检测的精度，在遥感图像建筑物检测[11]、水下生物目标识别[12]和磨损颗粒识别[13]、公路裂缝检测[14]等领域中取得了令人满意的结果。

2.2　改进的 Mask R-CNN 模型

本文提出的改进 Mask R-CNN 由五部分组成，如图 2 所示。其中，改进的特征提取部分由 ResNetV1d-50 和特征金字塔网络（FPN）组成，区域提交网络（RPN）、分类定位网络（box-head）和分割网络（mask-head）分别用于提取候选区域，对目标进行分类和定位，以及对目标个体进行分割输出。

2.2.1　特征提取网络改进

改进的特征提取网络如图 3 所示，包含 ResNetV1d 和特征金字塔网络的改进。为了同时保证矿石图像特征提取能力和模型运行速度，本文采用 ResNetV1d 作为特征提取部分。在改进的 ResNetV1d 中，对 ResNet 原始结构的一些不合理之处进行改进，减少了参数和计算量的同时提升了精度[15]。具体改进如下。

（1）改进 ResNet 的第一个 7×7 的卷积结构（Conv1）。与一层 7×7 卷积相比，Inception-v2 采用三层 3×3 卷积，可以减少参数量和计算量，并且能够达到与一层 7×7 卷积相当的感受野（receptive field），同时提升了非线性表现。因此，本文将第一个 7×7 卷积替换为三层 3×3 卷积，并将前两层卷积层的通道数减半，进一步减少参数量，具体的网络结构如图 3（a）所示。

（2）改进下采样模块。在原始的 ResNet 中，下采样是通过具有步长为 2 的 1×1 卷积来实现的。然而，这样的下采样会导致网络丢失 3/4 的特征信息。为了解决这个问题，本文使用了 3×3

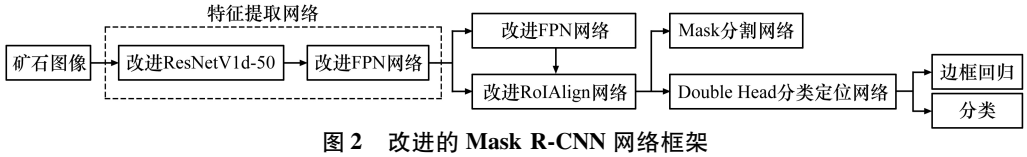


图 2　改进的 Mask R-CNN 网络框架

Fig. 2　Improved Mask R-CNN network framework

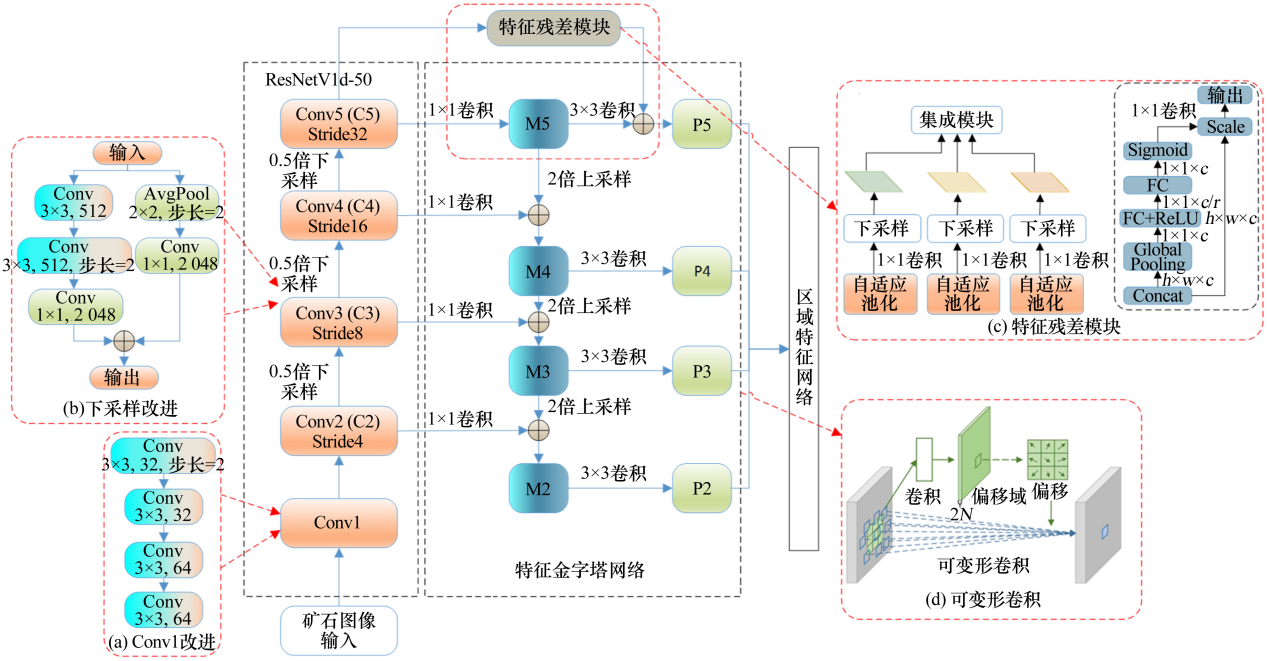


图 3　改进的多尺度特征提取网络

Fig. 3　Improved multi-scale feature extraction network

（注：h、w 分别为高度、宽度，m；c 为通道数；r 为分辨率；N 为偏移采样点的数量，个）

卷积层来进行下采样，而不会丢失特征图的信息。同时，在残差支路中，在1×1卷积之前增加了一个2×2的均值池化操作，进一步提高网络的性能，如图3(b)所示。

(3) 加入可变形卷积。为了增强模型对于异形矿石的建模能力，引入了可变形卷积(deformable conv)[16]，如图3(d)所示。可变形卷积是在传统卷积操作上加入了一个偏移量，从而使卷积变得不规则。这样的设计能够更好地适应具有不同形状的矿石，提高模型的识别能力。

2.2.2 特征金字塔网络改进

本文采用特征金字塔网络(FPN)来对改进的ResNetV1d生成的特征图进行处理，得到特征图{P2,P3,P4,P5}。其中，{P2,P3,P4}是经过金字塔信息融合后包含高层特征的特征层，而P5是主干模型的C5特征层经过1×1卷积降维后得到的特征层。然而，由于C5特征层的通道数减少，其信息损失较大。为了解决这个问题，本文设计了一个特征残差模块，将C5特征层与M5特征层逐像素相加，以补偿其语义信息的损失。

特征残差模块详细结构如图3(c)所示。首先，将主干网络的C5特征层进行降采样，分别得到$\alpha_1\times$C5、$\alpha_2\times$C5、$\alpha_3\times$C5大小的特征图，其中α_1、α_2、α_3分别为0.1、0.2、0.3；接下来，通过使用1×1卷积将自适应均值池化的结果进行降维，将每个输出的特征通道数变为256；然后，将下采样的特征图通过SENet中的注意力机制结构[17]进行融合，统一上采样成C5特征层大小的宽高，同时保持特征图通道数为256；随后，将这三个特征图拼接在一起，并进行特征图压缩操作，通过全局均值池化将每个二维特征通道变成一个实数，使其具有全局感受野，并将输出的维度与输入的特征通道数相匹配。

使用两个全连接层来捕捉特征通道之间的关系。首先，第一个全连接层起到降维的作用，然后，通过ReLU激活函数，再通过一个全连接层恢复原始的维度；最后，将学习到的各个特征通道的激活值(通过Sigmoid激活)与原始特征相乘，得到最终的特征输出。整个过程可以看作是学习了各个特征通道的权重系数，使得模型对各个通道的特征具有更好的辨别能力。

2.2.3 区域提议网络(RPN)

在区域提议网络(RPN)中，将改进的不同尺度特征图{P2,P3,P4,P5}传入RPN网络，以获取一组矩形候选区域。通常情况下，在Mask R-CNN的RPN阶段中，如果IOU大于0.7，就将其选为正样本，如果IOU小于0.3，则将其选为负样本[18]。然而，这种方法存在一些缺点，因为基于锚点的设置也是一组超参数，一旦参数设定完成，锚点的大小就固定了。这会导致如果数据集中存在一些具有异常长宽比的真实标注(ground truth，GT)，就可能无法匹配到这些真实标注，从而导致漏检或者误检。

针对该问题，本文设计了自适应正样本IOU选取方案。具体方法：首先，在特征图上铺设预设的锚点，然后在每个层级的特征图上为图像上每个GT选择k个锚点；选取规则是选择距离GT中心点的欧式距离最小的锚点，将这些锚点作为该GT的正样本候选集；假设共有L个金字塔层级，则正样本候选集有$L\times k$个锚点；接下来，计算该GT与正样本候选集中所有锚点的IOU，并计算IOU的均值m_g和标准差v_g；根据公式$t_g=m_g+v_g$计算IOU阈值；最后，从正样本候选集中选择大于t_g且锚点中心点落在GT内的锚点作为最终的正样本，将其余的锚点作为负样本。

RPN网络的损失函数$L(\{p_i\},\{t_i\})$如式(1)所示，由分类损失函数和回归损失函数构成。分类损失部分$L_{\text{cls}}(p_i,p_i^*)$使用Sigmoid和二分类交叉熵函数进行计算，如式(2)所示；回归损失部分$L_{\text{reg}}(t_i,t_i^*)$使用$Smooth_{L_1}(x)$函数进行计算，如式(3)所示。需要注意的是，当锚点被分配为负样本时，该函数不参与坐标框的回归。

$$L(\{p_i\},\{t_i\})=\frac{1}{N_{\text{cls}}}\sum_i L_{\text{cls}}(p_i,p_i^*)+\lambda\frac{1}{N_{\text{reg}}}\sum_i p_i L_{\text{reg}}(t_i,t_i^*) \tag{1}$$

$$L_{\text{cls}}(p_i,p_i^*)=-\log[p_ip_i^*+(1-p_i^*)(1-p_i)] \tag{2}$$

$$smooth_{L_1}(x)=\begin{cases}0.5x^2\times\dfrac{1}{\sigma^2} & \text{if } |x|<\dfrac{1}{\sigma^2}\\ |x|-0.5 & \text{otherwise}\end{cases} \tag{3}$$

式中：p_i 为第 i 个锚点预测为真实标签的概率；p_i^* 为正样本时值为 1，为负样本时值为 0；λ 为权重系数；N_{cls} 为一个最小批中的所有样本数量；N_{reg} 为锚点位置的个数；t_i 为预测第 i 个锚点的边界框回归参数；t_i^* 为第 i 个锚点对应的 GTBox 的边界框回归参数；x 为预测输出和实际标签之间的差值；σ 为阈值参数，用于界定平方损失和线性损失的切换点。

2.2.4　感兴趣区域对齐层(RoIAlign)

在感兴趣区域对齐层(RoIAlign)中，本文采用 Global-Context 技巧，将整张特征图作为一个感兴趣区域(region of interest，RoI)，并通过均值池化操作将其压缩到与 RoIAlign 处理后的特征图相同的大小；然后，将这两者拼接在一起，并将辅助信息传入后续的分类定位网络。

2.2.5　基于 Double Head 的分类定位网络改进

在 Mask R-CNN 中分类和检测都用同一个特征图，但实际上分类任务只需要保证平移和尺度的不变性，而对于坐标回归任务，需要保证目标位置和形状的变化反应在特征上，进而回归得到位置，即平移和尺度的相等性。这样共用一个特征图就限制了目标检测的性能提升。因此，本文采取 Double Head 结构[19]，除了公用的 RoI 特征，使用改结构将分类和坐标回归任务进行解耦。采用分类任务对两个全连接层进行细化的分类，而坐标回归任务则采用五个残差模块[类似图 3(c)]来回归目标的位置。采用这样的结构可以进一步提高目标检测的性能。

损失函数为多分类的交叉熵和 $smooth_{L_1}(x)$ 损失函数，其中，多分类交叉熵(L_{cls})如式(4)所示，$smooth_{L_1}(x)$ 损失函数如式(3)所示。

$$L_{\text{cls}} = \sum_{i=1}^{N} -p_i \times \log(p_i^*) \tag{4}$$

2.2.6　分割网络

RoIAlign 层的输出特征图同时输入到分割网络(mask-head)，用于预测分割掩膜。mask-head 采用 FCN 模型对每个输入的 $14 \times 14 \times 256$ 维特征图进行分割。具体操作包括首先进行四个 $3 \times 3 \times 256$ 的卷积操作，然后将结果输入到反卷积层进行上采样操作，得到 $28 \times 28 \times 256$ 的特征图，使其与输入的尺寸相同。这样可以得到每个像素属于各个类别的置信度，从而输出像素级实例分割结果。

Mask 分割网络使用的损失函数也是二分类的交叉熵函数(L_{mask})，如式(5)所示。

$$L_{\text{mask}} = -y\log y' - (1-y)\log(1-y') \tag{5}$$

式中：y' 为网络预测结果，取值范围为 0～1；y 为真实标签，取值为 0 或 1。

2.2.7　训练精度提升技巧

本文除了数据增强和改进模型外，还采用以下三种模型训练技巧来提高准确率。

(1) 学习率衰减策略：本文采用 gradual warmup 结合 Cosine decay 的学习率衰减策略来使模型学习的过程更加平滑。前半段采用 gradual warmup 策略，在前 200 个迭代周期内将学习率从 0.000 04 逐渐增加到预设的学习率 0.008 00，后半段采用 Cosine decay 策略，让学习率随着训练过程曲线下降。

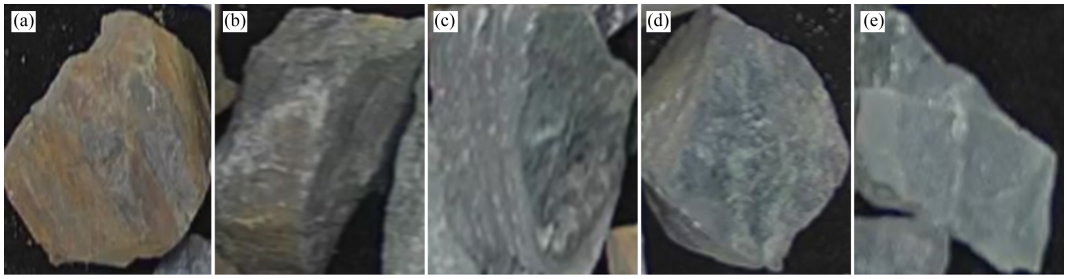
(2) 预训练模型使用：由于数据集经过数据增强后样本量较小，直接进行网络模型训练容易出现过拟合，因此本文采用 COCO 预训练模型来初始化原有网络层参数，以加快模型的训练并抑制模型的过拟合。

(3) 模型集成方法：本文采用随机权重平均(stochastic weight averaging，SWA)方法来得到损失值更低、泛化性更好的模型[20]。SWA 通过在不同的训练阶段组合相同网络的权重，然后使用这种组合权重的集成模型进行预测。SWA 只需要两个单独的模型，第一个模型 W_{swa} 用来存储模型权值的平均值，这将在训练结束后得到最终的模型，并且用于预测；第二个模型 W 用于穿越权值空间，并且使用循环学习率进行探索。

3　试验数据采集及预处理

3.1　矿石样本数据采集和标注

本文使用的原始数据来自鞍山某矿山，包括岩石、赤铁矿、碳酸铁、绿泥石和亚铁五类主要的矿岩类型，矿石的形态如图 4 所示。在专家指导下，首先给各类矿石贴上不同颜色的标签，具体如下：赤铁矿(红色)、碳酸铁(蓝色)、绿泥石(绿色)、亚铁(黑色)和岩石(没有贴标签)；然



（a）赤铁矿；（b）绿泥石；（c）亚铁；（d）碳酸铁；（e）岩石

图 4 矿石图像示意

Fig. 4 Ore image sketch

后，随机挑选各类矿石混合，将带矿石标签的一面朝上放置，拍摄一张图片，作为标签图片；接着，将带标签那一面朝下，其余三面没标签的朝上拍摄，作为正式数据；最后，利用 Labelme 进行标注，将得到的 json 文件转换成 coco 数据集格式，形成 362 张标签图片，其中 300 张用于训练集，62 张用于测试集。经过上述处理后，获得带标注图片，如图 5 所示。

3.2 图像预处理及图像增强

由于现场输送过程存在灰尘，数据中存在一些模糊情况。本文设计双边滤波去噪算法，并对图像进行 z-score 标准化操作，以加强模型训练时的稳定性。为了避免昂贵的数据标记工作，丰富数据集并提高模型鲁棒性，对标注图片进行数据增强操作。本文采用多尺度训练、随机左右翻转和随机 90°旋转等方法对图片进行增强。

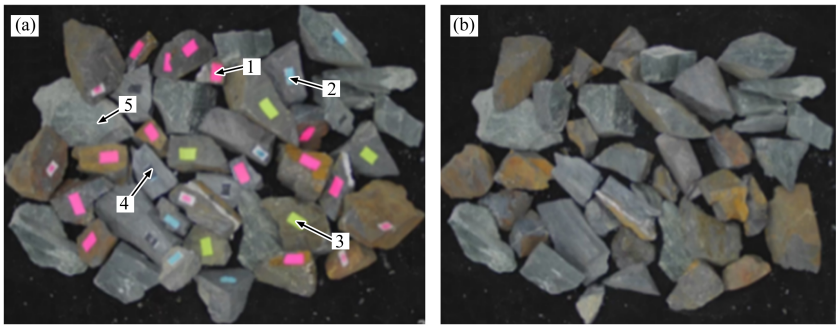
3.3 试验平台参数

试验运行平台采用 Inter E5 - 2640 处理器、两块 Nvidia GTX1080Ti 12GB 显卡，选用 Ubuntu 16.04 操作系统，采用 Pytorch 框架进行模型代码编写。主干网络采用 ResNetV1d-50，预训练模型使用在 COCO 数据集上预训练的 Mask R-CNN 的主干网络模型。归一化方法采用批归一化（batch normalization）。可变形卷积（DCN）在主干模型的第三至第五阶段，通过将 residual block 的中间 3×3 卷积替换来实现。

在 RPN 阶段，锚点的基础尺寸（base_scale）设置为 8，锚点比例设置为三个比例[0.5, 1.0, 2.0]，下采样倍数为[4, 8, 16, 32, 64]。RoIAlign 阶段，分类定位网络（box-head）输出的特征图统一大小为 7×7，分割网络（mask-head）输出的特征图统一大小为 28×28。

试验共使用两块显卡，分别同时处理两张图片，因此批次（batch_size）为 4。模型优化方法采用带动量的小批次随机梯度下降算法，初始学习率设置为 0.005，动量参数（momentum）设置为 0.9，权重下降参数（weight_decay）设置为 0.000 1，模型总共训练轮次为 24 轮。

在最后的模型预测阶段，RPN 阶段挑选了 1 000 个建议区域（proposal），而非极大值抑制



（a）带标签矿石图像；（b）矿石图像

1—赤铁矿；2—碳酸铁；3—绿泥石；4—亚铁；5—岩石（不带标签）。

图 5 矿石标注示意

Fig. 5 Schematic diagram of ore labeling

(NMS)的阈值设定为 0.7。在 R-CNN 阶段，只要输出置信度为大于 0.2 的结果，mask 的正样本阈值则设定为 0.5。此外，采用软化非极大值抑制(Soft-NMS)策略，阈值设定为 0.4。对于每张图片，最多输出 100 个结果。

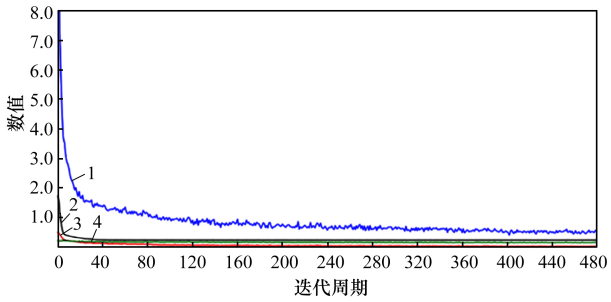
3.4　模型效果评价指标

在模型效果评价指标方面，采用实例分割领域公认的平均精度(average precision，AP)和平均交并比(mean intersection over union，MIoU)来评估模型性能，以衡量模型在复杂环境下对矿石分割的能力。

4　试验结果与讨论

4.1　模型训练过程

图 6 为模型的各部分损失函数值曲线图，包括分类、定位和分割损失，以及总的损失函数值。由图 6 可见：在前 20 轮训练时，分割损失、分类损失和定位损失函数曲线下降比较快，模型具有较好的学习能力；在后续迭代中，曲线下降并稳定，模型具有较好的泛化能力；在进行 400 次迭代后模型的总损失值趋于稳定，训练终止。



1—总损失函数；2—分割损失函数；3—分类损失函数；4—定位损失函数。

图 6　各部分损失函数曲线

Fig. 6　Loss function curve of each part

4.2　试验结果分析

4.2.1　不同数据增强方法对任务网络的影响

为了验证数据增强对减少过拟合的有效性，进行一系列对比试验，结果如表 1 所示。由表 1 可以看出：多尺度训练对于精度的提升效果最为明显，随机翻转和随机 90°旋转也能够得到不错的提高；而基于热力图的增强方法在两个评测指标上也有 0.46 和 0.51 的提升。

表 1　不同数据增强方法的识别效果对比

Table 1　Comparison of recognition effects of different data enhancement methods

评测指标	平均精度/%	平均交并比/%
Mask R-CNN	54.25	53.83
+多尺度训练/测试	56.93	54.96
+随机旋转	57.46	55.81
+随机 90°翻转	58.13	56.24
+基于热力图的增强	58.59	56.75

4.2.2　特征残差模块超参数试验对比

在对 C5 特征层进行自适应均值池化时，涉及到三个超参数$\{\alpha_1,\alpha_2,\alpha_3\}$。为此，本文进行 1 组超参数对比试验，结果如表 2 所示。在权衡模型参数和精度后，本文选取$\{0.1,0.2,0.3\}$作为$\{\alpha_1,\alpha_2,\alpha_3\}$的最终值。

表 2　特征残差模块对比试验

Table 2　Comparison test of characteristic residual module

α_1	α_2	α_3	平均精度/%	平均交并比/%
	基线		54.25	53.83
0.1	0.2	0.4	54.79	53.26
0.1	0.2	0.5	54.71	53.22
0.1	0.2	0.3	54.83	54.27

4.2.3　自适应 IOU 阈值试验对比

自适应正样本 IOU 阈值的选取需要在每个特征层上挑选 k 个锚点。为此，本文进行 1 组对比试验对超参数 k 进行调整，结果如表 3 所示。由表 3 可以看出，当 k 为 7 时，平均精度为最高值，考虑到 k 的增大会导致计算量增加，因此本文最终选取 k 为 7。

表 3　自适应 IOU 阈值对比试验

Table 3　Adaptive IOU threshold comparison experiment

k	平均精度/%	平均交并比/%
基线	54.25	53.83
3	55.04	54.16
5	55.52	54.41
7	55.84	54.67
9	55.79	54.68
11	55.81	54.65

4.2.4　改进模型及融入不同技巧后的试验对比

为了比较改进模型以及不同技巧的效果，进

行对比试验，结果如表 4 所示。由表 4 可知，通过数据增强、warmup + cosine decay、COCO 预训练、随机权重平均、ResNetV1d、DCN、Global-Context、Double-head 和全连接融合等改进和技巧的引入，模型的性能得到了显著的提升；其中，使用 COCO 数据集的预训练模型对精度提升最为明显。

表 4 试验结果分析

Table 4 Analysis of experimental results

评测指标	平均精度/%	平均交并比/%
Mask R-CNN + 数据增强	58.59	56.75
+ warmup + cosine decay	58.83	56.88
+ COCO 预训练	62.45	59.47
+ 随机权重平均	63.39	60.19
+ ResNetV1d	65.13	61.38
+ DCN	66.27	61.92
+ Global-Context	66.94	62.47
+ Double-head	67.86	62.81
+ 全连接融合	67.92	63.54

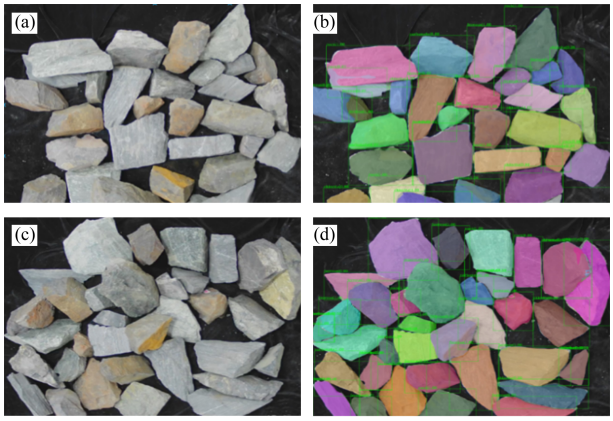
本文模型在测试集和现场图片的具体识别效果分别如图 7、8 所示。由图 7、8 可以看出，对于块度较大的矿石类型，识别效果较好，但是对于一些边缘的小石块的检测结果还有待提高。

5 结 论

（1）本文对 Mask R-CNN 模型进行了全面改进，使用 ResNetV1d-50 来提取矿石图像各阶段的特征图，并加入可变形卷积增强异形矿石的特征。通过学习率预热、预训练模型的使用和随机权值平均等模型训练技巧有效地提高了训练精度。

（2）针对特征金字塔网络高层特征的语义信息丢失，提出了特征残差模块，并针对预设锚点无法匹配长宽比异常的目标物，提出了自适应正样本 IOU 阈值方案，能够提高对小矿石以及异形矿石的识别精度。

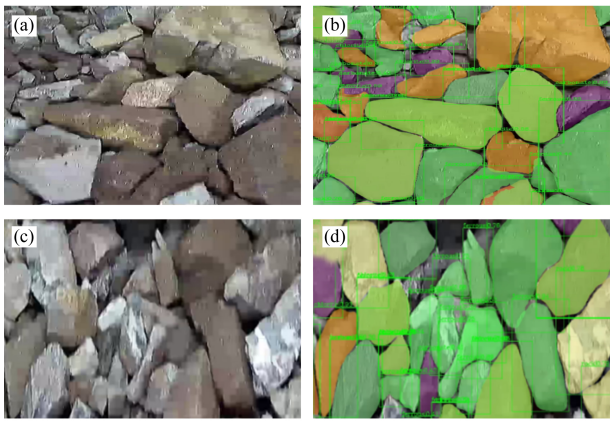
（3）Mask R-CNN 模型在矿石类型识别方面仍处于探索阶段，预测速度尚未达到现场需求，仍然存在较大改进空间，如使用模型裁剪以及 tensorRT 加速技术是后续研究方向。



（a）原图 1；（b）Mask R-CNN 结果 1；（c）原图 2；（d）Mask R-CNN 结果 2

图 7 测试环境矿石类型识别效果

Fig. 7 Identification effect of ore type in testing environment



（a）原图 1；（b）Mask R-CNN 结果 1；（c）原图 2；（d）Mask R-CNN 结果 2

图 8 现场环境矿石类型识别效果

Fig. 8 Recognition effect of ore types in the field environment

参考文献：

[1] 李燕. 基于深度学习的岩石图像识别[D]. 北京:北京林业大学,2020:4 - 19.
LI Yan. The study of rock image identification based on deep learning[D]. Beijing: Beijing Forestry University, 2020:4 - 19.

[2] 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.
CHENG Guojian, GUO Wenhui, FAN Pengshao. Study on rock image classification based on convolution neural network[J]. Journal of Xi'an Shiyou University(Natural

Science Edition),2017,32(4):116 – 122.

[3] 程国建,岳清清. 卷积神经网络在岩石薄片图像检索中的应用初探[J]. 智能计算机与应用,2018,8(2):43 – 46.
CHENG Guojian, YUE Qingqing. Feasibility study of convolutional neural network in image retrieval of rock slices[J]. Intelligent Computer and Applications,2018,8(2):43 – 46.

[4] 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.
ZHANG Ye, LI Mingchao, HAN Shuai. Automatic identification and classification in lithology based on deep learning in rock images[J]. Acta Petrologica Sinica,2018,34(2):333 – 342.

[5] 冯雅兴,龚希,徐永洋,等. 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(5):89 – 94.
FENG Yaxing, GONG Xi, XU Yongyang, et al. Lithology recognition based on fresh rock image sand twins convolution neural network[J]. Geography and Geo-Information Science,2019,35(5):89 – 94.

[6] 柳小波,王怀远,王连成. 岩石种类智能识别研究的 Faster R-CNN 方法[J]. 现代矿业,2019,601(5):60 – 64.
LIU Xiaobo, WANG Huaiyuan, WANG Liancheng. Faster R-CNN method for intelligent identification of rock species[J]. Modern Mining,2019,601(5):60 – 64.

[7] 王月明,李世翔,翟容清,等. 基于 YOLOv3 算法的烧结厂台车箅条缺失检测研究[J]. 烧结球团,2021,46(5):30 – 34;53.
WANG Yueming, LI Shixiang, ZHAI Rongqing, et al. Research on detection of missing trolley grate bars in sintering plant based on YOLOv3 algorithm[J]. Sintering and Pelletizing,2021,46(5):30 – 34;53.

[8] GIRSHICK R. Fast R-CNN[C]//IEEE International Conference on Computer Vision(ICCV). Las Vegas: IEEE,2016:1440 – 1448.

[9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]//Advances in neural information processing systems. Montreal: NeurIPS,2015:91 – 99.

[10] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision(ICCV). Venice: IEEE, 2017: 2961 – 2969.

[11] 胡舒,王树根,王越,等. 基于 Mask R-CNN 的高分遥感影像建筑物目标检测[J]. 测绘地理信息,2023,48(3):50 – 54.
HU Shu, WANG Shugen, WANG Yue, et al. Building object detection in high-resolution remote sensing image based on Mask R-CNN[J]. Journal of Geomatics,2023,48(3):50 – 54.

[12] 宋绍剑,朱靖旭. 基于 Mask R-CNN 和迁移学习的水下生物目标识别研究[J]. 计算机应用研究,2020,49(6):386 – 391.
SONG Shaojian, ZHU Jingxu. Feature residual module contrast test research on underwater biological target recognition based on Mask R-CNN and transfer learning[J]. Application Research of Computers,2020,49(6):386 – 391.

[13] 杨智宏,贺石中,冯伟,等. 基于 Mask R-CNN 网络的磨损颗粒智能识别与应用[J]. 摩擦学学报,2021,41(1):105 – 114.
YANG Zhihong, HE Shizhong, FENG Wei, et al. Intelligent identification of wear particles based on Mask R-CNN Network and Application[J]. Tribology,2021,41(1):105 – 114.

[14] 张跃飞,王敬飞,陈斌,等. 基于改进的 Mask R-CNN 的公路裂缝检测算法[J]. 计算机应用,2020,40(S2):162 – 165.
ZHANG Yuefei, WANG Jingfei, CHEN Bin, et al. Pavement crack detection algorithm based on improved Mask R-CNN[J]. Journal of Computer Applications,2020,40(S2):162 – 165.

[15] 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(CVPR). Las Vegas: IEEE,2016:770 – 778.

[16] ZHU X, HU H, LIN S, et al. Deformable ConvNets v2: More deformable, better results[C]//Proceedings of the IEEE Conference on Computer Cision and Pattern Recognition(CVPR). Salt Lake: IEEE, 2018: 9300 – 9308.

[17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Cision and Pattern Recognition (CVPR). Salt Lake: IEEE,2018:7132 – 7141.

(下转第 106 页)

[7] 黄卡玛,杨晓庆. 微波加快化学反应中非热效应研究的新进展[J]. 自然科学进展,2016,16(3):273-279.
HUANG Kama, YANG Xiaoqing. Recent advances in nonthermal effects of microwave accelerated chemical reactions[J]. Progress in Natural Science,2016,16(3):273-279.

[8] 唐军旺. 微波辐射下NO转化的研究[D]. 大连:中国科学院大连化学物理研究所,2001.
TANG Junwang. Conversion of NO with microwave irradiation[D]. Dalian: Dalian Institute of Chemical Physics of Chinese Acade,2001.

[9] ZAMUDIO M A,BENSAID S,FINO D,et al. Influence of the $MgCo_2O_4$ preparation method on N_2O catalytic decomposition[J]. Industrial & engineering chemistry research,2011,50(5):2622-2627.

[10] XU W T, WANG Q G, PENG K, et al. Development of $MgCo_2O_4$-$BaCO_3$ composites as microwave catalysts for the highly effective direct decomposition of NO under excess O_2 at a low temperature[J]. Catalysis Science & Technology,2019,9(16):4276-4285.

[11] MA Z X,SHENG L P,WANG X W,et al. Oxide catalysts with ultrastrong resistance to SO_2 deactivation for removing nitric oxide at low temperature[J]. Advanced Materials,2019,31(42):1903719.

[12] CONNER W C, TOMPSETT G A. How could and do microwaves influence chemistry at interfaces[J]. The Journal of Physical Chemistry B,2008,112(7):2110-2118.

[13] TANG J, ZHANG T, LIANG D, et al. Direct decomposition of NO by microwave heating over Fe/NaZSM-5[J]. Applied Catalysis B Environmental,2002,36(1):1-7.

[14] PENG K, ZHOU J C, XU W T, et al. Microwave irradiation-selective catalytic reduction of NO to N_2 by activated carbon at low temperature[J]. Energy & Fuels,2017,31(7):7344-7351.

[15] XU W T,ZHOU J C,OU Y P,et al. Microwave selective effect:A new approach towards oxygen inhibition removal for highly-effective NO decomposition by microwave catalysis over $BaMn_xMg_{1-x}O_3$ mixed oxides at low temperature under excess oxygen[J]. Chemical Communications,2015,51(19):4073-4076.

[16] TOFAN C, KLVANA D, KIRCHNEROVA J. Decomposition of nitric oxide over perovskite oxide catalysts: effect of CO_2, H_2O and CH_4[J]. Applied Catalysis B Environmental,2002,36(4):311-323.

[17] WANG X D,ZHANG T,XU C H,et al. Microwave effects on the selective reduction of NO by CH_4 over an In-Fe_2O_3/HZSM-5 catalyst[J]. Chemical Communications,2000(4):279-280.

(上接第73页)

[18] 王海云,王剑平,张果,等. 改进FPN的Mask R-CNN工业表面缺陷检测[J]. 制造业自动化,2020,12(42):35-40.
WANG Haiyun, WANG Jianping, ZHANG Guo, et al. Industrial surface defect detection based on improved FPN Mask R-CNN[J]. ManufacturingAutomation,2020,12(42):35-40.

[19] WU Y, CHEN Y, YUAN L, et al. Rethinking classification and localization for object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Seattle:IEEE,2020.

[20] IZMAILOV P, PODOPRIKHIN D, GARIPOV T, et al. Averaging weights leads to wider optima and better generalization[Z/OL]. arXiv,(2018.3.14)[2019.2.25]. https://arxiv.org/abs/1803.05407.

(上接第82页)

[21] 钟文,衷水平,唐定,等. 紫金山铜矿微裂纹特性及其对磨矿动力学的影响[J]. 有色金属(选矿部分),2022(2):58-65.
ZHONG Wen, ZHONG Shuiping, TANG Ding, et al. Characteristics of microcracks and its effect on grinding dynamics in Zijinshan copper mine[J]. Nonferrous Metals(Mineral Processing Section),2022(2):58-65.