

# 基于 CMI-UNet 的煤泥浮选泡沫分割算法

盖文东<sup>1</sup>, 王宁<sup>1</sup>, 张婧<sup>1</sup>, 李琳<sup>2</sup>

(1. 山东科技大学 电气与自动化工程学院, 山东 青岛 266590;

2. 山东科技大学 能源与矿业工程学院, 山东 青岛 266590)

**摘要:** 浮选泡沫边缘分割与识别是选煤厂核心工艺过程参数预测与优化的关键技术, 对于煤泥浮选智能化建设具有重要意义。目前, 基于 U 型结构网络(U-Net)的分割算法存在高层语义特征与底层信息的语义鸿沟和简单拼接带来的歧义等问题。本研究基于通道多尺度集成 U 型网络(CMI-UNet)提出一种新的煤泥浮选泡沫分割算法, 该算法在 U-Net 网络基础上, 增加了通道交叉融合 Transformer 模块、多分辨率融合模块和信息瓶颈模块, 从而有效弥合潜在的语义歧义, 提升特征表达能力、增强算法对复杂多尺度信息的捕捉和表达能力。在自建数据集上的实验结果表明, 所提 CMI-UNet 算法对煤泥浮选泡沫具有更好的分割效果。

**关键词:** 煤泥浮选; 泡沫图像; 深度学习; 语义分割; U 型结构网络

中图分类号: TD923

文献标志码: A

## Coal flotation froth segmentation algorithm based on CMI-UNet

GAI Wendong<sup>1</sup>, WANG Ning<sup>1</sup>, ZHANG Jing<sup>1</sup>, LI Lin<sup>2</sup>

(1. College of Electrical and Automation, Shandong University of Science and Technology, Qingdao 266590, China;

2. College of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

**Abstract:** Flotation froth edge segmentation is a key technology for the prediction and optimization of the core process parameters in coal preparation plants. It is significant to the intelligent construction of the coal flotation. So far, the segmentation algorithms based on U-Net still have the following problems: the semantic gap existing between the high-level semantic features and underlying information features, and semantic ambiguity due to simple splicing. This study proposed a new coal flotation froth segmentation algorithm based on CMI-UNet. By incorporating the channel cross fusion transformer (CCFT) module, multi-resolution fusion (MRF) module and information bottleneck (IB) module based on U-Net, this network effectively bridges the potential semantic ambiguities, enhances the ability of feature expression, and strengthens its capacity to capture and express complex multiscale information. The experimental results on self-built dataset show that the proposed CMI-UNet segmentation algorithm has a better segmentation effect on coal flotation froth.

**Key words:** coal flotation; froth image; deep learning; semantic segmentation; U-Net

泡沫浮选是从矿物中分离出有价值矿物的重要工业工艺, 能有效提高矿物资源利用效率。实践经验表明, 浮选泡沫表面信息反应了矿物分离过程的运行状态, 因此准确分割泡沫边缘和识别泡沫表面信息具有重要意义。但由于浮选过程的复杂性<sup>[1]</sup>, 仅依靠人工对泡沫进行肉眼观察是不够的, 需要借助图像处理等技术自动识别泡沫表面信息从而提高生产效率<sup>[2]</sup>。泡沫图像由大量边缘较弱的黏性气泡组成, 且复杂场景环境下气泡大小分布极不均匀, 因此准确地分割出图像中的泡沫是一项具有挑战性的工作。

收稿日期: 2024-11-05

基金项目: 山东省自然科学基金项目(ZR2023MF112)

作者简介: 盖文东(1982—), 男, 山东淄博人, 教授, 博士, 主要从事飞行控制、容错控制、智能控制理论及应用研究。

张婧(1982—), 女, 山东聊城人, 副教授, 博士, 主要从事模式识别与智能系统、控制理论及应用研究, 本文通信作者。E-mail: zjing133@sdust.edu.cn

浮选泡沫边缘分割方法,分为经典图像分割算法和基于深度学习的图像分割算法。经典图像分割算法包括阈值分割和分水岭分割等。文献[3]提出基于斜率差分布的阈值算法对 Gabor 滤波后的细胞图像进行分割。文献[4]将泡沫的标记区域、气泡光反射特征和子图像类别信息进行数据融合得到泡沫分割结果。文献[5]首先提取图像特征作为泡沫前景标签,边界先验条件作为约束,然后融合泡沫前景标签和约束得到泡沫最优分割线。文献[6]首先对浮选泡沫图像进行分类和阈值分割,然后将高亮标记作为局部最小点,最后采用分水岭算法分割。但经典分割算法难以适应浮选泡沫形态多样、颜色变化复杂以及背景干扰等复杂场景,在实际应用中存在局限性。

深度学习图像分割算法则展现出了较大优势,通过模型训练学习图像的特征,并进行像素级别的分类<sup>[7]</sup>。在深度学习图像分割领域中,U型结构网络(U-Net)<sup>[8]</sup>凭借卓越的分割性能成为主流分割网络之一。U-Net由获取高层语义信息的收缩路径和融合浅层位置信息的扩张路径组成。文献[9]基于U-Net设计了一种气泡分割方法,并通过合成图像和实验图像验证了模型的有效性;文献[10]在U-Net中融入了卷积神经网络,实现两相流图像的气泡分割。为解决分割任务中遇到的语义歧义问题,注意力机制被引入到U-Net中。文献[11]提出ResUNet++,通过整合注意力机制、残差块提高分割精度;文献[12]通过在U-Net中加入注意力机制,有效提升了细胞分割的精度;文献[13]从空间注意力和通道注意力的角度改进U-Net,减少了像素之间的语义歧义;文献[14]对自注意力门控单元进行改进,提升了网络对泡沫边缘的精准分割能力;文献[15]在U-Net的基础上引入注意力模块,加强了模型对气泡前景区域的关注,同时降低了无气泡矿浆背景区域的权重。Transformer的引入进一步推动了分割任务的发展。文献[16]基于Transformer特征融合模块,通过融合多尺度特征和捕捉通道间依赖,有效解决了U-Net中跳跃连接的局限性;文献[17]将Transformer、自注意力机制、多尺度模块与分层上下文提取模块集成于U-Net的跳跃连接中;文献[18]提出融合自注意力与交叉注意力机制的关系Transformer模块以提升语义捕获能力;文献[19]结合Transformer和压缩-激励注意力机制(squeeze-and-excitation attention, SE-Attention)模块优化了U-Net分割性能;文献[20]在融合Transformer提取远程和局部特征的同时,通过自注意力机制减少特征差异。在多尺度特征融合方面,文献[21]提出的神经网络中的多分辨率卷积块,为分层特征学习和上下文理解提供了支持。此外,信息瓶颈也被应用于分割网络,文献[22]将信息瓶颈理论应用于语义分割,可以过滤掉不相关信息,提升与前景相关的特征范式。

尽管上述工作使得图像分割效果得到显著提升,但U-Net在如何有效捕获煤泥浮选泡沫图像的非局部语义信息,以及如何弥合编码器和解码器间语义歧义方面仍存在不足。另外,煤泥浮选泡沫具有黏连性大、尺度不一、图片前景与背景对比度低等特点。对此,针对U-Net用于煤泥浮选泡沫分割时无法准确分割泡沫边界和信息丢失等问题,提出基于通道多尺度集成U型网络(channel multi-scale integration U-Net, CMI-UNet)的煤泥浮选泡沫分割算法,主要贡献如下。

1) CMI-UNet通过通道交叉融合Transformer(channel cross fusion transformer, CCFT)模块和多分辨率融合(multi resolution fusion, MRF)模块,有效融合不同尺度下的特征信息,减小不同层次特征之间的语义歧义;通过信息瓶颈(information bottleneck, IB)模块,在压缩和再扩展特征图通道数时,保证特征信息的完整性和有效性,增强多尺度信息捕捉和表达能力。

2) 构建煤泥浮选泡沫图像数据集,与U-Net、ResUNet++、UCTransNet进行对比实验,结果表明,CMI-UNet具有更好的煤泥浮选泡沫分割效果;同时对比分割的泡沫面积表明,CMI-UNet分割泡沫面积更接近实际面积。最后,进行消融实验验证了所提算法的有效性。

## 1 基于CMI-UNet的分割算法

CMI-UNet以U-Net为基础,结构如图1所示,包括编码器(encoder)模块(图1蓝色块)和解码器(decoder)模块(图1浅粉色块)。编码器通过卷积层提取图像特征,并使用最大池化层缩小图像尺寸,逐步获取高层语义信息。解码器将高级特征还原为高分辨率图像,融合了浅层位置信息和深层语义信息,最终生成具有像素级语义信息的图像。针对解码器和编码器之间的语义歧义,为有效融合多尺度通道的功能与消除语义歧义,提出通道交叉融合Transformer模块,该模块从通道维度出发,融合具有交叉注意力

的多尺度上下文信息,可以有效整合多尺度通道特征,解决潜在的尺度语义歧义问题。同时,为进一步缓解编码器和解码器之间的语义歧义,在跳跃连接中添加多分辨率融合模块。此外,为有效地压缩和再扩展特征图的通道数,提出信息瓶颈模块,增强模型对多尺度信息的捕捉和表达能力。在 CMI-UNet 中,下采样将输入特征图的分辨率减半的同时提取更高层次的特征;信息瓶颈模块进一步提炼特征,增强关键信息的表示能力,同时减少冗余或无关特征的干扰;上采样将低分辨率特征图还原至高分辨率特征图的同时,还将高分辨率特征图与跳跃连接输出的特征图进行融合。整个过程实现了从特征提取到信息恢复的端到端图像语义分割。

图 1 中黄色箭头表示由多分辨率融合模块和上采样得到特征块进行卷积和 ReLU 函数激活,其中 ReLU 函数表示为:

$$f(t) = \max(0, t)。$$
 (1)

式中,  $t$  为神经元输入。

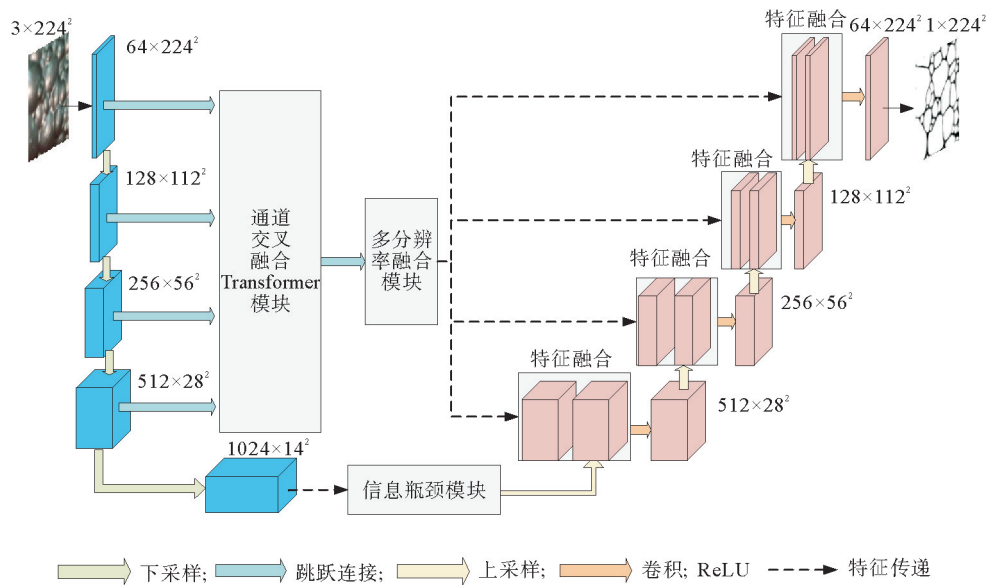


图 1 CMI-UNet 深度学习网络结构图

Fig. 1 Structure of CMI-UNet deep learning network

### 1.1 通道交叉融合 Transformer 模块

为解决浅层编码器和解码器之间存在的语义歧义、促进多尺度编码器的特征融合,提出通道交叉融合 Transformer 模块,该模块可以输出具有融合多尺度语义信息的特征图。通道交叉融合 Transformer 模块包括多尺度特征嵌入、多头通道交叉注意和多层感知机 (multilayer perceptron, MLP) 三个部分。通道交叉融合 Transformer 模块结构如图 2 所示。图 2 中圆圈加号表示特征的元素级加法操作, LN 表示层归一化 (layer normalization, LN), 计算式为:

$$L(x) = \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}}。$$
 (2)

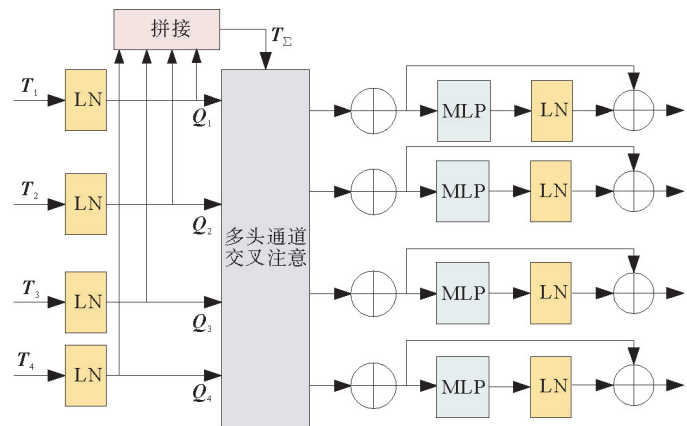


图 2 通道交叉融合 Transformer 模块结构图

Fig. 2 Structure of channel cross fusion Transformer module

式中: $x$  为特征图中的特征; $\mu$  和  $\sigma^2$  分别表示特征图的均值和方差; $\epsilon$  是小的正常数,用于避免分母取 0 而导致运算错误。

### 1.1.1 多尺度特征嵌入

为统一位置编码与补丁通道数的维度,首先将 4 个不同分辨率下的特征图  $T_i (i=1,2,3,4)$  划分为固定大小的补丁;然后在每个补丁中嵌入位置编码,生成带位置信息的标记补丁,标记补丁映射到编码器特征的四个尺度的相同区域时原始通道维度保持不变;最后将标记的  $T_i$  进行拼接得到  $T_\Sigma$ 。该流程可以整合多分辨率特征图的信息,为后续的多头通道交叉注意提供规范化输入。

### 1.1.2 多头通道交叉注意

多头通道交叉注意通过并行的多个注意力头来处理输入特征,从不同的表示空间中提取信息。多头通道交叉注意结构如图 3 所示。

图 3 中, $Q_j (j=1,2,3,4)$  表示从不同分辨率特征图提取的多尺度特征信息, $l$  为序列长度, $C_i (i=1,2,3,4)$  表示不同通道下的特征维度,其中  $C_1=128, C_2=256, C_3=512, C_4=1\ 024, C_\Sigma$  表示不同通道下特征组合后的总维度。多头通道交叉注意输入包括 4 个标记  $T_i$  和一个  $T_\Sigma$  :

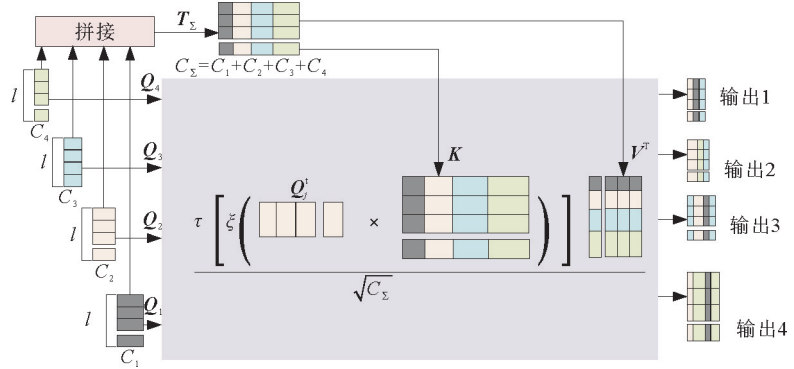


图 3 多头通道交叉注意结构图

Fig. 3 Structure of multi-head cross attention

$$Q_j = T_i W_{Q_j}, K = T_\Sigma W_K, V = T_\Sigma W_V. \quad (3)$$

式中: $K$  和  $V$  为从输入特征图中提取的特征信息,其中  $K$  表示每个位置的特征, $V$  表示每个位置的内容, $W_{Q_j}, W_K, W_V$  分别表示查询、键和值的输入权重。

计算查询向量与键向量的点积,并利用 Softmax 函数将点积结果转换为概率分布,得到注意力权重矩阵  $W_e$  :

$$W_e = \text{Softmax}\left(\frac{Q_j K^T}{\sqrt{d_k}}\right). \quad (4)$$

式中: $d_k$  为键向量维度,用于防止点积结果过大。Softmax 函数为:

$$\text{Softmax}(z_j) = \frac{e^{z_j}}{\sum_j e^{z_j}}. \quad (5)$$

式中: $\text{Softmax}(z_j)$  表示第  $j$  类的 Softmax 输出,即第  $j$  类的概率, $z_j$  表示输入的第  $j$  个元素。

对注意力权重矩阵  $W_e$  进行加权求和,得到加权特征矩阵  $W_r$  :

$$W_r = W_e V^T = \tau \left[ \xi \left( \frac{Q_j^T K}{\sqrt{C_\Sigma}} \right) \right] V^T = \tau \left[ \xi \left( \frac{W_{Q_j}^T T_i^T T_\Sigma W_K}{\sqrt{C_\Sigma}} \right) \right] W_V^T T_\Sigma^T. \quad (6)$$

式中: $r$  表示通道的索引, $\tau()$  和  $\xi()$  分别表示 Softmax 函数和实例归一化。

实例归一化计算式为:

$$\hat{x}_{m,n,p} = \frac{x_{m,n,p} - \mu_{m,n}}{\sqrt{\sigma_{m,n}^2 + \epsilon}}. \quad (7)$$

式中: $\hat{x}_{m,n,p}$  为归一化后的特征, $x_{m,n,p}$  为输入特征图的值,下标  $m$  表示第  $m$  个样本, $n$  表示第  $n$  个通道, $p$  表示第  $p$  个样本的位置, $\mu_{m,n}, \sigma_{m,n}^2$  分别表示第  $m$  个样本第  $n$  个通道的均值和方差。

多头通道交叉注意后的输出为:

$$M_\Sigma = \frac{W_r^1 + W_r^2 + \dots + W_r^N}{N}. \quad (8)$$

式中,  $N$  为头的数量。

将  $Q_j$  和  $M_\Sigma$  输入到 MLP 后获得最终输出:

$$O_m = M_\Sigma + \zeta(Q_j + M_\Sigma). \quad (9)$$

式中,  $\zeta(\cdot)$  表示 MLP。

### 1.2 多分辨率融合模块

如图 4 所示,多分辨率融合模块通过提取不同分辨率特征图的全局和局部信息,有效整合语义与细节特征,增强特征表示能力。首先,低分辨率和高分辨率特征图分别通过全局平均池化提取通道全局信息,经过多层感知机生成通道权重信息。然后,低分辨率特征图的通道信息通过适应性加权调整后,经过上采样操作与高分辨率特征图对齐,并在空间和通道维度上逐元素相加完成初步融合,融合后

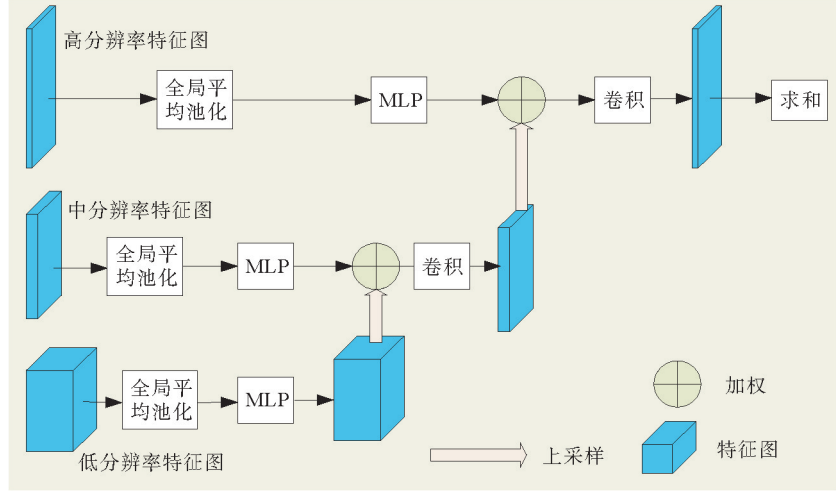


图 4 多分辨率融合模块结构图

Fig. 4 Structure of multi-resolution fusion module

通过卷积进一步提取特征图。融合过程从底层开始,逐层向上传递,逐步与中间层和高层特征图进行融合。最后,所有分辨率的特征图在相同分辨率下逐元素相加,输出多尺度信息的融合特征图。

### 1.3 信息瓶颈模块

信息瓶颈模块用于特征的提取和重构,即对通道数进行压缩后再进行扩展。

假设深度神经网络输入是  $X$ , 输出是  $Y$ , 中间层为随机编码  $Z$ , 其中  $Z$  是网络的瓶颈层(或称为中间表示层), 是输入数据  $X$  的压缩表示, 可以保留  $X$  和  $Y$  最相关的特征信息。其网络结构为:

$$X \rightarrow Z \rightarrow Y. \quad (10)$$

信息瓶颈目标是找到一个  $Z$ , 尽量减少  $X$  与  $Y$  不相关的信息, 保留  $X$  和  $Y$  最相关的那部分信息。

$$\max I(Z; Y) \quad \text{st.} \quad \min I(X; Z). \quad (11)$$

式中:  $I(X; Z)$  是输入  $X$  和中间  $Z$  的互信息, 反映了  $X$  和  $Z$  的信息量;  $I(Z; Y)$  是输出  $Y$  和中间  $Z$  的互信息, 反映了  $Y$  和  $Z$  的信息量。互信息定义为:

$$I(X; Z) = E_{p(x, z)} \left[ \ln \frac{p(x, z)}{p(x)p(z)} \right] = E_{p(x, z)} \left[ \ln \frac{p(z|x)}{p(z)} \right], \quad (12)$$

$$I(Z; Y) = E_{p(z, y)} \left[ \ln \frac{p(z, y)}{p(z)p(y)} \right] = E_{p(z, y)} \left[ \ln \frac{p(y|z)}{p(y)} \right]. \quad (13)$$

式中:  $p(x, z)$  为变量  $X$  和  $Z$  的联合概率分布,  $p(z|x)$  为  $Z$  基于  $X$  的条件概率分布,  $p(x)$  和  $p(z)$  分别为  $X$  和  $Z$  的边缘概率分布,  $p(z, y)$  为变量  $Z$  和  $Y$  的联合概率分布,  $p(y|z)$  为  $Y$  基于  $Z$  的条件概率分布,  $p(y)$  为  $Y$  的边缘概率分布;  $E_{p(x, z)}$  为  $p(x, z)$  的期望,  $E_{p(z, y)}$  为  $p(z, y)$  的期望,  $[\ ]$  内是  $p(x, z)$  和  $p(z, y)$  的具体说明。

引入拉格朗日乘子  $\beta$ , 将上述优化问题转化为拉格朗日优化问题, 即信息瓶颈的目标为优化目标函数  $L$ :

$$L = I(Z; Y) - \beta I(X; Z). \quad (14)$$

通过 Jensen 不等式构造互信息的下界和上界, Jensen 不等式表示为:

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i). \quad (15)$$

式中:  $\lambda_i > 0, i = 1, 2, \dots, n$ , 且  $\sum_{i=1}^n \lambda_i = 1$ ;  $x_i$  是随机变量的取值或样本点。

对于互信息  $I(Z;Y)$ ,引入变分分布  $q(y|z)$  近似真实的  $p(y|z)$ ,通过 Jensen 不等式得到互信息  $I(Z;Y)$  的下界:

$$I(Z;Y) \geq E_{p(z)} \left\{ E_{q(y|z)} \left[ \ln \frac{q(y|z)}{p(y)} \right] \right\}. \quad (16)$$

式中:  $E_{p(z)}$  表示  $p(z)$  分布期望,  $E_{q(y|z)}$  表示在条件分布  $q(y|z)$  下对  $y$  的期望。

对于互信息  $I(X;Z)$ ,引入变分分布  $q(z|x)$  近似真实的  $p(z|x)$ ,  $r(z)$  作为边缘分布  $p(z)$  的近似,通过 Jensen 不等式得到互信息  $I(X;Z)$  上界:

$$I(X;Z) \leq E_{p(x)} \left\{ E_{q(z|x)} \left[ \ln \frac{q(z|x)}{r(z)} \right] \right\}. \quad (17)$$

式中:  $E_{p(x)}$  表示  $p(x)$  分布期望,  $E_{q(z|x)}$  表示在条件分布  $q(z|x)$  下对  $z$  的期望。

结合以上推导得到目标函数  $L$ :

$$L = E_{p(z)} \left\{ E_{q(y|z)} \left[ \ln \frac{q(y|z)}{p(y)} \right] \right\} - \beta E_{p(x)} \left\{ E_{q(z|x)} \left[ \ln \frac{q(z|x)}{r(z)} \right] \right\}. \quad (18)$$

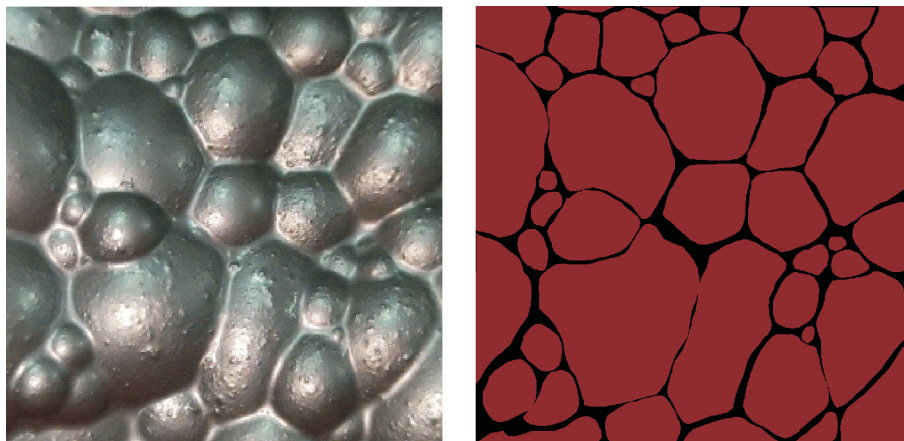
通过以上步骤得到目标函数  $L$ ,使其在捕捉对  $Y$  有用信息的同时,尽量减少与  $X$  的不相关信息。

## 2 实验结果与分析

### 2.1 数据集

在煤泥浮选实验过程中,通过摄像设备拍摄浮选过程的视频,并从视频中定时截取原始尺寸为  $1024 \times 1024$  像素的图像,构建了包含 239 张图片的数据集,并按训练集 80%、验证集 8%、测试集 12% 的比例划分为不同的子集。

为了确保所提算法能够准确识别和分割泡沫区域,对图像中的每个像素进行了相应的物体类别标注。在本实验中,标注的区域即为泡沫边缘。使用在线 Javascript 图像标注工具 LabelMe,对 239 张图像进行手动标注和分割。具体过程为,将图像中的泡沫边缘标注“Froth edge”标签,以实现泡沫边缘的准确描述。每张图像中的泡沫边缘都经过细致的手工分割,提供了高质量的标注数据。泡沫图像及标签图像如图 5 所示。



(a) 泡沫图像

(b) 标签图像

图 5 采集的泡沫图像及经手工分割后的标签图像

Fig. 5 Acquired froth image and labeled image after manual segmentation

### 2.2 测试平台与参数设置

实验所用 GPU 为 NVIDIA GeForce RTX 2080 Ti, GPU 加速计算库为 Pytorch 1.11.1+cu113。

通过 Dice 系数(dice similarity coefficient)和交并比(intersection over union, IoU)评价指标对分割效果进行衡量,用变量  $D_{ice}$ 、 $I_{oU}$  表示,计算式分别为:

$$D_{\text{ice}} = \frac{2 \times T_P}{(T_P + F_N) + (T_P + F_P)}, \quad (19)$$

$$I_{\text{oU}} = \frac{T_P}{T_P + F_N + F_P}. \quad (20)$$

式中:  $T_P$  为真阳性, 表示预测为正样本且标签为正样本的像素数量;  $F_P$  为假阳性, 表示预测为正样本但标签为负样本的像素数量;  $F_N$  为假阴性, 表示预测为负样本但标签为正样本的像素数量。  $D_{\text{ice}}$ 、 $I_{\text{oU}}$  取值范围均为 0~1, 其中 0 表示完全不重叠, 1 表示完全重叠, 评价指标越接近 1, 分割效果越好。

### 2.3 训练与测试

使用 CMI-UNet 对数据集进行训练和验证, 训练时从输入图像中逐层提取多尺度语义和细节特征, 同时通过对验证集的分割性能进行评估, 及时调整训练策略以确保模型的鲁棒性和泛化能力, 训练完成后, 将模型应用于测试集, 得到二值化分割图像。为更清楚地显示泡沫的原始状态, 使用卷积操作来遍历二值化图像, 记录二值化图像中 0 和 1 的位置, 并在原图上用红色轮廓标注出这些分割结果。煤泥浮选泡沫的原图、二值化图像和分割结果如图 6 所示。

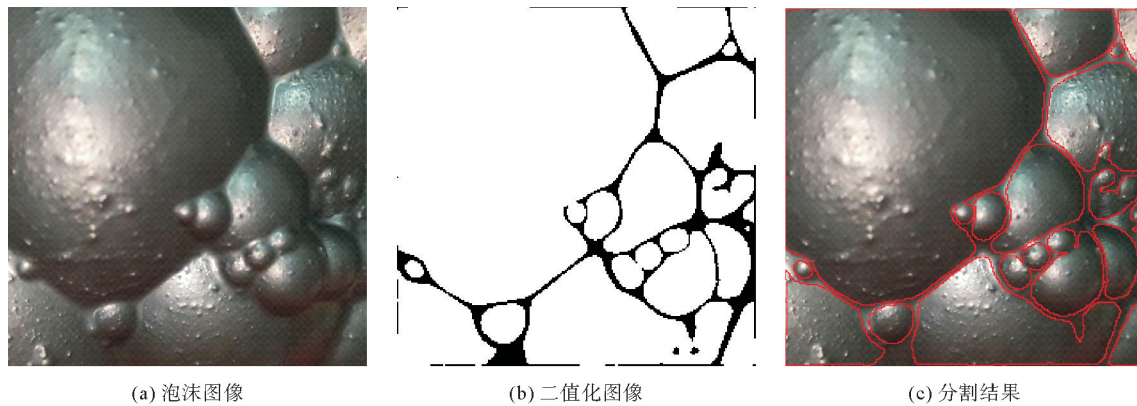


图 6 分割实验结果

Fig. 6 Results of the segmentation experiment

### 2.4 分割效果对比

#### 2.4.1 对比实验

为了验证 CMI-UNet 对煤泥图像分割的有效性, 选取 U-Net<sup>[8]</sup>、ResUNet++<sup>[11]</sup> 和 UCTransNet<sup>[16]</sup> 进行对比, 结果如表 1 所示。

可以看出, CMI-UNet 的 Dice 为 94.81%, 较

UCTransNet、ResUNet++ 和 U-Net 分别提高了 1.14%、2.71% 和 3.25%, CMI-UNet 的 IoU 为 90.17%, 较 UCTransNet、ResUNet++ 和 U-Net 分别提高了 2.00%、4.73% 和 5.59%。

图 7 给出了基于 U-Net、ResUNet++、UCTransNet 和 CMI-UNet 的煤泥浮选泡沫分割效果对比。由图 7 可看出, 基于 U-Net 分割算法得到的浮选泡沫分割图像存在边缘不清和欠分割等问题, 且大多数泡沫边缘未清楚识别; ResUNet++ 得到的分割结果更加清晰, 但仍存在欠分割问题; 相比于 U-Net 和 ResUNet++, UCTransNet 和 CMI-UNet 分割结果较好, 可以分割出图像中大多数泡沫边缘。

为对比 CMI-UNet 和 UCTransNet 的性能, 引入泡沫面积, 将 CMI-UNet 和 UCTransNet 分割泡沫面积与标签图像 Labels 泡沫面积进行对比, 结果如图 8 所示, 其中横坐标表示测试集图像索引, 纵坐标表示泡沫面积, 单位是像素 px。

表 1 分割结果对比

Table 1 Comparison of segmentation results %				
模型	U-Net	ResUNet++	UCTransNet	CMI-UNet
Dice	91.56	92.10	93.67	<b>94.81</b>
IoU	84.58	85.44	88.17	<b>90.17</b>

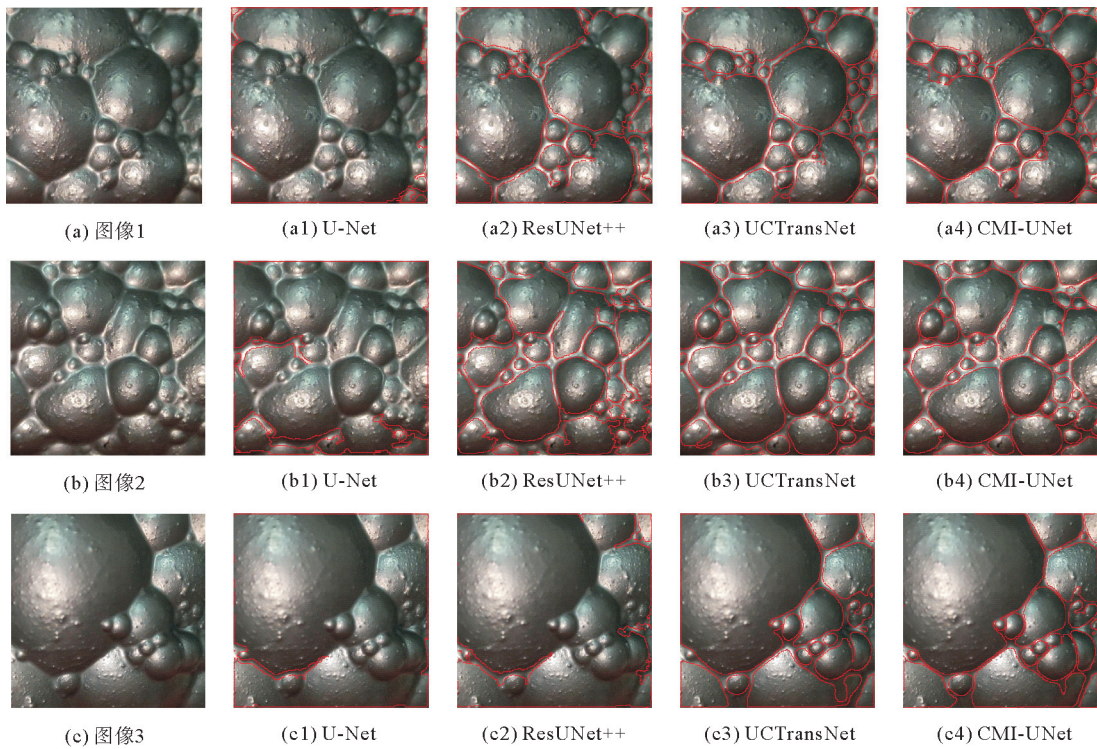


图 7 不同算法分割泡沫图像效果实验对比

Fig. 7 Experimental comparison of the effect of different algorithms for segmenting froth images

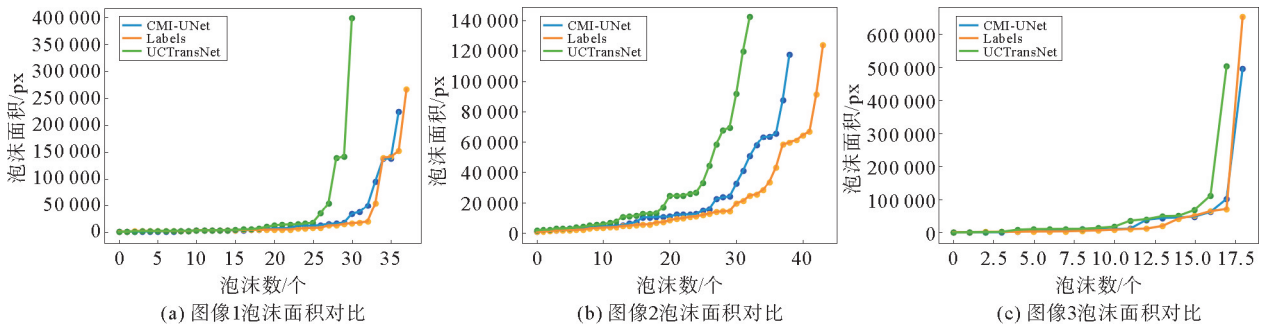


图 8 面积对比

Fig. 8 Area comparison

图 8(a)展示了图 7(a3)、图 7(a4)泡沫面积对比效果,图 8(b)展示了图 7 (b3)、图 7(b4)泡沫面积对比效果,图 8(c)展示了图 7 (c3)、图 7(c4)泡沫面积对比效果。从图 7 可知,CMI-UNet 和 UCTransNet 对浮选泡沫边缘的分割都取得较好效果,但由图 8 可知,CMI-UNet 得到的泡沫面积(蓝色曲线)相比于 UCTransNet 得到的泡沫面积(绿色曲线)更接近标签图像的泡沫面积(橙色曲线)。

由图 7 和图 8 可知,CMI-UNet 和 UCTransNet 对较小泡沫的分割效果都较好,且得到的泡沫面积比较接近;但对于较大泡沫,CMI-UNet 得到的泡沫面积更接近于真实面积。以图 8(a)中图像 1 为例,标签图像中最大泡沫面积为 265 163 px,CMI-UNet 计算最大泡沫面积为 224 036 px,UCTransNet 计算最大泡沫面积为 397 580 px。以标签图像面积为基准,CMI-UNet 泡沫面积误差为 15.5%,UCTransNet 泡沫面积误差为 49.9%。

#### 2.4.2 消融实验

为研究各个模块对算法分割效果的影响,设计消融实验,结果如表 2 所示。结果显示,在 U-Net 中加入 CCFT 后,Dice 和 IoU 分别为 94.16%和 89.03%,相比 U-Net 分割效果分别提升了 2.60%和 4.45%,

相比 CMI-UNet 分割效果降低 0.65% 和 1.14%; baseline 基础上加入 CCFT 和 MRF 模块, Dice 和 IoU 分别 94.33% 和 89.33%, 相比 U-Net 分割效果分别提升了 2.77% 和 4.75%, 相比 CMI-UNet 分割效果降低 0.48% 和 0.84%; baseline 基础上加入 CCFT 和 IB 模块, Dice 和 IoU 分别为 94.71% 和 89.01%, 相比于 U-Net 分割效果分别提升了 3.15% 和 4.43%, 相比 CMI-UNet 分割效果还差 0.10% 和 1.16%。

由表 2 结果可知, CMI-UNet 中的 CCFT 模块凭借其全局建模能力, 显著提升了分割性能, 有效捕捉了图像中的复杂上下文信息, 弥合了潜在的语义歧义; MRF 模块进一步融合了不同分辨率下的图像特征, 对高层语义信息、底层细节信息和全局结构处理至关重要; IB 模块通过压缩和选择性地传递重要信息来减少网络中的冗余, 从而提高了模型对于煤泥浮选泡沫图像复杂多尺度信息的捕捉与表达能力。

表 2 消融实验

Table 2 Ablation experiments		%
模块	Dice	IoU
U-Net(baseline)	91.56	84.58
baseline+CCFT	94.16	89.03
baseline+CCFT+MRF	94.33	89.33
baseline+CCFT+IB	94.71	89.01
CMI-UNet	94.81	90.17

### 3 结论

针对煤泥浮选泡沫边缘分割中高层抽象语义与底层细节信息融合不足的问题, 提出一种 CMI-UNet 的煤泥浮选泡沫分割算法, 通过引入通道交叉融合 Transformer 模块和多分辨率融合模块提取并融合多尺度上下文信息, 减少尺度语义歧义, 引入信息瓶颈模块来提高特征表达能力。在自建的煤泥浮选泡沫图像数据集上, 相较于 UCTransNet、ResUNet++ 和 U-Net, CMI-UNet 的 Dice 分别提高了 1.14%、2.71% 和 3.25%, IoU 分别提高了 2.00%、4.73% 和 5.59%, 并且 CMI-UNet 得到的泡沫面积值更接近实际面积。对 CMI-UNet 进行消融实验, 验证了 CMI-UNet 算法的有效性。

#### 参考文献:

- [1] ALDRICH C, AVELAR E, LIU X. Recent advances in flotation froth image analysis[J]. Minerals Engineering, 2022, 188:107823-107839.
- [2] JAHEDSARAVANI A, MASSINAEI M, ZARIE M. Measurement of bubble size and froth velocity using convolutional neural networks[J]. Minerals Engineering, 2023, 204:108400-108414.
- [3] WANG Z H, WANG Z Z. A generic approach for cell segmentation based on gabor filtering and area constrained ultimate erosion[J/OL]. Artificial Intelligence in Medicine, 2020, 107. DOI:10.1016/j.artmed.2020.101929.
- [4] ZHANG H, TANG Z H, XIE Y F, et al. A watershed segmentation algorithm based on an optimal marker for bubble size measurement[J]. Measurement, 2019, 138:182-193.
- [5] PENG C, LIU Y K, GUI W H, et al. Bubble image segmentation based on a novel watershed algorithm with an optimized mark and edge constraint[J/OL]. IEEE Transactions on Instrumentation and Measurement, 2021, 71. DOI:10.1109/TIM.2021.3129873.
- [6] LIANG X M, TIAN T, LIU W T, et al. Flotation froth image segmentation based on highlight correction and parameter adaptation[J]. Mining, Metallurgy and Exploration, 2020, 37(2):467-474.
- [7] MINAEI S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7):3523-3542.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer Assisted Intervention. 2015:234-241.
- [9] JU Y Y, WU L C, LI M, et al. A novel hybrid model for flow image segmentation and bubble pattern extraction[J/OL]. Measurement, 2022, 192. DOI:10.1016/j.measurement.2022.110861.
- [10] WANG S J, GUAN H F, WANG Y Q, et al. Intelligent recognition of gas liquid two phase flow based on optical image [J]. The International Arab Journal of Information Technology, 2023, 20(4):609-617.
- [11] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Resunet++: An advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on Multimedia. 2019:225-230.
- [12] 代茵, 刘维宾, 董昕阳, 等. 基于注意力机制的 U-Net 脑脊液细胞分割[J]. 东北大学学报(自然科学版), 2022, 43(7):

944-950.

DAI Yin, LIU Weibin, DONG Xinyang, et al. U-Net CSF cells segmentation based on attention mechanism[J]. Journal of Northeastern University (Natural Science), 2022, 43(7): 944-950.

- [13] 马思珂, 赵萌, 石凡, 等. 注意力驱动的细胞团簇细胞核分割算法[J]. 西安电子科技大学学报, 2022, 49(2): 198-206.  
MA Sike, ZHAO Meng, SHI Fan, et al. Attention driven nuclei segmentation method for cell clusters[J]. Journal of Xidian University, 2022, 49(2): 198-206.
- [14] 唐朝晖, 郭俊岑, 张虎, 等. 基于改进 I-Attention U-Net 的锌浮选泡沫图像分割算法[J]. 湖南大学学报(自然科学版), 2023, 50(2): 12-22.  
TANG Zhaohui, GUO Juncen, ZHANG Hu, et al. Zinc flotation froth image segmentation algorithm based on improved I-Attention U-Net[J]. Journal of Hunan University (Natural Science Edition), 2023, 50(2): 12-22.
- [15] 徐宏祥, 李神舟, 徐培培. 基于 DA-Attention U-Net 编码-解码结构的浮选矿浆相气泡图像分割[J]. 有色金属(选矿部分), 2024(6): 106-115.  
XU Hongxiang, LI Shenzhou, XU Peipei. Segmentation of flotation pulp phase bubbles image based on DA-Attention U-Net[J]. Nonferrous Metals (Mineral Processing Section), 2024(6): 106-115.
- [16] WANG H, CAO P, WANG J, et al. UCTransNet: Rethinking the skip connections in U-Net from a channel wise perspective with transformer[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(3): 2441-2449.
- [17] MA M J, XIA H Y, TAN Y M, et al. HT-Net: Hierarchical context attention transformer network for medical CT image segmentation[J]. Applied Intelligence, 2022, 52(9): 10692-10705.
- [18] HUANG S Q, LI J N, XIAO Y Z, et al. RTNet: Relation transformer network for diabetic retinopathy multilesion segmentation[J]. IEEE Transactions on Medical Imaging, 2022, 41(6): 1596-1607.
- [19] JIANG S, LI J L, HUA Z. DPCFN: Dual path cross fusion network for medical image segmentation[J/OL]. Engineering Applications of Artificial Intelligence, 2022, 116. DOI: 10. 1016/j. engappai. 2022. 105420.
- [20] WANG C, WANG L, WANG N Q, et al. CFATransUnet: Channel wise cross fusion attention and transformer for 2D medical image segmentation[J/OL]. Computers in Biology and Medicine, 2024, 168. DOI: 10. 1016/j. complbiomed. 2023. 107803.
- [21] DHEIVYA I, SARAVANA KUMAR G. Multi resolution neural network for road scene segmentation[C]//International Conference on Deep Learning, Artificial Intelligence and Robotics. Cham: Springer Nature Switzerland, 2023: 195-205.
- [22] HU Y T, HUANG X, LUO X Y, et al. Learning foreground information bottleneck for few shot semantic segmentation[J/OL]. Pattern Recognition, 2024, 146. DOI: 10. 1016/j. patcog. 2023. 109993.

(责任编辑: 齐敏华)