

融合双重注意力与多尺度级联网络的 密集球团粒度测量方法

吴鑫^{1,2†}, 李靓杰^{1,2}, 魏建好¹

(1. 湖南工商大学 计算机学院, 长沙 410205; 2. 湘江实验室, 长沙 410205)

摘要: 球团矿是现代高炉炼铁的重要原料, 合理的粒度分布是影响球团矿质量的重要因素之一。目前, 在球团矿生产过程中, 粒度分布检测主要依赖人工离线筛分的方式, 该方法效率低, 测量结果滞后, 难以满足工业生产现场实时性要求, 且密集球团颗粒存在重叠、曲面光影等噪声的干扰, 现有基于图像处理的粒度检测方法对其边缘轮廓检测能力有限, 易出现漏检或过分割现象。鉴于此, 提出一种融合双重注意力和多尺度级联网络的密集球团粒度测量方法 (LMAD-UNet)。所提出方法将两个 UNet 以并行级联的方式拼接作为骨干网络, 增加网络宽度, 降低网络下采样特征损失; 然后, 设计一种轻量多尺度融合模块 (LMulti-Res Block), 以实现多尺度特征提取, 减少模型参数量, 提升推理速度, 同时, 引入双重注意力机制, 增强网络对轮廓特征的提取能力; 最后, 改进损失函数, 加强对不平衡轮廓点数据的学习。实验结果表明, 所提出方法能够对密集球团颗粒进行精准分割, F_1 -score 可达到 96.85%, 整体优于其他对比方法, 且粒度测量速度能够满足现场生产过程中的实时性要求。

关键词: 球团矿; 粒度测量; 机器视觉; 级联网络; 双重注意力; 深度学习

中图分类号: TP751 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1333

引用格式: 吴鑫, 李靓杰, 魏建好. 融合双重注意力与多尺度级联网络的密集球团粒度测量方法 [J]. 控制与决策, 2025, 40(10): 3155-3166.

Dense pellet size measurement method integrating dual attention and multi-scale cascaded network

WU Xin^{1,2†}, LI Liang-jie^{1,2}, WEI Jian-hao¹

(1. College of Computer Science, Hunan University of Technology and Business, Changsha 410205, China;
2. Xiangjiang Laboratory, Changsha 410205, China)

Abstract: Pellet ore is a crucial raw material for modern blast furnace ironmaking, and an appropriate particle size distribution (PSD) is one of the key factors affecting its quality. Currently, the detection of the PSD in pellet ore production mainly relies on offline manual sieving, which is inefficient, provides delayed results, and cannot meet the real-time requirements of industrial production sites. Existing image processing-based particle size detection methods have limitations in accurately detecting the edges and contours of densely packed pellets, often leading to missed detections or over-segmentation. In light of this, a dense pellet size measurement model integrating dual attention mechanisms and multi-scale cascaded networks, termed light multi-residual attention dual-UNet (LMAD-UNet), is proposed. The proposed method concatenates two UNet architectures in a parallel cascading manner to form the backbone network, thereby increasing network width and reducing feature loss during downsampling. Subsequently, a lightweight multi-scale fusion module, referred to as the lightweight multi-residual block (LMulti-Res Block), is designed to facilitate multi-scale feature extraction, minimize model parameters, and consequently enhance inference speed, while a dual attention mechanism is introduced to enhance contour feature extraction. The loss function is also improved to strengthen the learning of imbalanced contour point data. Experimental results show that the proposed method can accurately segment densely packed pellets, achieving an F_1 -score of 96.85%, outperforming other comparison methods, and meeting the real-time requirements of the production process.

Keywords: pellet ore; particle size measurement; machine vision; cascade network; double attention; deep learning

收稿日期: 2024-11-15; 录用日期: 2025-04-03.

基金项目: 湖南省自然科学基金项目 (2023JJ40238, 2023JJ40237); 湖南省教育厅优秀青年基金项目 (23B0596, 22B0648); 国家自然科学基金项目 (62402177).

†通信作者. E-mail: xinwu@hnu.edu.cn.

0 引言

面对严峻的节能减排政策要求,现代高炉炼铁技术正朝高效、低耗、智能化的方向发展.为实现以上目标,改进炼铁原料的质量和炉料配比等条件是最有效的途径.球团矿作为现代大型高炉冶炼的优质原料之一,具有含铁品位高、粒度均匀、透气性好、易还原等优点^[1-2].在高炉冶炼过程中,加大球团矿的使用比例能够有效优化高炉炉料结构,提高高炉利用系数,降低能源消耗,实现节能减排.因此,大力发展和使用球团矿,是我国钢铁工业走向低碳绿色发展的必经之路^[3].目前,国内球团矿生产主要采用圆盘造球工艺,将细磨的铁精矿粉与少量添加剂混合后,在加水湿润的条件下通过圆盘造球机滚动成球,再经过干燥焙烧而成.粒度是衡量球团矿质量好坏的重要指标之一,根据 GB/T27692-2024 高炉用铁球团矿质量标准要求,一级球团矿粒度指标在 8 mm ~ 16 mm 间占比要大于 95%,粒度指标小于 5 mm 占比应小于 3%.粒度分布不均会影响球团的焙烧时间,增加高炉冶炼过程的能源消耗以及二氧化硫等有害气体的排放.因此,在造球过程中,需要持续对粒度分布进行实时监测,并对造球参数进行及时调整,使得粒度分布在合理范围内^[4].现有球团矿粒度测量方法主要采用离线人工取样筛分法,通过筛分后获取球团矿粒度分布,然后调整造球过程参数,但是该方法过程繁琐、耗时长、效率低、检测结果存在明显的滞后性,无法满足工业生产过程实时性测量的要求,且操作工人长期暴露于造球现场高噪声、强粉尘恶劣环境中,存在身体损害以及职业病风险^[5].近年来,随着人工智能技术的快速兴起,基于机器视觉的非接触式测量方法得到了广泛应用,该方法可提供精准、实时的测量结果,已成功应用于钢铁冶炼^[6]、浮选矿^[7]等复杂流程工业的工况识别与检测.然而,造球过程工作在强粉尘、光照不均等复杂场景下,采集图像质量低、噪声大,加之造球盘内生球一直处于运动状态,生球间产生相互粘连重叠,甚至遮挡的现象,导致图像分割易出现误差.因此,基于机器视觉的球团矿粒度测量方案仍然存在诸多困难,针对这些问题,国内外学者已经开展了一些相关研究,取得了重要的研究进展.下面对相关领域的研究进展进行介绍.

1) 基于传统图像算法的粒度测量方法.传统图像粒度测量方法首选采用对球团颗粒图像进行滤波、阈值分割等处理,然后利用分水岭、凹点检测等算法分割球团颗粒,最后对球团轮廓进行圆拟合并进行粒度分布统计测量. Thurlley 等^[8]采用三维视觉检测系统对传送带上的球团粒径分布进行检测,但

是,该方法利用传统的分水岭算法对相互粘连和重叠的球团颗粒进行分割,极易产生过分割现象,导致粒径测量误差大,且三维视觉系统对于环境要求高,高粉尘环境影响其稳定性; Heydari 等^[9]分别采用 Hough 变换、形态学与 SVM 方法、分水岭算法对粘连和重叠的球团进行分割测量,最后利用多数表决技术对各种分割方法进行评价,选取最好的测量结果,但是,该方法无法对粘连重叠严重的生球进行分割,导致测量误差大,且速度上难以满足在线测量; Nanda 等^[10]利用傅里叶变换分析方法对实验室模拟传送带上采集的粘连球团颗粒的边缘图像进行了分析,并以此获取粒度分布测量结果,但是,测试结果局限于对稳定实验室采集的图像进行处理; 刘小燕等^[11]根据出球区生球的灰度分布特征建立了 GMM (Gauss mixture model) 模型,通过对高斯混合模型的求解来实现对重叠粘连生球的分割,最后利用圆拟合方法对分割后的球团粒度进行测量,由于高斯混合模型参数求解过程复杂耗时,该算法难以满足生产现场实时性测量要求; 白志城等^[12]针对球团粒径检测,设计开发了基于图像处理的粒径检测系统,并提出了一种基于自适应标记的优化分水岭分割算法,与手动筛分结果相比,该算法更快捷、准确,但是,目前该系统只局限于对实验室传送带上的球团粒度进行检测.

2) 基于深度学习的粒度测量方法.深度学习可从大量标记的训练样本图像中提取到表达能力强、通用性好的特征.国内外学者将其应用于粒度测量算法中,可大幅提高颗粒图像分割的准确率. Duan 等^[13]针对出料区球团颗粒图像光照不均的情况,提出了一种轻量化深度网络 Lunet 将球团颗粒从背景中提取,然后利用分水岭算法对粘连重叠颗粒进行分割; Duan 等^[14]提出了一种融合 VGG16 与 UNet 的深度特征融合网络用于球团粒度分布的测量,并设计了基于 K -means 聚类的雾气图像识别方法来增强球团颗粒图像的特征显著性,提升了雾气图像中球团粒度测量的准确率; 张学锋等^[15]提出了基于语义分割的球团粒径在线检测算法,该方法在 UNet 的基础上融入了可分离卷积结构,设计了轻量化的 DSC-UNet 前景分割网络,保证了前景分割的鲁棒性; Deo 等^[16]提出了一种基于改进的轻量化的 UNet 网络用于从复杂背景中提取出重叠粘连的球团颗粒,该方法通过在卷积层之间增加的瓶颈层来实现通道信息融合,提高了图像分割精度.以上测量方法均是以出球口稀疏球团颗粒目标为测量对象,虽然与传统方法相比,在分割性能方面有较大的提升,但是,

以上方法均需要利用分水岭算法对分割后的球团颗粒进行分离才能完成粒度测量,在算法实时性方面难以保证.针对密集球团颗粒粒度测量,李福民等^[17]提出了基于 Re-UNet 网络的粒度测量识别模型,该方法摆脱了原始 UNet 网络中的标准特征连接,并将这些特征映射合并到每个空间尺度上,有效增强了图像特征提取,但是,该方法局限于对实验室光照条件良好的球团颗粒图像进行粒度分布测量;林双等^[5]提出了结合 Mask-RCNN 和最小二乘法的球团粒度识别模型,并通过凹点检测算法和最小二乘圆拟合改进算法来解决球团颗粒测量过程中存在的堆叠效应;与其他粒度测量方法不同,Deo 等^[18]将球团颗粒粒度分布定义为 5 种类别,分别利用 VGG16、MobileNet、ResNet50 作为图像特征分类提取器,对采集的球团粒度分布进行分类,实验结果表明,该方法能够取得较好的分类效果,但是,不能给出具体粒度级的分布占比,无法用于对球团粒度质量进行评估.综上,传统粒度测量参数方法与人工特征选取过程受到外界环境光照影响较大,易产生过分割或欠分割现象,导致测量结果出现较大的误差.而基于深度学习的粒度测量方法则存在以下问题:① 粒度分布范围较大,深度网络感受野与粒度尺寸不匹配,导致特征信息提取不准确;② 为增大感受野,深度网络模型往往包含多个下采样池化层,易造成球团颗粒间的微弱轮廓特征信息丢失;③ 球团颗粒图像中轮廓信息与背景像素数量间存在严重不平衡,造成深度网络在训练过程中难以收敛的问题.针对以上问题,本文提出一种融合双重注意力机制与多尺度级联网络的密集球团粒度分布测量方法,能够实现密集球团粒度分布的在线精准检测,对推动球团矿的智能化生产具有重要的理论意义和现实价值.

本文主要内容如下.

1) 提出一种轻量化多尺度融合模块 LMultiRes Block,该模块能够有效融合不同空间分辨率下的球团矿图像特征,获取更丰富的多尺度图像特征信息,降低模型参数量.

2) 提出基于位置注意力和通道注意力结合的双重注意力特征提取方法,加强对球团轮廓特性的提取,实现上下文感知特征的融合,精准的轮廓特征定位能够提高轮廓提取精度.

3) 针对球团颗粒轮廓与背景像素类别不均衡的问题,提出构建二进制交叉熵损失函数 (BCE) 与 Dice Loss 损失函数加权融合的方法,以提高对球团轮廓的检测精度.

1 图像数据采集系统

为实现对球团粒度的精确测量,本文首先设计了视觉测量系统,并构建了图像数据集用于模型训练和测试.钢铁厂造球流程工作在强粉尘、强振动、光照不均等复杂环境中,加之造球盘内颗粒一直处于运动状态,颗粒间存在相互粘连重叠的现象.为了采集清晰稳定的颗粒图像,本文设计了如图 1 所示的图像采集装置视觉测量系统,包括图像采集模块、照明模块和终端处理模块.

该系统安装在安全围栏外,图像采集模块距离采集区域较远,经测算物距为 2.6 m,因此,选用了 50 mm 长焦距镜头,使得颗粒在图像中更加清晰易辨.同时,为提高颗粒图像的亮度,克服外界环境光照的干扰,本文选用了两个大拱形光源灯组共同照明.为使得采集的图像幅面大,颗粒数量更具代表性,本文选用了堡盟 VCXG-13 型号工业相机,其相机分辨率为 1280×1024 ,像素元大小为 $4.8 \times 4.8 \mu\text{m}$,传感器尺寸大小为 1/2 英寸 CMOS 芯片,采集帧率为 145 帧/s,完全能够满足粒度测量的实时性要求.为了得到球团的实际粒径,本文根据针孔成像模型的原理,在图像校准的过程中计算校准因子 α ,表达式如下所示:

$$\alpha = S_c \times \frac{d}{f}. \quad (1)$$

其中: S_c 为相机的像元大小,系统中所选用相机的感光芯片中像素尺寸为 $4.8 \times 4.8 \mu\text{m}^2$,因此 $S_c = 4.8 \mu\text{m}$; d 为物距,在本系统中为 2600 mm; f 为镜头焦距,取 50 mm; 因此,校准因子 $\alpha = 0.25 \text{ mm}$,即图像上每个像素对应的实际物理长度为 0.25,能够满足现在测量 0.5 mm 的精度要求.本文生球球团图像数据集采自国内某大型炼铁厂造球产线,包含正常光照、轻微粉尘干扰 ($\text{PM}_{2.5}$ 浓度为 $80 \sim 120 \mu\text{g}/\text{m}^3$)、设备振动 (振幅 $0.5 \sim 2 \text{ mm}$) 等工况.标注采用 Labelme 工具并进行交叉验证,标注标准为完整颗粒标注闭合轮廓,重叠颗粒标注可见弧段.

2 本文模型构建

2.1 检测框架

深度学习是当前人工智能技术重点研究和应用的方向,其中卷积神经网络是一种强大的特征提取工具,其在图像分类、目标检测、图像分割、关键点预测等领域取得了巨大成果.针对目前球团颗粒尺寸测量的难题,以及现有深度学习方法在测量过程中存在的不足,本文提出一种融合双重注意力和多尺度级联网络的密集球团粒度测量方法 (LMAD-UNet),用于对球团颗粒边缘进行精准提取和定位,

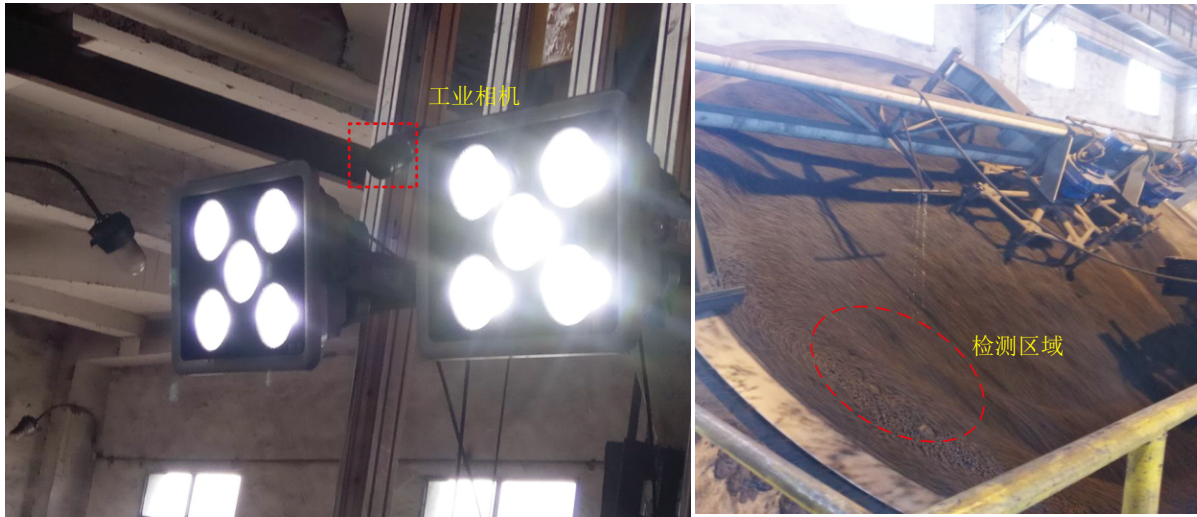
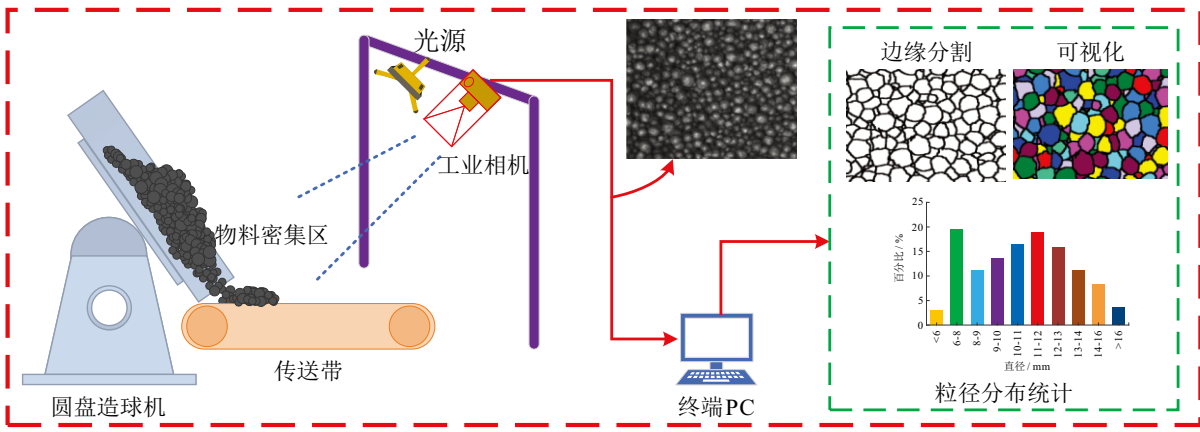


图1 图像采集装置

提高球团颗粒分割精度, 以实现球团颗粒粒度的精准测量. LMAD-UNet 整体结构如图 2 所示. 其主要有 3 个特征: 1) LMAD-UNet 算法的主干特征提取网络选用轻量级网络 UNet, 将两个 UNet 串接为双重 UNet, 以减少网络结构的深度, 增加网络结构的宽度, 用于加强对图像特征的提取, 同时, 为了增加信息传递路径, 减少特征信息前向传递时的丢失率, 在两个 UNet 间添加了跳链接, 并添加了双重注意力模块; 2) 增加了改进的多尺度 Inception 模块和双重注意力机制, 多尺度融合模块能够适应密集颗粒粒度变化范围大的检测要求, 双重注意力机制融合了位

置注意力机制 (PAM) 与通道注意力机制, 分别在空间和通道维度对特征图像进行权值建模, 突出表层颗粒物的流态化特征信息, 加强界面流态化物料信息的提取而抑制背景噪声, 有利于改善边缘细节丢失的弊端, 获得信息丰富的特征图, 以提升算法的精度; 3) 考虑工业环境实时性的要求, 利用深度可分离多尺度卷积核融合方式设计轻量化多尺度模块来替换网络中的常规卷积层, 能够降低模型参数量以及计算量, 提高检测速度.

2.2 轻量多尺度特征融合模块

密集颗粒粒度分布范围广, 具有不同的尺度, 若

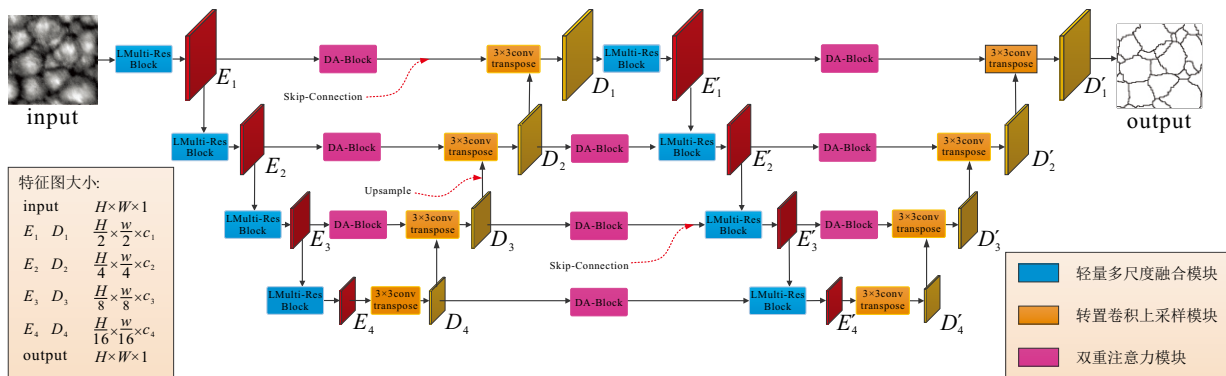


图2 网络总体结构

使用固定的卷积尺度特征进行分割,则可能会导致较小轮廓信息被忽略或较大颗粒物体分割不准确.为了增强卷积网络的鲁棒性,获取足够清晰完整的密集颗粒边界轮廓特征,本文提出了轻量多尺度特征融合模块(LMulti-Res),并将级联UNet网络中的常规卷积模块替换为轻量多尺度融合模块.现有研究主要采用Inception模块的方式来实现网络的多尺度性能,Inception模块将 3×3 、 5×5 和 7×7 的卷积运算并行结合,以融合不同感受野的特征图,从而增强网络的多尺度能力,但是Inception模块卷积核较大,采用并行计算的方式,参数量多,导致计算复杂.为克服以上并行计算存在的问题,减少参数量,提高计算效率,本文提出了如图3所示的LMulti-Res模块.该模块的设计受到卷积核分解思想的启发,即 7×7 、 5×5 等较大的卷积核可分解为若干连续级联的 3×3 卷积核来实现^[19].在LMulti-Res模块中利用3个 3×3 卷积模块通过级联的方式串接,并将级联模块中的3个分支在通道维度上进行拼接即可实现对多尺度特征的融合提取.其中:黄色分支表示 3×3 卷积核提取的特征,绿色分支相当于 5×5 卷积核提取的特征,红色分支相当于 7×7 卷积核提取的特征.通过以上卷积核分解的方式构建多尺度特征提取模型,可提升网络的深度,同时有效减少参数计算量,提高计算效率.为了缓解级联模块训练过程中存在的梯度消失现象,本文在LMulti-Res模块中增加了残差分支路径(图中紫色分支),保留第1层输出中的重要信息,加快模型训练的收敛速度.同时,为进一步降低模块的参数量,实现轻量化,将 3×3 的普通卷积替换为深度可分离卷积(DP-Conv).

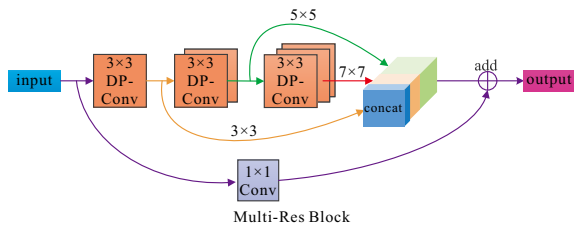


图3 LMulti-Res 模块

2.3 注意力机制

为解决UNet网络中跳跃连接以及网络级联过程将提取的关键特征信息传入解码层时引入过多的无用信息或噪声,同时,使得深度网络模型更关注对于密集颗粒轮廓特征信息的学习,本文在级联网络中引入了双重注意力机制模块(dual-attention block).如图4所示:本文双重注意力机制模块包含通道注意力模块和位置注意力模块,将两者提取的特征图进行融合,加强对图像中感兴趣区域和通道两方面

特征的学习.

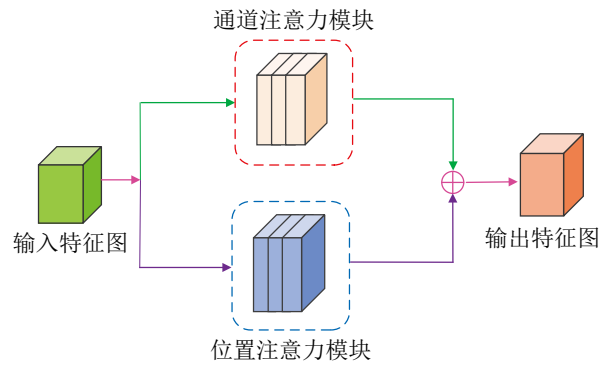


图4 双重注意力机制

通道注意力机制结构如图5所示,通道注意力(CAM)由压缩模块、激励模块和特征加权模块组成^[20].其中:压缩模块利用平均池化(GAP)和全局最大池化(GMP)将原始特征图 $f^{c \times h \times w}$ 中每个通道压缩为 $1 \times 1 \times C$,聚合特征图的通道信息,有效地增大了对精细轮廓特征的感受野;激励模块利用两个 1×1 Conv组成的多层感知机减少通道个数,降低计算复杂度,提高网络模型对特征信息的提取能力;特征加权模块在通道维度上对原始特征图进行加权融合,可获得不同通道高响应区域的特征.经过压缩激励后的权重为

$$\omega_c = \text{MLP}(\text{AvgPool}(f)) + \text{MLP}(\text{MaxPool}(f)). \quad (2)$$

经过Sigmoid激活函数操作后的通道注意力特征 $m_c(f)$ 为

$$m_c(f) = \sigma(\omega_c) = \frac{1}{1 + \exp(-\omega_c)}. \quad (3)$$

通道注意力机制在通道维度上进行特征加权后的特征 f_{CAM} 为

$$f_{\text{CAM}} = m_c(f) \otimes f. \quad (4)$$

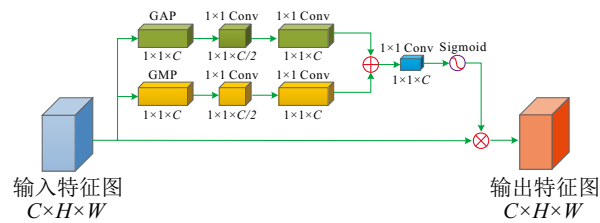


图5 通道注意力机制

位置注意力机制结构如图6所示.位置注意力机制(PAM)用于捕获特征图中任意两个像素位置间的依赖性,即通过所有位置特征的加权和来更新特定的特征,因此,PAM在提取空间特征方面具有重要作用^[21].

首先,PAM结构输入特征 $A \in R^{C \times H \times W}$;然后,将特征A输入卷积层得到3个新的特征图B、C、D,

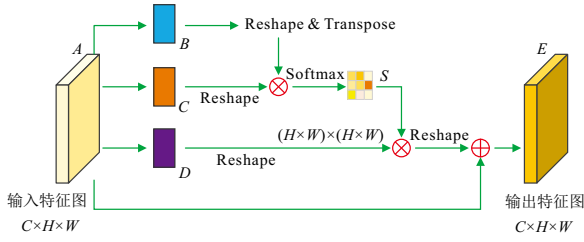


图6 位置注意力机制

大小为 $R^{C \times H \times W}$, 对 B 、 C 进行 Reshape 操作将大小调整为 $R^{C \times N}$ ($N = H \times W$, 表示总像素个数); 接着, 将 B 的转置与 C 做矩阵乘法, 再用 Softmax 层来计算空间注意力图 $S \in R^{N \times N}$, 计算公式如下所示:

$$S_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^N \exp(B_i \cdot C_j)}, \quad (5)$$

其中 S_{ji} 为第 i 个位置的像素对第 j 个位置像素的影响; 然后, 对 D 矩阵做 Reshape 操作, 将大小调整为 $R^{C \times H \times W}$; 最后, 乘以参数 α , 并与特征 A 进行逐元素求和运算, 得到最终输出 $E \in R^{C \times H \times W}$, 有

$$E_j = \alpha \sum_{i=1}^N (S_{ji} \cdot D_i) + A_j. \quad (6)$$

PAM 具有很强的空间特征提取能力. 由于特征 E 是通过所有位置特征与原始特征加权和生成的, 它拥有全局上下文特征并基于空间注意力聚合上下文, 这确保了位置特征的有效提取, 同时维持了全局上下文信息.

2.4 损失函数构建

通常, 在边缘检测任务中存在类别不均衡的问题, 属于背景类别的像素远远少于属于边缘对象的像素, 这会导致模型倾向于关注背景像素而忽略了边缘轮廓像素. 球团矿混合料密集堆叠现象严重, 属于高度不平衡的图像数据, 而传统的损失函数 (BCE Loss) 易因背景像素主导梯度更新而导致模型失效. 经过查阅研究现状发现, Dice Loss 通过计算重叠度, 更关注区域整体的匹配, 对类别不平衡更具有鲁棒性, 从而避免背景像素数量压制前景 (颗粒轮廓) 梯度, 此损失函数已经广泛应用, 其有效性也已经在 Milletari 等^[22] 的医学图像分割文献中得到了验证. 而 BCE Loss 能够作为补充, 对每个像素独立优化, 保留图像细节, 能够修正 Dice Loss 可能忽略的细小断裂. 基于此, 本文使用 BCE Loss 损失函数与 Dice Loss 损失函数加权融合的方法来构建损失函数, Dice 宏观约束区域形状, BCE 微观优化边界细节, 其中二进制交叉熵损失函数 (BCE) 为

$$L_1 = -\frac{1}{N} \sum_{i \in N} [y_i \log(P_i) + (1 - y_i) \log(1 - P_i)]. \quad (7)$$

这里: L_1 为二进制交叉熵损失, N 为图像中像素的总个数, y_i 为真实标签, P_i 为预测值. Dice Loss 计算公式如下所示:

$$L_2 = 1 - 2 \times \frac{\sum_{i=1}^k p_i g_i + \varepsilon}{\sum_{i=1}^k p_i^2 + \sum_{i=1}^k g_i^2}. \quad (8)$$

其中: k 为像素总个数; P_i 表示预测值, 为 Dice 损失; g_i 为真实值; ε 为平滑因子. 最终加权损失函数为

$$L_{\text{final}} = \lambda_1 L_1 + \lambda_2 L_2. \quad (9)$$

这里: λ_1 和 λ_2 分别为 BCE 损失和 Dice 损失的加权因子, 本实验分别取用 0.7 和 0.3 达到实验最佳效果; L_{final} 为最终损失.

3 实验结果与分析

本文的实验环境如下: CPU 为 Xeon(R) Silver 4214R 12 核, GPU 为 GeForce GTX 3080Ti, 显存 12 GB. 使用 Python3.9 编程语言, Pytorch 深度学习框架. 本文中使用的工业相机采集图像大小为 1280×1024 . 由于图像分辨率太大, 直接利用原始图像作为训练数据, 对 GPU 性能要求较高, 本文对原始图像裁剪为 128×128 小图像进行训练, 测试过程将小图像测试结果拼接为完整原始图像. 训练数据集共有图片 600 张, 按照 7:2:1 的比例划分训练集、验证集和测试集. 网络训练轮数 Epoch 设置为 200, 训练批量大小 BatchSize 设置为 6.

3.1 评价标准

本节实验对球团数据集进行了精确率 (Precision)、平衡分数 (F_1 -score)、Jaccard 系数、推理时间 (Inference time)、召回率 (Recall) 等性能指标的分析, 计算了 P-R 曲线以分析模型在数据集上的分割效果. 语义分割的各项指标主要是以混淆矩阵为基础, 如表 1 所示. 其中: T 为预测结果正确 (True), F 为预测结果错误 (False), P 为正样本 (Positive), N 为负样本 (Negative).

表1 混淆矩阵

	正样本(Positive)	负样本(Negative)
预测正确(True)	TP	TN
预测错误(False)	FP	FN

上述性能指标的公式如下所示:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (11)$$

其中: TP为预测正确的正样本数量, FP为预测错误的正样本数量, TN为预测正确的负样本数量, FN为预测错误的负样本数量. 精确率 (Precision) 体现了模型对负样本的分辨能力, 召回率 (Recall) 体现了模型对正样本的识别能力. 但是, 准确率与召回率间相互制约, 需要使用 F_1 分数对两者进行综合, 如下所示, F_1 分数越高, 模型的鲁棒性越好:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (12)$$

P-R 曲线能够直观地体现模型的好坏程度, P-R 曲线下面积越大, 模型的精确率越高, 性能越好.

3.2 分割结果图对比

为了表明 LMAD-UNet 模型分割球团能力的卓越性, 本文将其与最常用的分割算法在相同实验条件和训练策略下进行对比实验. 图 7 为不同算法模型分割效果图, 图 7(a) 为原始图像, 图 7(b) 为图像标签真实值, 图 7(c) ~ 图 7(g) 分别为 U-Net、U-Net++、PSPNet、LadderNet 和所提出算法的分割结果图像. 由图 7 可见, 其他模型对于球体轮廓的分割出现断裂、模糊现象, 而所提出算法对球团边缘整体轮廓分割地更加精确和清晰, 算法的健壮性更强. 其中: U-Net、U-Net++ 和 PSPNet 无法有效提取球团边缘模糊部分的轮廓, 且受到噪声干扰而出现误分割的现象; 所提出算法则在边缘模糊阴影处的分割结果更加清晰.

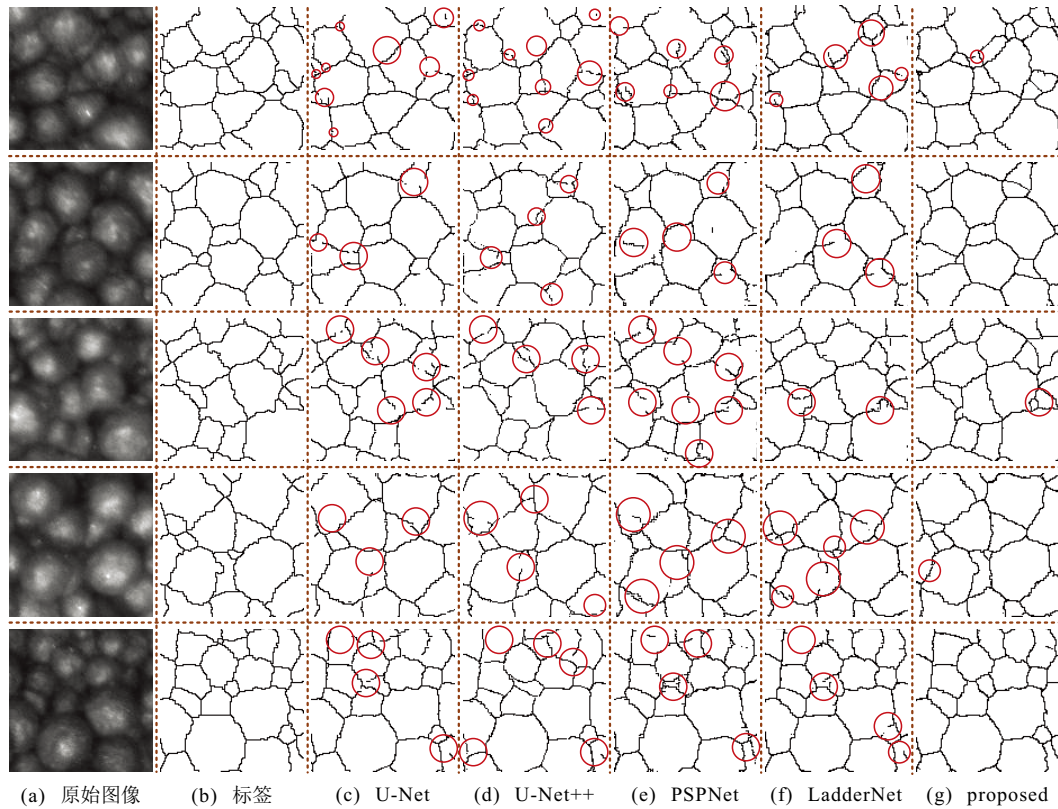


图7 不同算法模型分割效果

由图 7 还可以看出: U-Net、U-Net++ 和 PSPNet 的分割结果中, 受光影的影响导致均出现了许多噪声, 将阴影部分当成球团轮廓分割, 造成了较高的假阳性. 这是因为传统 U-Net 系列方法在密集颗粒场景表现受限, 其主要原因包括: 1) 单一 U 型结构感受野固定, 难以适配 8 ~ 16 mm 大跨度粒度分布; 2) 最大池化操作导致弱边缘特征丢失, 如图 7(c) 中红色圆所示断裂区域; 3) 缺乏注意力机制致使阴影误检率偏高. 相较而言, LadderNet 和所提出算法则

较好地抑制了光影的干扰并减少了误分割现象, 表明了多路径多尺度特征融合的有效性. 基于上述分析, 所提出算法在球团分割上能够更精确地区分背景与轮廓图, 能够呈现出更连续平滑的分割效果, 同时, 能够很好地抑制光影噪声的干扰以提取出边缘轮廓的详细信息. 图 8 为利用不同算法对颗粒图像进行分割后得到的连通区域可视化结果. 由图 8 可见: 当球团颗粒轮廓出现欠分割现象, 边缘出现断裂时, 会导致不同颗粒的连通区域融合在一起, 造成球

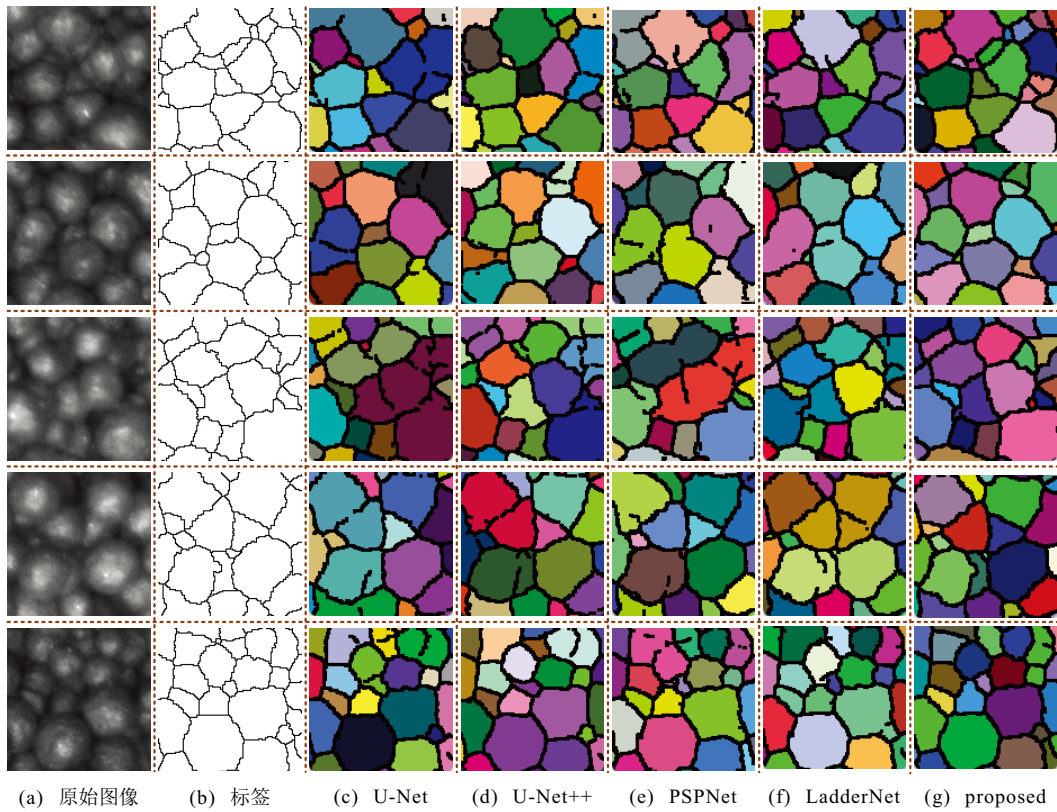


图8 不同算法模型分割效果可视化

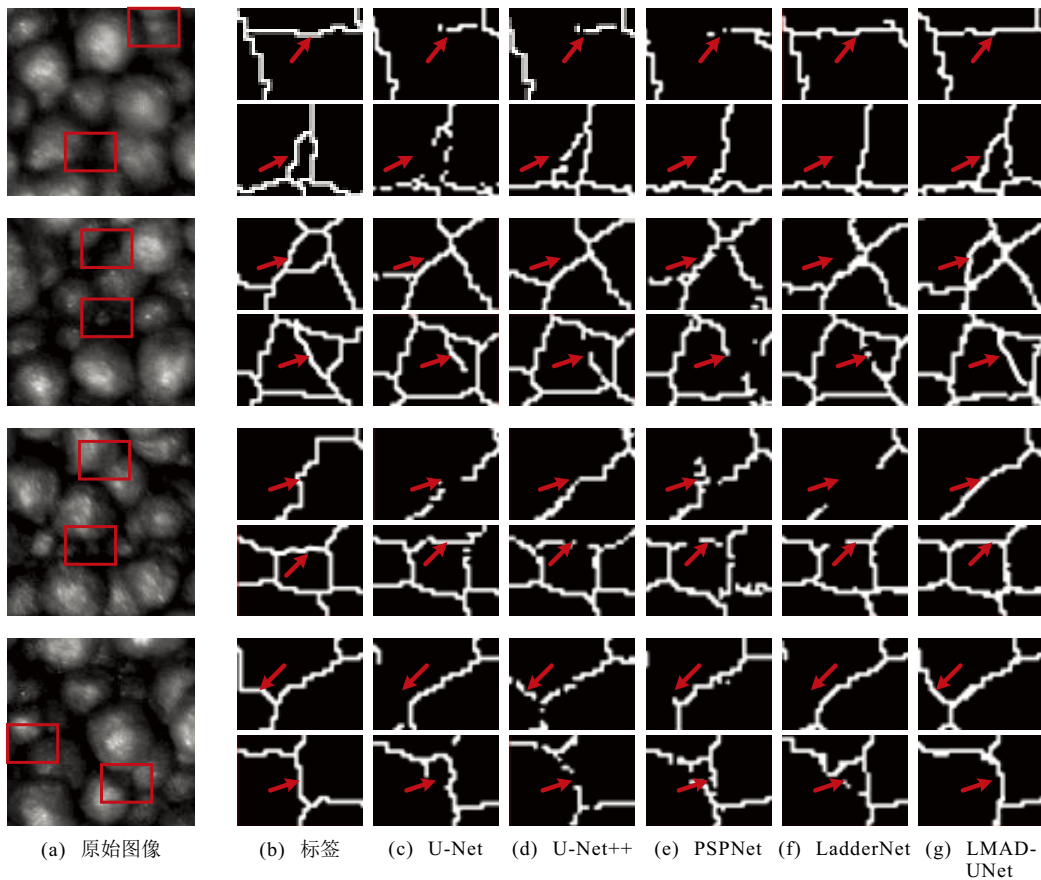


图9 不同模型分割效果细节对比

团颗粒分割错误. 相较于其他对比算法 (图 7(c) ~ 图 7(f)), 所提出算法 (图 7(g)) 对球团颗粒边缘检测较为精准, 欠分割现象较少, 因此, 被错误划分的区

域也较少, 能够提高后续颗粒粒度检测的准确率.

为了更清楚地突显所提出 LMAD-UNet 在细微特征上良好的分割效果, 对不同算法球团分割细节

进行详细对比. 图 9 为不同模型分割效果细节对比, 展示了 LMAD-UNet 与其对比分割算法的分割细节图像. 图 9(a) 为原始图像, 图 9(b) 为图像标签真实值, 图 9(c) ~ 图 9(g) 分别为 U-Net、U-Net++、PSPNet、LadderNet 和所提出算法的分割结果图像. 由图 9 可见: 在轮廓边缘细节处, 其他模型在分割这些地方时均明显出现了不同程度的断裂和缺失现象; 相比之下, 所提出算法可以精确地检测到上述部分, 这主要是由于轻量多尺度特征提取模块能够获得更多轮廓的细节信息, 双重注意力机制能够更进一步地获取边缘纹理细节信息, 抑制了阴影以及噪声的干扰.

所提出算法分割出的边缘轮廓比较清晰、平滑合理, 阴影和模糊边缘处细节信息保存完好, 表明并联 UNet 结构能够有效减少纹理信息在特征传递过程中的损失. 结果表明, 通过用并联 UNet 结构并引入轻量多尺度融合模块和双重注意力模块, 能够有效分割球团轮廓, 获得比其他对比算法更好的分割结果.

3.3 性能分析

在表 2 不同模型的性能比较中显示了球团矿数据集在不同模型的精确率、召回率、 F_1 分数、推理时间等性能指标的比较结果. 对比结果表明, LMAD-UNet 算法与其他网络相比, 明显有更高的准确率, 可以精确地检测密集区域的球团边界. 由表 2 可见: LMAD-UNet 算法在召回率、 F_1 -score 性能指标上至少比其他算法高 10%, 整体性能明显优于其他算法; 但是, 在图像检测的速度上与原始 UNet 相比, 所提出方法相对慢一点, 这是因为使用了双 UNet 结构和通道注意力机制、位置注意力机制, 增加了算法处理时间, 但是对比常用的几个分割算法, 更加快速, 因为使用了比传统卷积参数量更小的 LMultires 模块进行替代, 牺牲部分运行时间, 以提高模型的准确性, 可以更好地适应密集区域和不同工况环境下的分割任务.

表2 不同模型的性能比较

模型	球团矿数据集				
	Precision/%	Recall/%	F_1 -score/%	AUC/%	t /ms
U-Net	83.45	81.72	77.02	97.59	12.79
U-Net++	86.33	82.10	80.54	98.21	30.11
PSPNet	78.12	76.82	83.41	96.57	25.38
LadderNet	90.61	90.32	92.02	98.26	17.61
LMAD-UNet	96.77	92.67	96.85	99.37	15.65

图 10 和图 11 分别为模型训练时的损失和 P-R 曲线. 由图 10 和图 11 可见: 所提出方法的损失下降得最平滑, Loss 也更低; 而 UNet 的 Loss 虽然稍微低

一点, 但是波动大, 不平滑, 表明所提出方法的精度更高, 模型更稳定.

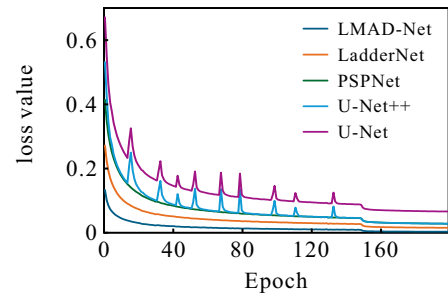


图10 训练损失曲线

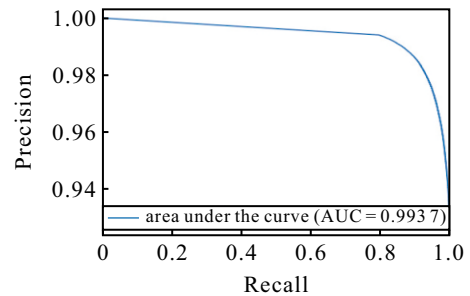


图11 P-R 曲线

P-R 曲线下的面积 (AUC-PR) 作为评估模型性能的重要指标, 其数值越大, 模型在精确率 (Precision) 与召回率 (Recall) 的权衡中表现越优异. 如图 11 所示, 所提出模型的 AUC-PR 值已达到 0.9937, 逼近理论最大值 1 (0.983 ± 0.008 , 置信区 95%), 这一显著结果印证了模型在复杂场景中具备极强的正类识别能力和误判抑制能力. 实验发现, 所提出模型在低召回阈值 (Recall < 0.2) 时即可维持 0.98 以上的精准度, 当召回率提升至 0.9 时仍然能够保持 0.96 的精确水平, 展现出了优异的鲁棒性. 特别是在类别不平衡 (正负样本比大于 1 : 50) 的极端测试集中, 模型仍然维持 0.9937 的 AUC-PR 值, 验证了其强大的泛化能力. 实验结果充分表明, 所提出多尺度特征融合架构和动态阈值调整机制有效提升了模型的综合性能.

3.4 消融研究

为进一步表明所提出模型的有效性, 对其进行消融实验, 以分析不同模块的有效性. 由于所提出模型是以 UNet 为 Baseline 进行的改进, 本文消融实验选用 UNet 为基准, 在原始结构上进行逐步增添模块进行调整, 表 3 为消融实验分析. 其中: 加粗值为对应指标的最优值; M_1 (Baseline-UNet) 为 U-Net 基线; M_2 (+ Dual-UNet) 表示将两个 UNet 进行并联, 在它们之间对应的尺度层级加入跳链接; M_3 (+ LMultires) 表示基于 M_2 加入改进的多尺度融合模块; M_4 (+ DualAttention) 表示在 M_3 中加入双重注意力机制; M_5 (+ DepthWiseConv) 表示在 M_4 中引入深度可分

离卷积,即所提出 LMAD-UNet. 这里: UNet 各项性能标准均较低,表明该模型在提取球团边缘轮廓这一任务的能力不足,主要原因是原始 UNet 对局部和全局特征信息的提取能力较弱,以及该模型对噪声的抗干扰能力不足导致出现断裂和欠分割现象. M_2 为并联 UNet 结构,各项性能指标均有显著上升,但是模型参数量和推理时间均增加了,这是因为并联结构的 UNet 增加了多条信息流,减少了特征信息传递过程中的损失,从而提升了分割精度,由于有两个 UNet 模块,参数量会翻倍; M_3 相较于 M_2 的 Acc

和 F_1 分别提升了 2.08% 和 1.40%,参数量显著降低,表明轻量多尺度融合模块将高分辨率与低分辨率的信息融合,有效地提取了更多的特征信息,提高了边缘的准确度; M_4 在加入了注意力机制后,各项性能有显著地提升,表明该模块使得模型的主要关注目标在球团边缘区域,减少了噪声和阴影的干扰; M_5 在加入深度可分离卷积后形成了所提出的最终模型,实现了模型在没有损失过多精度的情况下,显著减少了参数量和推理时间. 实验结果表明,所提出 LMAD-UNet 具有合理性和有效性.

表3 消融实验分析

模型配置	Acc/%	Recall/%	F_1 -score/%	AUC/%	Param/M	t /ms
M_1 (Baseline-UNet)	88.17	81.72	77.02	97.59	534	22.79
M_2 (+ Dual-UNet)	90.61	90.32	86.44	98.36	1251	26.31
M_3 (+ LMulti-Res)	91.21	90.94	88.36	98.41	256	20.03
M_4 (+ DualAttention)	94.88	91.42	97.05	99.86	358	18.41
M_5 (+ DepthWiseConv)	93.94	92.67	96.85	99.37	126	15.65

为了更直观地展现所提出算法的高效性,本文提供了消融模型效率对比,如图 12 所示. 其中: 红色标记点显示所提出模型在进行分割任务时只需 16 ms,参数量也小于 200 M. 可以得出所提出算法的推理时间和模型参数量均处于具有一定的优势位置,体现了该模型的轻量性和高效性.

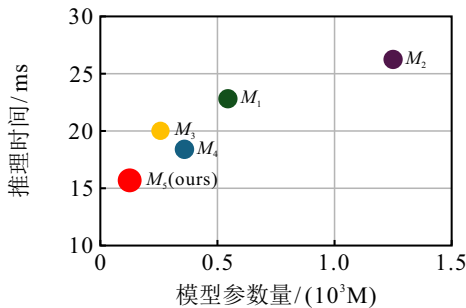


图12 消融模型效率对比

本文还进行了消融实验对两个损失函数的效果进行分析,用以验证加权融合损失函数的有效性,实验结果如表 4 所示.

表4 损失函数消融实验结果

损失函数	F_1 -score/%	Precision/%	参数量/M
BCE Only	77.30	87.91	126
Dice Only	82.11	89.45	126
BCE + Dice	96.85	96.77	126

实验结果表明,单一损失函数存在显著局限: BCE 损失虽然保持 87.91% 的高精确率,但是受类别不平衡影响导致严重漏检 (F_1 -score 仅 77.30%); Dice

损失虽然提升了召回性能 (F_1 -score 为 82.11%), 却因梯度不稳定引发边界模糊. 二者融合后产生协同效应, BCE + Dice 组合在参数量不变条件下实现了 F_1 -score 为 96.85%, 较最佳单损失提升了 14.74%.

3.5 粒径可视化

对密集球团颗粒轮廓分割图像进行圆拟合是测量球团粒度分布的最后阶段. 通过对密集球团颗粒图像的观察和分析发现,颗粒间重叠和粘连现象较为严重. 如图 13 所示: 对于上层颗粒,其轮廓完整;而被覆盖在底层的颗粒的轮廓分割结果呈现不规则的形态. 由于圆盘造球机的工作特性,球团矿颗粒不断翻滚最终形成类球形颗粒,在图像中呈圆形状. 根据以上生产工艺过程的先验知识,为了减小颗粒堆叠效应带来的粒度测量误差,本文在分割模型的基础上利用最小二乘方法对分割颗粒轮廓进行外接圆拟合,对遮挡球团颗粒轮廓缺失部分进行修复,实现粒径可视化,达到测量粒度分布的目的.

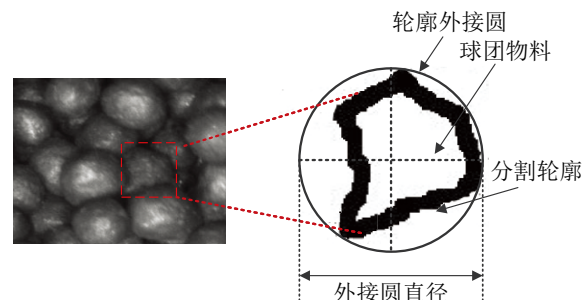


图13 外接圆测量方法

轮廓外接圆的测量方式如图 13 所示. 通过采用

轮廓外接圆对球团矿颗粒进行粒径统计,并计算每个级别的粒径个数占总个数的百分比,从而得到球团矿的粒度分布情况.图14为圆拟合和粒度分布示意图.获得轮廓拟合圆的图像后,能够方便快速地区分不同粒度颗粒分布情况,对球团矿物料的粒度做出定量统计.通过对同一时间段多幅图像粒度分布统计与人工筛分结果比较,如图14(d)所示,本文视觉测量系统得到的粒度分布与人工筛分结果较为吻

合.进一步统计8~16mm之间合格球团颗粒的占比,视觉检测结果与人工筛分相对误差为4.2%.在硬件环境为CPU为Xeon(R) Silver 4214R 12核,GPU为GeForce GTX 3080Ti,显存12GB条件下,所提出算法单幅图像测试时间为1.3s,能够满足工业现场实时检测的需求,能够对球团矿生产过程中参数的调整进行精准指导,对钢铁行业中生产效率的提升具有积极意义.

4 结论

为了实现对密集球团颗粒粒度分布的在线测量,本文提出了一种融合双重注意力与多尺度级联网络的密集颗粒测量方法.该方法采用级联网络的方式来实现对图像特征的提取,增加了深度网络的宽度,减少了深度网络下采样以及池化过程中导致的微弱轮廓特征丢失现象.并提出了一种轻量多尺度特征融合模块,实现了对密集颗粒多尺度特征的提取,减少了深度网络的参数量,提高了计算效率.同时,引入了双重注意力机制模块,利用通道注意力和位置注意力聚合颗粒边缘特征空间信息,加强了深度网络模型对轮廓细节特征的提取能力,抑制了背景噪声,改善了边缘细节的可视化效果.另外,本文还改进了损失函数,使得深度网络着重于对密集球团轮廓区域特征的学习.分割实验结果表明,所提出方法能够对密集球团颗粒进行精准分割, F_1 -score达到了96.85%,整体优于其他对比方法.粒度测量可视化结果表明,所提出方法能够有效对密集球团颗粒粒度分布进行测量,且时间效率可以满足工业生产现场在线测量的要求.

参考文献 (References)

- [1] 陈许玲, 蒋文浩, 黄晓贤, 等. 基于特征选择与梯度提升决策树的球团矿抗压强度预测[J]. 钢铁研究学报, 2024, 36(3): 319-327.
(Chen X L, Jiang W H, Huang X X, et al. Prediction of compressive strength of pellets based on feature selection and gradient boosting decision tree[J]. Journal of Iron and Steel Research, 2024, 36(3): 319-327.)
- [2] 顾宝澍, 马叙叙, 朱子康, 等. 不同比表面积铁精矿球团氧化动力学[J]. 过程工程学报, 2023, 23(5): 755-762.
(Gu B S, Ma X X, Zhu Z K, et al. Oxidation kinetics of pellets prepared from iron concentrate with different prepared from iron concentrate with different specific surface areas[J]. The Chinese Journal of Process Engineering, 2023, 23(5): 755-762.)
- [3] 王新东, 金永龙. “双碳”背景下高炉使用高比例球团的展望[J]. 过程工程学报, 2022, 22(10): 1379-1389.
(Wang X D, Jin Y L. Prospect on high ratio pellet utilized in blast furnace under the background of carbon peaking and carbon neutrality[J]. The Chinese Journal of Process Engineering, 2022, 22(10): 1379-1389.)
- [4] Wu X, Liu X Y, Duan J X. Online size distribution

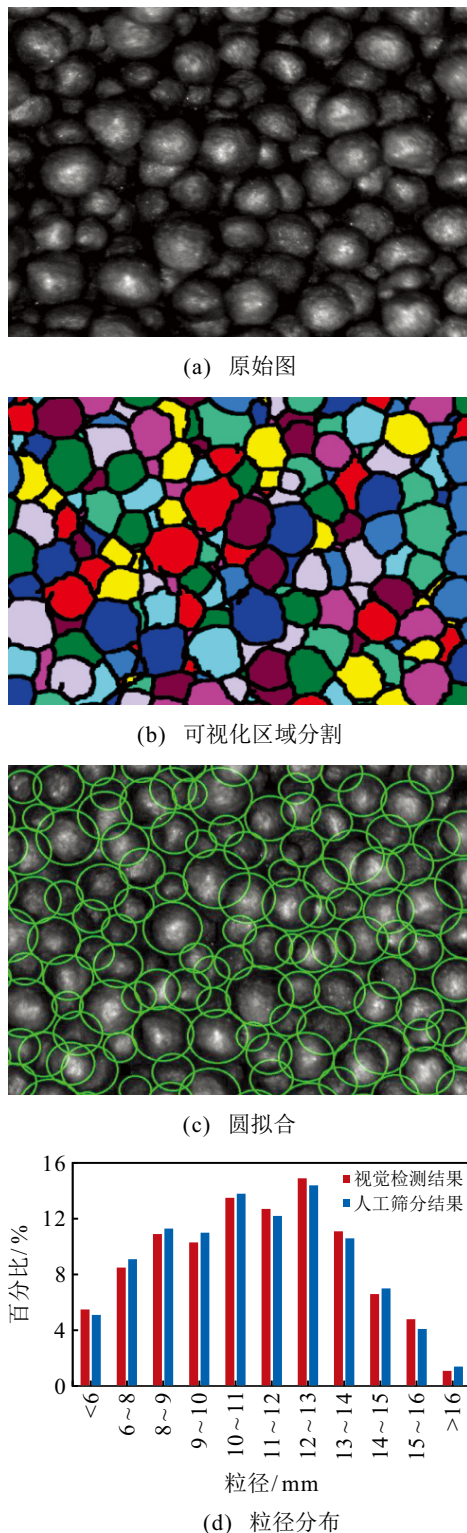


图14 圆拟合和粒径分布示意图

- measurement of dense iron green pellets using an efficient and multi scalenested U-net method[J]. *Powder Technology: An International Journal on the Science and Technology of Wet and Dry Particulate Systems*, 2021, 387(1): 584-600.
- [5] 林双, 陆伟文, 胡守景, 等. 结合 Mask-RCNN 和最小二乘法的球团粒度识别模型[J]. *钢铁研究学报*, 2024, 36(7): 855-865.
(Lin S, Lu W W, Hu S J, et al. A pellet size recognition model combining Mask-RCNN and least squares method[J]. *Journal of Iron and Steel Research*, 2024, 36(7): 855-865.)
- [6] 刘旭琛, 刘辉, 陈甫刚, 等. 一种 DDMCN 火焰图像特征提取的转炉炼钢碳含量实时预测方法[J]. *控制与决策*, 2023, 38(10): 2795-2804.
(Liu X C, Liu H, Chen F G, et al. A real-time prediction method of carbon content in converter steelmaking based on DDMCN flame image feature extraction[J]. *Control and Decision*, 2023, 38(10): 2795-2804.)
- [7] 安迪, 王姝, 关展旭, 等. 基于跨域流形正则化特征域适应的浮选工况识别[J]. *控制与决策*, 2023, 38(9): 2597-2605.
(An D, Wang S, Guan Z X, et al. Identification of flotation working condition based on feature domain adaptation of cross-domain manifold regularization[J]. *Control and Decision*, 2023, 38(9): 2597-2605.)
- [8] 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.
- [9] Heydari 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.
- [10] Nanda A, Patra A K, Panigrahi S. Pellet sphericity and size analysis by image processing technique[C]. *International Conference on Applied Machine Learning*. Bhubaneswar, 2019: 222-227.
- [11] 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 X Y, Wu X, Sun W, et al. Image segmentation of pellet particles based on morphological reconstruction and GMM[J]. *Chinese Journal of Scientific Instrument*, 2019, 40(3): 230-238.)
- [12] 白志城, 魏子赫, 谢文卓, 等. 基于图像处理的球团矿粒度检测系统及分割算法研究[J]. *烧结球团*, 2021, 46(5): 41-46.
(Bai Z C, Wei Z H, Xie W Z, et al. Research on pellet particle size measurement system and segmentation algorithm based on image processing[J]. *Sintering and Pelletizing*, 2021, 46(5): 41-46.)
- [13] Duan J X, Liu X Y, Wu X, et al. Detection and segmentation of iron ore green pellets in images using lightweight U-Net deep learning network[J]. *Neural Computing and Applications*, 2020, 32(10): 5775-5790.
- [14] Duan J X, Liu X Y. Online monitoring of green pellet size distribution in haze-degraded images based on VGG16-LU-net and haze judgment[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 5006316.
- [15] 张学锋, 史桢, 周思雨, 等. 基于改进 UNet 网络的烧结矿粘连图像分割研究[J]. *烧结球团*, 2023, 48(6): 83-89.
(Zhang X F, Shi Z, Zhou S Y, et al. Research on image segmentation of sinter adhesion based on improved UNet network[J]. *Sintering and Pelletizing*, 2023, 48(6): 83-89.)
- [16] Deo A J, Behera S K, Das D P. Online monitoring of iron ore pellet size distribution using lightweight convolutional neural network[J]. *IEEE Transactions on Automation Science and Engineering*, 2024, 21(2): 1974-1985.
- [17] 李福民, 侯炬才, 刘颂, 等. 基于计算机视觉和 Re-UNet 网络的烧结混合料粒度识别模型[J]. *钢铁研究学报*, 2023, 35(11): 1347-1357.
(Li F M, Hou J C, Liu S, et al. Particle size identification model of sintering mixture based on computer vision and re-UNet network[J]. *Journal of Iron and Steel Research*, 2023, 35(11): 1347-1357.)
- [18] Deo A J, Sahoo A, Behera S K, et al. Vision-based size classification of iron ore pellets using ensembled convolutional neural network[J]. *Neural Computing and Applications*, 2022, 34(21): 18629-18641.
- [19] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, 2016: 2818-2826.
- [20] 刘嘉敏, 谢文杰, 黄鸿, 等. 基于空间和通道注意力机制的目标跟踪方法[J]. *电子与信息学报*, 2021, 43(9): 2569-2576.
(Liu J M, Xie W J, Huang H, et al. Spatial and channel attention mechanism method for object tracking[J]. *Journal of Electronics & Information Technology*, 2021, 43(9): 2569-2576.)
- [21] Huang Z L, Wang X G, Wei Y C, et al. CCNet: Criss-cross attention for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(6): 6896-6908.
- [22] Milletari F, Navab N, Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]. *Proceedings of the 4th International Conference on 3D Vision*. Stanford, 2016: 565-571.

作者简介

吴鑫 (1987-), 男, 讲师, 博士, 主要研究方向为机器视觉检测、深度学习、语义分割和场景解析, E-mail: xinwu@hnu.edu.cn;

李靛杰 (1999-), 男, 硕士生, 主要研究方向为机器视觉、图像语义分割, E-mail: 220720854027@stu.hutb.edu.cn;

魏建好 (1989-), 男, 讲师, 博士, 主要研究方向为图像安全、人工智能, E-mail: jianhao@hutb.edu.cn.