

融合 CBAM 的 Mask R-CNN 模型在球团识别与 粒径测量中的应用

王 猛^{a,d}, 刘卫星^{b,d}, 李 喆^{a,d}, 李 浩^{a,d}, 齐西伟^d, 杨爱民^{c,d}

(华北理工大学 a. 冶金与能源学院; b. 综合测试分析中心; c. 理学院; d. 铁矿石优选与铁前工艺
智能化河北省工程研究中心 河北 唐山 063210)

摘要: 球团粒径的大小是影响高炉透气性、高炉冶炼效率与能源消耗的主要因素之一。本文针对工业条件下球团粒径难以精准测量的问题,采用融合注意力机制 Mask R-CNN 模型对球团进行分割与粒径测量。在对球团图像进行预处理后,构建了球团数据集,对比了多种主干网络的训练表现,并与多个分割模型进行了精度对比。此外,利用像素点统计分割掩膜面积实现了球团粒径的测量。结果表明,ResNet50 作为主干网络在球团的特征提取中更具优越性。引入 Convolutional Block Attention Module (CBAM) 的 Mask R-CNN 模型对比初始模型 A_{mean} 提高了 2.18%。对比 BlendMask、SOLOv2、YOLACT 以及 CondInst 等分割模型,改进后的模型在分割精度上也有优势,并能更好地处理分割细节。此外,与 Image J 测量的球团粒径相比,本文所提出的球团粒径测量方法的最大误差保持在 ± 1.8 mm 之内, $A_{\text{IoU}=0.5}$ 可达到 0.948 3。

关键词: 球团粒径; Mask R-CNN; 迁移学习; ResNet; CBAM

中图分类号: TF046.6; TP18

文献标志码: A

文章编号: 1000-8764(2025)01-0085-10

doi: 10.13403/j.sjqt.2025.01.011

Application of Mask R-CNN model combined with CBAM in pellets identification and particle size measurement

WANG Meng^{a,d}, LIU Weixing^{b,d}, LI Zhe^{a,d}, LI Hao^{a,d}, QI Xiwei^d, YANG Aimin^{c,d}

(North China University of Science and Technology a. Institute of Metallurgy and Energy; b. Comprehensive Test and Analysis Center; c. College of Science; d. Hebei Engineering Research Center for Iron Ore Optimization and Pre-Iron Process Intelligence, Tangshan 063210, Hebei, China)

Abstract: The pellet size is one of the main factors affecting blast furnace air permeability, blast furnace smelting efficiency and energy consumption. In order to solve the problem that the pellet size is difficult to measure accurately under industrial conditions, the Mask R-CNN model combined with attention mechanism is used to segment and measure the pellet size. After preprocessing the pellets image, the pellet dataset is constructed, the training performance of a variety of backbone networks is compared, and the accuracy is compared with multiple segmentation models. In addition, the pellet size is measured by using the pixel statistically segment mask area. The results show that ResNet50 is superior as the backbone network in feature extraction of pellets. The Mask R-CNN model of Convolutional Block Attention Module (CBAM) is introduced, which improve by 2.18% compared to the initial model A_{mean} . Compared with BlendMask,

收稿日期: 2024-07-23; 修回日期: 2024-07-30

基金项目: 河北省教育厅青年科学基金资助项目(QN2024226)。

作者简介: 王 猛(2000—)男,硕士研究生,从事冶金智能制造、工业大数据等方面的研究。

通信作者: 刘卫星(1985—)男,副教授,从事烧结球团智能制造、矿物结构特征定量化表征、钢铁大数据深度挖掘及建模、冶金资源高效利用等方面的研究。

SOLOv2, YOLAT and CondInst, the improved model also has advantages in segmentation accuracy and can better handle segmentation details. In addition, compared with the pellet size measured by Image J, the maximum error of the pellet size measurement method proposed in this paper is kept within ± 1.8 mm, and $A_{IoU=0.5}$ can reach 0.948 3.

Key words: pellet size; Mask R-CNN; transfer learning; ResNet; CBAM

球团生产工艺包括原料准备、配料、混料、造球、干燥、焙烧、冷却、成品以及返矿处理等工序。球团粒径在高炉冶炼的过程中直接影响反应效率,进而影响能源的消耗,根据文献资料以及球团矿质量国家标准(GB/T 27692—2011)规定,造球过程中制备的生球粒径一般要求控制在 8 ~ 16 mm^[1]。在造球过程中,混合料经过水的毛细力和机械力的作用,在旋转和挤压的过程中逐渐形成球状颗粒。但是由于物料的流动性不同,以及球团的压制力度不匀等一系列不确定因素,经常会出现球团粒径分布不均匀的问题,进而影响球团的焙烧质量^[2]。另外,在高炉冶炼过程中,球团粒径分布会影响炉渣量和渣铁比。球团粒径分布不均匀会导致部分球团无法被完全还原,进而生成较高比例的渣铁,影响高炉冶炼效率。目前,球团筛分的方法主要是圆孔网筛筛分法。该方法效率低下,耗时耗力,具有一定的滞后性^[3-4]。因此,对球团粒径检测方法的研究仍有很大的发展空间。

随着计算机视觉技术的迅猛发展,球团粒径的检测方法有了新的思路。将工业相机放置于工业球团生产线进行实时拍摄取样,再将采集到的图像进行数字图像处理,使得非接触式实时测量球团粒径成为可能^[5]。很多学者在颗粒检测领域已经取得了较多研究进展。张建立等^[6]结合实际工业环境,采用遗传算法结合阈值的方法对矿石图像进行识别分割。实验结果理想,验证了此方法在实际矿石粒度检测中的可行性。而阈值分割法分割效果依赖于阈值的选取,阈值选取不当会造成图像分割效果不佳;温春友等^[7]、BAI 等^[8]将分水岭算法分别引入球团粒度以及煤粉粒度图像分割研究之中。但基于传统分水岭算法对噪声极其敏感,易造成过分割;针对重叠颗粒的图像分割问题,刘小燕等^[9]、WU 等^[10]提出了一种基于形态学重建和高斯混合模型的球团图像分割算法,在一定程度上改善了黏连球团图像的分割。

以上传统图像分割算法虽然分割效果都相对理想,但大都需要人们手动设计人工特征,严重的依赖设计者的知识储备和设计经验,并且分割效果针对不同的场景也难尽人意,存在较大的偶然性^[11]。近年来,基于深度学习的图像分割算法在图像分割领域得到了迅速发展。而在复杂的图像背景以及噪声干扰的条件下,深度学习模型中的卷积神经网络(CNN)表现出其优异的性能。目前 CNN 已经广泛应用到目标检测领域,主要由两部分组成:一是一阶段目标检测算法如 YOLO、YOLACT、SSD 等^[12-14];二是二阶段目标检测算法如 R-CNN、Fast R-CNN、Faster R-CNN 和本文所采用的 Mask R-CNN 等^[15-18]。相比较而言,二阶段算法通常表现更高的检测精度,更适用于环境复杂的识别场景。YANG 等^[19]使用 Faster R-CNN 作为铁矿石目标检测算法的框架,实现了 97.0% 的识别率。HE 等^[20]提出 ME Mask R-CNN 算法对 11 种障碍物进行细粒度检测,使列车障碍物检测精度达到 91.3%。WANG 等^[21]通过设计多尺度融合的新型特征金字塔对 Mask R-CNN 模型改进,在钢轨缺陷检测识别任务上实现了 98.7% 的平均精度。Mask R-CNN 多应用于实例分割,其保留了二阶段检测算法中所有的优点,在分类和回归基础上又加入了一个全卷积网络(fully convolutional networks, FCN)构建的 Mask 分支,这就为检测复杂工业环境下球团粒径及分布提供了新思路。

1 模型理论与方法设计

1.1 Mask R-CNN 架构原理

HE 等^[18]在 Faster R-CNN 框架基础上提出了一种创新的目标检测算法,被称为 Mask R-CNN。它的优势在于引入了一个全连接的分割网络模块,将原本的两个任务(分类和回归)转化为三个任务(分类、回归和分割)。这一结合目标检测与图像分割的新算法,显著提高了检测结果的精度和准确性,为复杂场景下的物体识别提供了更强的能力。

Mask R-CNN 作为二阶段的目标检测算法,在第一阶段利用残差网络(residual networks, ResNet)作为主干网络,并结合了特征金字塔(feature pyramid networks, FPN)、区域建议网络(region proposal networks, RPN)生成一系列候选建议框^[22-23]。在第二阶段,通过 ROI Align 操作,将这些建议框准确地映射到特征图上,得到每个候选区域的特征表示。最后,通过分类、边界框回归和掩膜三个并行的分支进行目标预测。整体网络结构如图 1 所示。

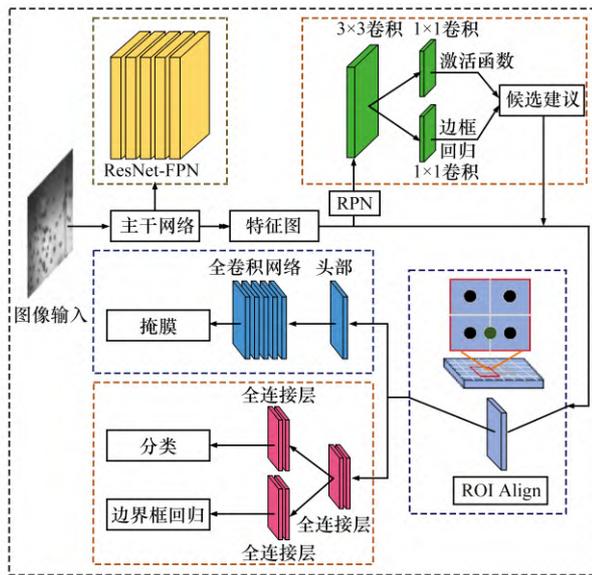


图 1 Mask R-CNN 网络结构

Fig. 1 Structure of Mask R-CNN network

1.2 Mask R-CNN 架构改进

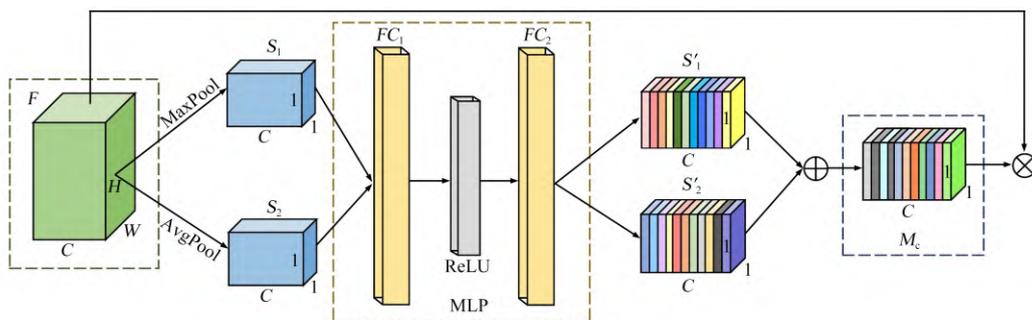
通过工业摄像机现场拍摄得到的数据图像,虽然进行了图像预处理等一系列图像对比度增强操作,但其像素还是较低,存在目标被遮挡、烟尘大、光线不好等问题。为减少上述问题对模型

识别的干扰,提高模型精度、鲁棒性和泛化性,本文拟添加注意力机制对 ResNet 改进。

在计算机视觉中,注意力机制主要被应用于三个维度:通道维度、空间维度以及混合维度。通道维度着重注意特征通道,空间维度重点关注重要区域,而混合维度则是二者的结合。目前在计算机视觉中常见的注意力机制包括通道注意力网络模型(squeeze-and-excitation networks, SENet)、坐标注意力模型(coordinate attention, CA),以及结合了空间和通道注意力的 CBAM 模型。SENet 注意力机制对于模型性能的提升具有一定效果。然而,由于引入了一系列额外的计算过程,SENet 注意力机制存在较高的计算成本。CA 注意力机制涉及对整个特征图执行全局平均池化和全局最大池化操作,以及进行一次全连接层计算,这也将导致计算开销较高。此外,CA 注意力机制未考虑特征图的空间信息,可能在需要空间位置感知的任务中导致性能下降,并限制通道之间的有效交互。这种情况可能导致某些有益通道被忽略,从而影响模型性能。

CBAM 作为一个轻量级注意力模块,主要包含通道注意力和空间注意力。通过引入 WOO 等^[24]提出的 CBAM 模块后,能够显著增强神经网络对图像中目标的关注度,从而有效改善网络的特征学习和表达能力。

图 2 展示了通道注意力模块的结构。图 2 中,尺寸为 $C \times H \times W$ 的输入特征图首先通过两个并行的操作,分别是最大化池(MaxPool)和平均化池(AvgPool),生成了两个尺寸为 $C \times 1 \times 1$ 的一维特征图 S_1 、 S_2 。这些特征图是通过共享具有多层感知机(multi-layer perception, MLP)结构的连接层



⊕—按通道求和操作; ⊗—哈达玛积操作。

图 2 通道注意力机制

Fig. 2 Mechanism of channel attention

FC_1 、 FC_2 ，并使用 ReLU 非线性激活函数进行处理，生成尺寸为 $C \times 1 \times 1$ 的 S'_1 和 S'_2 。接下来，对特征图按通道进行求和，并通过 Sigmoid 函数进行归一化处理，生成权重 M_c 。然后，将与每个通道权重 M_c 的元素逐一相乘，对 F 进行特征重标定处理，生成能够有效捕捉关键通道信息的特征映射 F' 。整个过程可以用公式(1)、公式(2)来表示：

$$M_c(F) = \sigma\{ \text{MLP}[\text{AvgPool}(F)] + \text{MLP}[\text{MaxPool}(F)] \} = \sigma\{ W_1[W_0(F_{\text{avg}}^c)] + W_1[W_0(F_{\text{max}}^c)] \} \quad (1)$$

$$F'_c = M_c \otimes F \quad (2)$$

式中： σ 表示 Sigmoid 函数，用于产生 0 到 1 之间的输出， W_0 和 W_1 分别是 MLP 的第一层和第二层的权重参数。

空间注意力模块结构如图 3 所示。其中，在经过通道特征重标定后的特征图 F' 为输入，先分别进行 MaxPool 和 AvgPool 操作来获得 $P_1 \in R^{1 \times H \times W}$ 和 $P_2 \in R^{1 \times H \times W}$ 特征图。再进行拼接操作，生成 $P_3 \in R^{2 \times H \times W}$ 特征图。随后，经过卷积层处理，对 P_3 中不同位置信息进行编码融合，从而生成空间加权信息 M_s ，用于区分图像中不同空间位置的重要程度。将 F' 作为空间注意力模块的输入，并将其与空间权重系数矩阵 M_s 的对应元素相乘，从而获得显著特征图 F'' ，该过程可用公式(3)、公式(4)表示：

$$M_s(F) = \sigma\{ f^{7 \times 7}[\text{AvgPool}(F) ; \text{MaxPool}(F)'] \} \quad (3)$$

$$F'' = M_s \otimes F' \quad (4)$$

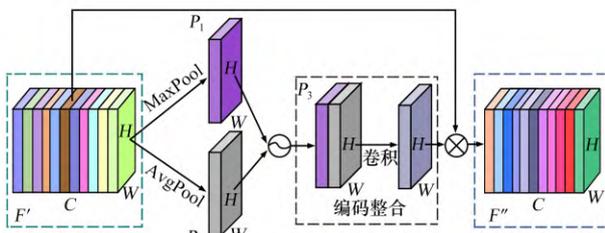


图 3 空间注意力机制示意

Fig. 3 Mechanism of spatial attention mechanism

采用 CBAM 对 ResNet 主干网络进行改进，如图 4 所示。图 4 中，在 C_2 、 C_3 、 C_4 和 C_5 的 1×1 卷积层后分别添加了 CBAM 注意力模块。这样改进的目的是增强 Mask R-CNN 模型在处理球团图像的通道和空间聚焦能力。由于 CBAM 注意力模块

直接整合到 Mask R-CNN 模型中，并参与整个模型的参数迭代更新，随着模型的训练进行，CBAM 注意力权重矩阵也会逐步优化并最终收敛到最优状态。

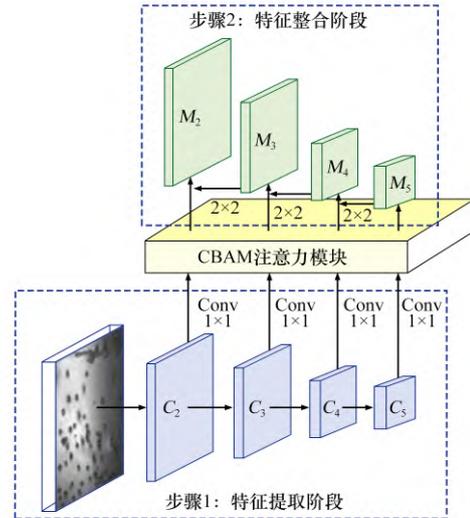


图 4 引入 CBAM 的 ResNet 主干网络

Fig. 4 Backbone network of ResNet introduced into CBAM

未改进前，由图 3 可以将 $M_2 \sim M_5$ 的计算公式表示为

$$M_5 = \text{Conv}_{1 \times 1}(C_5) \quad (5)$$

$$M_i = \text{Conv}_{1 \times 1}(C_i) \oplus f_u(M_{i+1}) \quad i \in \{4, 3, 2\} \quad (6)$$

式中： M_i 为 C_i 经过特征融合后输出的特征图； C_i 为 ResNet 网络第 i 阶段所提取的特征图； $\text{Conv}_{1 \times 1}$ 为进行 1×1 卷积的层次； \oplus 为对应元素的相加操作； f_u 用于描述 2×2 的上采样。

经过改进后， $M_2 \sim M_5$ 的计算公式：

$$M_5 = \text{Conv}_{1 \times 1}(C_4) \otimes A_4[\text{Conv}_{1 \times 1}(C_4)] \quad (7)$$

$$M_i = \text{Conv}_{1 \times 1}(C_i) \otimes A_i[\text{Conv}_{1 \times 1}(C_i)] \oplus f_u(P_{i+1}) \quad i \in \{4, 3, 2\} \quad (8)$$

式中： A_i 为第 i 阶段特征图的 CBAM 注意力权重矩阵； \otimes 为对应元素相乘操作。

2 数据处理

2.1 图像增强

在数据选取上，本文选用工业生产环境下球团圆盘生产图像(通过拍照，采集了 71 张图片)。试验所有球团图像均在实际工业条件下拍摄，为避免球团颜色和造球盘颜色相近而对神经网络提

取球团特征产生的干扰,进一步对图像质量进行增强处理,以提高图像的对比度。

在图像处理中,伽马(γ)变换是一种非线性调整图像对比度的方法,特别是在处理过度曝光或曝光不足的情况下,能够增强图像中较暗区域的灰度值,同时抑制灰度较高区域的对比度。通过这种变换处理后,图像不仅在整体上显得更加清晰和均衡,还能够突出细节^[25]。其数学公式见式(9)。

$$S = cr^\gamma \quad (9)$$

式中: c 为缩放系数,用于调整图像的灰度值范围; r 为灰度值,取值范围在 $0 \sim 1$ 之间, γ 能控制图片亮度的变换。将工业球团图像进行伽马变换,原图像与变换图像如图5所示。

取 $\gamma = 0.5$ 时,对球团图像进行批量处理,得到 γ 变换之后的球团图像。Labelme是一个用于计算机视觉的开放注释工具^[26]。本研究使用Labelme对经过图像增强的球团图像进行数据标注,每张图像约含80颗球团颗粒,总计5982颗球团颗粒。原始数据与标签掩膜示例如图6所示。

2.2 数据增强

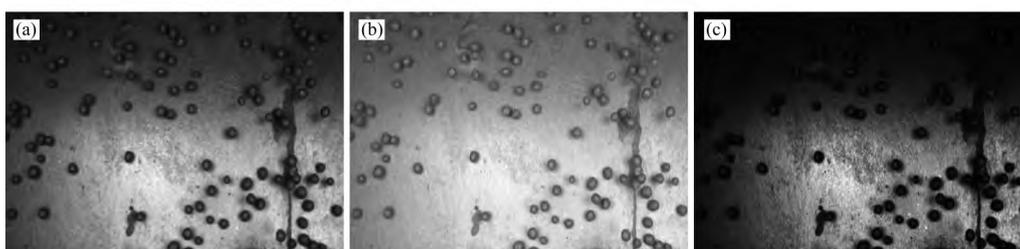
对标注的71张数据图像进行数据增强,其中

操作包括旋转、翻转及其组合。对数据集在原来的基础上进行7倍扩充(图7),以提高模型训练的泛化性及鲁棒性。在经过数据增强后总共有497张图像,总计41874粒球团。将数据集随机地以7:2:1的比例划分为训练集、测试集和验证集。

2.3 试验环境与过程

本文试验环境是Ubuntu 9.4.0-1ubuntu1 ~ 20.04.2操作系统,硬件采用4块NVIDIA GeForce RTX 3090(24 GB)。软件编译环境为Python 3.7,具有CUDA 11.3图形计算平台,编译软件为PyCharm。使用pytorch框架进行训练测试。由于试验设备的限制,仅将批量训练图像设置为1。

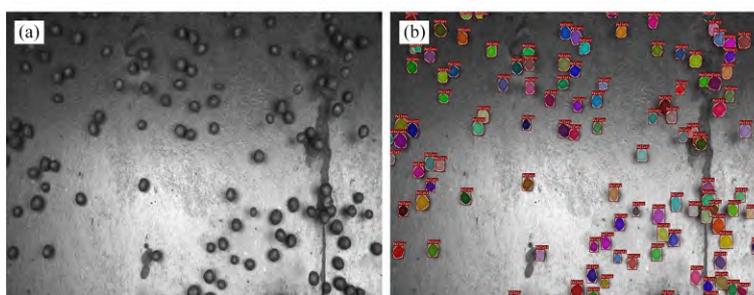
图8为本文所提出的球团分布识别以及粒径测量过程。图8中:首先将经过数据增强处理后的数据集作为输入,使用改进后网络进行训练,得到球团分割之后的图像;再根据分割的结果,对分割图像的像素点面积以及坐标进行统计,以取球团边缘最小外接椭圆短轴的思想,得到球团的数目以及粒径。为了评估结果,试验将测得的球团粒径结果与Image J测量的球团的粒径进行误差比对,最后得到球团粒径分布^[27]。



(a) 原始图像; (b) $\gamma = 0.5$ 时图像的变换效果; (c) $\gamma = 2$ 时图像的变换效果

图5 不同 γ 值下图像的变换效果

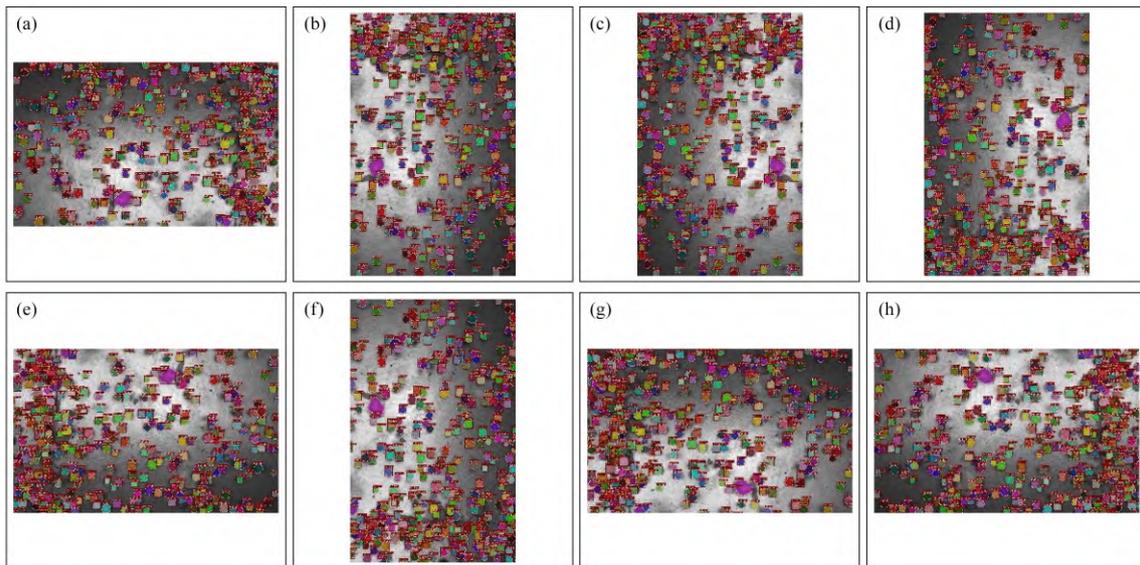
Fig. 5 Effect of image transformation under different γ values



(a) 原始图像; (b) 标记后的图像

图6 原始数据与数据可视化

Fig. 6 Raw data and data visualization



(a) 原始图像; (b) 左旋转 90°; (c) 右旋转 90°; (d) 左旋转 90°加上下翻转;
 (e) 左旋转 180°; (f) 左旋转 270°; (g) 左右翻转; (h) 上下翻转

图 7 数据增强可视化示意

Fig. 7 Schematic of data enhancement visualization

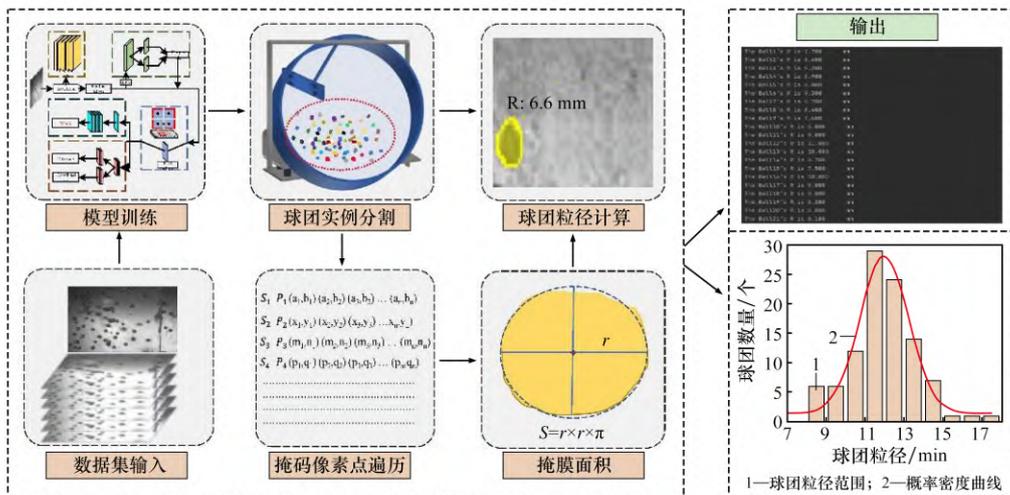


图 8 球团分布识别以及粒径测量过程

Fig. 8 Pellet distribution identification and particle size measurement process

3 试验结果与讨论

3.1 模型评价指标

分类识别问题的模型性能有两个重要的评价指标。一个是精准率，用于评估分类结果中正确识别了多少对象。另一个是召回率，用于评估在总正样本中正确预测了多少个阳性样本。

真实值与预测值关系如图 9 所示。图 9 中， S_{TP} 为正样本被正确区分为正样本的数量， S_{FP} 为负样本被错误区分为正样本的数量， S_{FN} 为正样本

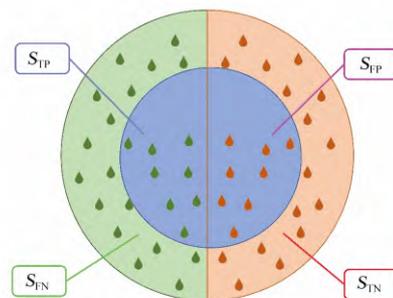


图 9 真实值与预测值的关系

Fig. 9 Relationship between true values and predicted values

被错误区分为负样本数量, S_{TN} 为负样本被正确分类为负样本的数量。将待检测的球团目标定义为正类, 其他类别归为负类, 则精准率、召回率可以分别用式(10)与式(11)表示:

$$P = \frac{S_{TP}}{S_{TP} + S_{FP}} \quad (10)$$

$$R = \frac{S_{TP}}{S_{TP} + S_{FN}} \quad (11)$$

然而精准率和召回率是一对矛盾的度量。当检测器的置信度阈值增加时, 精准率会有所波动并上升, 而召回率下降。因此我们综合考虑精准率和召回率, 采用 $P-R$ 曲线下的面积 Average Precision 作为模型评价指标, 本文用 A 表示。

交并比(intersection over union, IoU) 是用于衡量两个边界框接近程度的指标, 同时考虑了大小和位置, 其值等于交集面积除以并集面积, 范围在 0 到 1 之间。

尽管 $P-R$ 曲线能提供一定的性能评估, 但在球团识别分割的具体细节上还是有所不足。平均精度均值(A_{mean}) 可以评估模型对所有类别的平均精度。小物体平均精度(A_s) 可以评估模型对小尺寸目标(面积小于 32×32 像素)的性能, 中等物体的平均精度(A_m) 可以评估模型对中等尺寸目标(面积大于等于 32×32 像素并且小于 96×96 像素)的性能。因此, 本文以 A_{mean} 、 $A_{IoU=0.5}$ 、 $A_{IoU=0.75}$ 、 A_s 、 A_m 以及模型测试对于一张图片的处理时间(五次测量取平均)作为模型好坏的评价标准。

3.2 主干网络选取

在深度学习中, 主干网络通常由多个卷积层组成, 是特征提取的主要部分。不同的主干网络在不同模型以及不同数据集上的表现也各有差异。本文选取目标检测模型常应用的几种主干网络(包括 ResNet、VGG、DenseNet 以及 MobileNet) 分别作为 Mask R-CNN 的主干网络, 用来对比验证对工业环境下球团特征的提取能力。为了缩短不必要的试验时间, 试验选取部分球团数据集, 设置相同训练参数分别训练 200Epoch, 得到结果见表 1。由表 1 可知, ResNet 作为 Mask R-CNN 的主干网络在球团数据集上的特征提取能力相较 VGG、MobileNet 以及 DenseNet 明显更好。本文应用迁移学习的策略, 将在公开数据集(microsoft common

objects in context, MS COCO) 上预训练的 ResNet50 模型权重应用于球团数据集, ResNet50 作为球团特征提取的主干网络^[28]。

表 1 不同主干网络下的球团特征提取评估结果

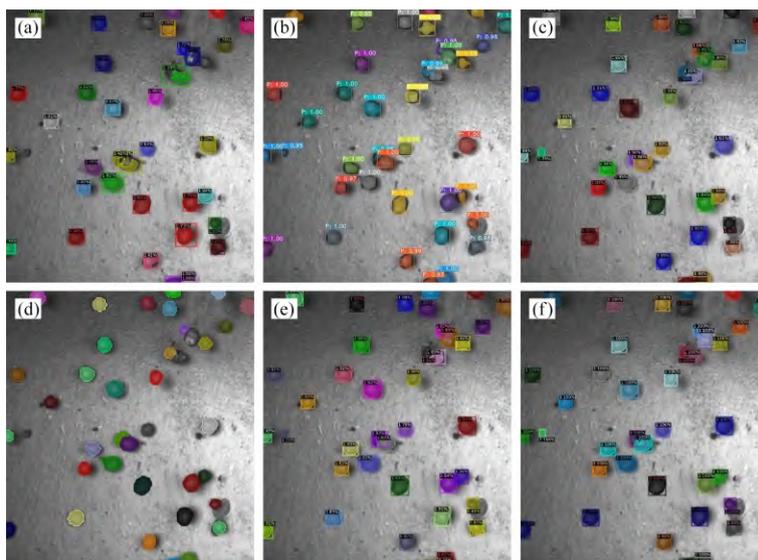
Table 1 Pellet feature extraction evaluation results under different backbone networks

主干网络	A_{mean}	$A_{IoU=0.5}$	$A_{IoU=0.75}$	A_s	A_m
ResNet101	0.129	0.394	0.035	0.019	0.136
ResNet50	0.569	0.849	0.726	0.230	0.591
VGG13	0.200	0.496	0.127	0.001	0.212
VGG16	0.292	0.566	0.277	0.015	0.309
MobileNet_v2	0.032	0.129	0.003	0.020	0.033
MobileNet_v3_small	0.087	0.292	0.016	0.011	0.092
MobileNet_v3_large	0.050	0.182	0.008	0.006	0.052
DenseNet121	0.074	0.256	0.017	0.013	0.078

3.3 球团分割结果

为了验证本模型在球团粒径识别分割的优越性, 本文分别采用相同数据集下均训练 200 000 次迭代的 Mask R-CNN、YOLACT、Blendmask、SOLOv2、CondInst 作为改进的 Mask R-CNN 模型的对比模型, 结果如图 10 所示。由图 10 可知, 在球团的识别效果上, YOLACT 模型表现最差, 在检测上, YOLACT 模型表现出良好的效果, 其识别置信度分数相对较高都在 0.9 以上, 但其分割基本未覆盖球团边缘, 这对于球团的粒度识别会造成巨大误差。SOLOv2、Mask R-CNN、BlendMask 以及 CondInst 在球团数据集的分割上表现了模型的优越性, 但其置信度分数并不高, 大多数球团识别的置信度分数仅在 0.8 以上, 在分割细节上也仍存在一定的提升空间。基于改进 Mask R-CNN 模型的球团识别, 其对于球团的置信度分数更加高, 基本都在 0.95 以上, 甚至多数为 1。另外在边缘处理上也较为平滑, 达到了相对较好的效果, 满足球团识别分割的基础要求。

为了进一步证明改进后的 Mask R-CNN 模型在球团数据集识别分割表现的优越性, 对比了其与其他几个模型的 A_{mean} 指标, 结果如表 2 所示。由表 2 可知, 引入 CBAM 注意力机制的 Mask R-CNN 模型在球团数据集的分割表现最好, A_{mean} 达到了 0.729 6, 相对 Mask R-CNN 模型, 分割精度提高了 2.18%。在另外几个模型中, BlendMask 算法



(a) Mask R-CNN; (b) YOLACT; (c) BlendMask; (d) SOLOv2; (e) CondInst; (f) 改进的 Mask R-CNN

图 10 不同算法下的球团图像识别对比

Fig. 10 Comparison of pellet image recognition under different algorithms

表 2 不同模型的评估结果

Table 2 Evaluation results of different models

模型	A_{mean}	$A_{\text{IoU}=0.5}$	$A_{\text{IoU}=0.75}$	A_s	A_m	执行时间/s
Mask R-CNN	0.714 0	0.942 2	0.850 5	0.342 3	0.783 8	1.14
YOLACT	0.227 9	0.600 8	0.235 1	—	—	0.52
BlendMask	0.719 5	0.958 7	0.821 4	0.325 1	0.740 7	1.20
Solov2	0.527 8	0.784 7	0.611 4	0.359 5	0.541 3	1.34
CondInst	0.627 4	0.932 2	0.783 8	0.178 2	0.648 4	1.67
Mask R-CNN improved	0.729 6	0.948 3	0.874 1	0.385 9	0.744 6	1.56

在球团数据集上的分割表现最好, A_{mean} 达到了 0.719 5, 其次是 Mask R-CNN 和 CondInst, A_{mean} 也都稳定在 0.6 以上。表现相对较差的是 SOLOv2 和 YOLACT 模型, A_{mean} 相对较低。此外, 测定了 6 种模型处理同一张 $1\ 920 \times 1\ 200$ 分辨率的工业造球图像所花费的平均时间, 在耗费时间上看, YOLACT 模型所花费的时间是最短的, 但其精度偏差较大。Mask R-CNN 模型在改进前后的, 在处理图像所花费的时间上耗费时常增加了 0.42 s。以上数据的差异在球团的实际识别过程与效果中也有所体现。

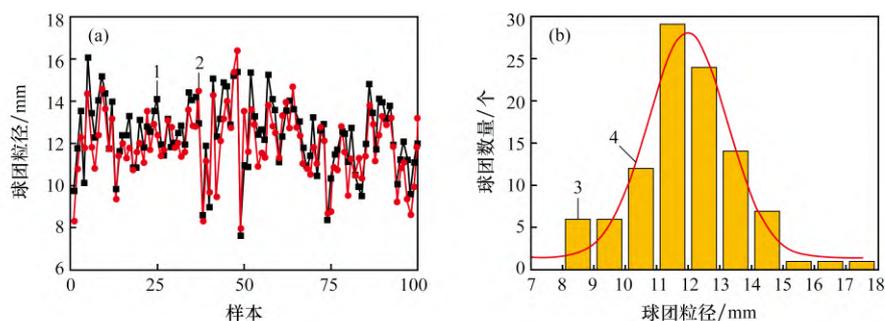
3.4 球团粒度统计

在基于图像的球团粒径测量过程中, 需要将图像像素尺寸以比例尺的形式转换为实际的球团粒径, 再对图像进行计算分析。本文提出了一种基于掩膜求解的方法。在经过实例分割后, 球团

被分割成每一个标记有随机颜色掩膜的实例, 分别对图像中被探测到的球团实例进行编号, 对掩膜像素点进行统计, 遍历掩膜像素点, 进而计算得到二维图像下球团的切面面积。将切面面积近似为标准圆面积, 根据圆的面积公式, 进而得到圆的粒径长度, 并将其作为球团粒径。

与 Image J 测量的球团粒径作对比(为了减少随机误差, 本文进行 4 次试验测量, 取平均值进行对比), 结果如图 11 与表 3 所示。由于数据量相对较大, 表 3 省略了一部分数据, 只给出部分数据。

结合图 11 与表 3 可知, 本文提出的球团粒径测量办法与 Image J 测量最大误差保持在 ± 1.8 mm 之内, 对比 5 次 Image J 测量平均结果与试验结果, 本文提出的方法测量得到球团粒径与 Image J 测量得到球团粒径的平均相对误差为 7.29%。



(a) 粒径识别对比; (b) 粒径分布

1—真实值; 2—预测值; 3—球团粒径范围; 4—概率密度曲线。

图 11 球团粒径识别对比及分布结果

Fig. 11 Comparison of pellet size identification and distribution results of pellet size

表 3 不同方法下球团粒径测量结果对比

Table 3 Comparison of pellet size measurement results under different methods

试验编号	Image J 测量/mm	掩膜求解方法/mm	绝对误差差值/mm	相对误差/%
1	11.756	10.8	0.956	8.132 0
2	13.539	12.3	1.239	9.151 3
3	13.958	13.2	0.758	5.430 6
4	9.796	9.3	0.496	5.063 3
5	11.639	11.4	0.239	2.053 4
6	12.410	12.0	0.410	3.303 8
7	15.164	14.6	0.564	3.719 3
8	13.539	12.9	0.639	4.719 7
9	11.915	11.4	0.515	4.322 3
10	12.559	12.0	0.559	4.451 0
11	11.912	11.2	0.712	5.977 2
12	15.064	14.3	0.764	5.071 7

4 结 论

(1) 本文分别对比了 ResNet、VGG、DenseNet 和 MobileNet 作为 Mask R-CNN 的主干网络在球团数据集上的训练表现, 结果表明 ResNet50 对于复杂环境下的球团特征提取能力更强。

(2) 基于 ResNet50 主干网络, 本文引入注意力机制 CBAM, 更好地帮助主干网络忽略背景区域的影响, 着重关注球团特征部分, 提高了模型的精度和鲁棒性。

(3) 经试验验证, 改进的 Mask R-CNN 模型在 A_{mean} 上较初始模型提高了 2.18%, $A_{\text{IoU}=0.5}$ 可达

到 0.948 3。在球团粒径测量上, 误差可以控制在 ± 1.8 mm 之内, 能够为工业球团生产监测提供一定指导。

参考文献:

- [1] THURLEY M J, ANDERSSON T. An industrial 3D vision system for size measurement of iron ore green pellets using morphological image segmentation[J]. Minerals Engineering 2008, 21(5): 405–415.
- [2] EYDARI M, AMIRFATTAHI R, NAZARI B, et al. An industrial image processing-based approach for estimation of iron ore green pellet size distribution[J]. Powder Technology 2016, 303: 260–268.
- [3] 张学锋, 祝忠阳, 黄永鹤, 等. 基于图像增强的球团粒度检测方法[J]. 烧结球团 2024, 49(3): 81–88. ZHANG Xuefeng, ZHU Zhongyang, HUANG Yonghe, et al. Pellets particle size detection method based on image enhancement[J]. Sintering and Pelletizing, 2024, 49(3): 81–88.
- [4] 覃德波, 陈天宇, 余正伟, 等. 基于图像增强的生球粒度识别方法研究[J]. 烧结球团 2021, 46(6): 82–88. QIN Debo, CHEN Tianyu, YU Zhengwei, et al. Research on recognition method of fresh pellets particle size based on image enhancement[J]. Sintering and Pelletizing, 2021, 46(6): 82–88.
- [5] 江天宇, 薛涛, 李泽政, 等. 智能算法在球团制造流程中的应用[J]. 中国冶金 2022, 32(5): 25–32. JIANG Tianyu, XUE Tao, LI Zezheng, et al. Application of intelligent algorithms in pellet manufacturing process[J]. China Metallurgy 2022, 32(5): 25–32.
- [6] 张建立, 孙深深, 秦书棋. 基于遗传算法最佳阈值分割的矿石图像分割[J]. 科学技术与工程 2019, 19(7):

- 105 – 109.
- ZHANG Jianli, SUN Shenshen, QIN Shuqi. Ore image segmentation based on optimal threshold segmentation based on genetic algorithm [J]. *Science Technology and Engineering* 2019, 19(7): 105 – 109.
- [7] 温春友, 邵志勇. 基于数字图像处理的球团矿粒度检测 [J]. *烧结球团* 2004, 29(2): 38 – 40.
- WEN Chunyou, SHAO Zhiyong. A measuring method of pellets diameter based on digital image processing [J]. *Sintering and Pelletizing* 2004, 29(2): 38 – 40.
- [8] BAI F Y, FAN M Q, YANG H L, et al. Image segmentation method for coal particle size distribution analysis [J]. *Particuology* 2021, 56: 163 – 170.
- [9] 刘小燕, 吴鑫, 孙炜, 等. 基于形态学重建和 GMM 的球团颗粒图像分割 [J]. *仪器仪表学报* 2019, 40(3): 230 – 238.
- LIU Xiaoyan, WU Xin, SUN Wei, et al. Image segmentation of pellet particles based on morphological reconstruction and GMM [J]. *Chinese Journal of Scientific Instrument* 2019, 40(3): 230 – 238.
- [10] WU X, LIU X Y, SUN W, et al. An image-based method for online measurement of the size distribution of iron green pellets using dual morphological reconstruction and circle-scan [J]. *Powder Technology*, 2019, 347: 186 – 198.
- [11] 张华博. 基于深度学习的图像分割研究与应用 [D]. 成都: 电子科技大学 2018.
- ZHANG Huabo. Research and application of image segmentation based on deep learning [D]. Chengdu: University of Electronic Science and Technology of China 2018.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE 2016: 779 – 788.
- [13] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT: Real-time instance segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE 2019: 9156 – 9165.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2016: 21 – 37.
- [15] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE 2014: 580 – 587.
- [16] GIRSHICK R. Fast R-CNN [C]//2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440 – 1448.
- [17] 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.
- [18] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE 2017: 2980 – 2988.
- [19] YANG A M, JIANG T Y, HAN Y, et al. Research on application of on-line melting in-situ visual inspection of iron ore powder based on Faster R-CNN [J]. *Alexandria Engineering Journal* 2022, 61(11): 8963 – 8971.
- [20] HE D Q, QIU Y F, MIAO J, et al. Improved Mask R-CNN for obstacle detection of rail transit [J]. *Measurement*, 2022, 190: 110728.
- [21] WANG H, LI M J, WAN Z B. Rail surface defect detection based on improved Mask R-CNN [J]. *Comput Electr Eng* 2022, 102: 108269.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE 2016: 770 – 778.
- [23] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE 2017: 936 – 944.
- [24] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018: 3 – 19.
- [25] 翟月. 基于小波变换和伽马变换的图像增强技术 [J]. *西安文理学院学报(自然科学版)* 2019, 22(6): 41 – 44.
- ZHAI Yue. Image enhancement technology based on wavelet transform and gamma transform [J]. *Journal of Xi'an University (Natural Science Edition)*, 2019, 22(6): 41 – 44.

(下转第 125 页)

- Experimental research on carbothermal reduction of converter fly ash [J]. *Sintering and Pelletizing*, 2022, 47(6): 76–83.
- [12] MATSUKEVICH I, KULINICH N, ROMANOVSKI V. Direct reduced iron and zinc recovery from electric arc furnace dust [J]. *Journal of Chemical Technology and Biotechnology* 2022 97(12): 3453–3458.
- [13] 刘纲, 干勇, 李士琦, 等. 高炉含锌粉尘还原性影响因素分析[J]. *中国冶金* 2019 29(10): 20–25.
LIU Gang, GAN Yong, LI Shiqi, et al. Analysis of reducing factors of zinc-bearing blast furnace dust [J]. *China Metallurgy* 2019 29(10): 20–25.
- [14] 陈栋. 含多金属硫酸渣制备预还原球团工艺及机理研究[D]. 长沙: 中南大学 2012.
CHEN Dong. Study on the preparation process and mechanism of pre-reduced pellets from multi-metal sulfate slag [D]. Changsha: Central South University, 2012.
- [15] 胡佩伟, 胡超, 胡兵, 等. 钒钛磁铁海砂矿还原过程有价元素变化规律的研究[J]. *钢铁钒钛* 2021 42(5): 10–17.
HU Peiwei, HU Chao, HU Bing, et al. Study on the regulation mechanism of valuable elements in the reduction process of vanadium-titanium magnetite marine placer [J]. *Iron Steel Vanadium Titanium* 2021 42(5): 10–17.
- [16] 李洋, 张建良, 袁骥, 等. 铁酸锌碳热还原动力学及反应机理[J]. *工程科学学报* 2023 45(1): 82–90.
LI Yang, ZHANG Jianliang, YUAN Xiang, et al. Kinetics and reduction mechanism of non-isothermal analysis carbothermal reduction of zinc ferrite [J]. *Chinese Journal of Engineering* 2023 45(1): 82–90.
- [17] 柴志鹏, 丁学勇, 居天华, 等. 含碳球团气固协同还原特点及动力学研究[J]. *材料与冶金学报*, 2021, 20(4): 239–245.
CHAI Zhipeng, DING Xueyong, JU Tianhua, et al. Study on the characteristics and kinetics of gas-solid reduction of iron ore-coal composite pellets [J]. *Journal of Materials and Metallurgy* 2021 20(4): 239–245.
- [18] CHAI Y F, FAN Y J, LI Z C, et al. Kinetics of reduction in stages of pellets prepared from the Bayan Obo iron ore concentrate [J]. *ACS Omega* 2022 7(9): 7759–7768.
- [19] 明守禄, 张芳, 彭军, 等. 氧化铁皮非等温碳热还原动力学研究[J]. *粉末冶金工业* 2023 33(5): 17–24.
MING Shoulu, ZHANG Fang, PENG Jun, et al. Study on the kinetics of non-isothermal carbon thermal reduction of iron oxide sheet [J]. *Powder Metallurgy Industry*, 2023, 33(5): 17–24.

(上接第 60 页)

WU Shengli, PEI Yuandong, CHEN Hui, et al. Evaluation on liquid phase fluidity of iron ore in sintering [J]. *Journal of University of Science and Technology Beijing*, 2008 30(10): 1095–1100.

- [14] 郝德胜. 含硼复合添加剂对烧结矿性能和烧结的影响[D]. 沈阳: 东北大学 2008: 15.
HAO Desheng. Study on the experiment and mechanism of boracic complex additive to the sinter [D]. Shenyang: Northeastern University 2008: 15.

(上接第 94 页)

- [26] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: A database and web-based tool for image annotation [J]. *International Journal of Computer Vision*, 2008 77(1): 157–173.
- [27] ABRÀMOFF M D, MAGALHÃES P J, Ram S J. Image processing with Image J [J]. *Biophotonics international*, 2004 11(7): 36–42.
- [28] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Transactions on knowledge and data engineering*, 2010 22(10): 1345–1359.