

文章编号 1004-924X(2025)24-3901-14

浮选泡沫表面红外成像显著目标分割及 稳定度估计

廖一鹏^{1*}, 廖玉华¹, 董则希¹, 白森杰¹, 何建军²

(1. 福州大学 物理与信息工程学院, 福建 福州 350108;

2. 福建金东矿业股份有限公司, 福建 三明 365101)

摘要:针对可见光图像易受光照影响而导致浮选泡沫稳定度视觉检测不准确的问题,对泡沫表面进行红外热成像,将不稳定气泡检测转换为红外显著目标分割,提出一种泡沫红外显著目标分割的 FIRSegNet 模型及稳定度估计方法。首先,基于 U-Net 架构进行 FIRSegNet 模型构建,采用 Swin Transformer 模块构建编码器网络捕捉长距离依赖关系,提升模型对图像全局上下文的理解能力;其次,在跳跃连接中嵌入双重交叉注意力模块(DCA),以更好地融合不同尺度红外泡沫图像的特征信息,增强模型分割不同大小红外显著性区域的能力;然后,提出一种残差细化模块(RRM),学习分割图与真值之间的残差以改善红外显著区域的边缘细节,实现显著目标的精细化分割;最后,根据显著目标分割结果计算泡沫稳定度。实验结果表明:本文方法在自建数据集能取得较好的分割效果,较主流分割模型的 mIoU 平均提升 1.73%, mDice 平均提高 0.63%, MAE 平均降低 11.6%, 分割精度得到一定提升;正常光照下不同工况泡沫稳定度预测的 RMSE, MAE, MAPE 和 AMAPE 平均值分别为 0.014 5, 0.0133, 2.384 8% 和 0.025 6, 在过曝光和低照度光照条件下的预测误差显著低于现有方法。本文方法具有较强的抗光照变化影响能力,保证了复杂光照环境下泡沫稳定度的准确定量描述,为后续的浮选生产工况与指标预测奠定基础。

关键词:浮选泡沫;红外显著目标分割;泡沫稳定度估计;Swin Transformer 网络;双重交叉注意力;残差细化网络

中图分类号: TP394.1; TH691.9 **文献标识码:** A

doi: 10.37188/OPE.20253324.3901

CSTR: 32169.14. OPE.20253324.3901

Infrared imaging-based salient object segmentation and stability estimation of flotation foam surfaces

LIAO Yipeng^{1*}, LIAO Yuhua¹, DONG Zexi¹, BAI Senjie¹, HE Jianjun²

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China;

2. Fujian Jindong Mining Co. Ltd., Sanming 365101, China)

* Corresponding author, E-mail: fzu_yp@163.com

Abstract: To address the inaccuracy in visual assessment of flotation foam stability caused by illumination variations in visible-light imaging, infrared thermal imaging was employed to capture the foam surface, thereby transforming the detection of unstable bubbles into a task of infrared salient object segmentation. Accordingly, FIRSegNet was proposed for salient object detection in infrared foam imagery, accompanied by a stability estimation method. FIRSegNet, an infrared salient object segmentation network, was con-

收稿日期: 2025-08-10; 修订日期: 2025-09-30.

基金项目: 国家重点研发计划项目(No. 2022YFB3603704); 国家自然科学基金资助项目(No. 62271149, No. 62271151)

structured based on the U-Net architecture. Swin Transformer module was employed in the encoder to capture long-range dependencies and to enhance the understanding of global contextual features. Dual Cross Attention (DCA) module was embedded in the skip connections to better fuse multi-scale features of infrared foam images, thereby improving the segmentation of salient regions of varying sizes. Furthermore, a Residual Refinement Module (RRM) was introduced to learn the residual between the predicted segmentation map and the ground truth, which refined boundary details and enabled fine-grained segmentation of salient objects. Finally, foam stability was estimated based on the segmentation results of the salient regions. Experimental results demonstrated that the proposed method achieved superior segmentation performance on a self-constructed dataset. Compared with mainstream segmentation models, it achieves average improvements of 1.73% in mIoU, 0.63% in mDice, and reduces MAE by 11.6%, indicating enhanced segmentation accuracy. Under normal lighting conditions, RMSE, MAE, MAPE, and AMAPE for foam stability prediction are 0.014 5, 0.013 3, 2.384 8%, and 0.025 6, respectively. Under overexposed and low-illumination conditions, the prediction errors are significantly lower than those of existing methods. The proposed method demonstrates strong robustness against illumination variations, ensuring accurate quantitative characterization of foam stability under complex lighting conditions. This provides a solid foundation for subsequent flotation process monitoring and performance prediction.

Key words: flotation foam; infrared salient object segmentation; foam stability estimation; Swin Transformer network; dual cross attention; residual refinement network

1 引言

泡沫浮选是目前矿物分选中应用最广的技术之一,也是矿物加工领域的重要核心手段,该技术利用矿物与杂质的亲疏水性质,使可浮性高的固体随气泡上浮至泡沫层^[1]。实验表明浮选过程中气泡的生成和破裂可有效反映矿物含量,通过图像分割技术,可以对浮选过程中新生成的气泡以及崩塌的气泡进行检测,这一技术在计算机视觉和工业生产等领域具有广泛应用前景^[2]。泡沫稳定度定义为泡沫中气泡抵抗合并和破裂的能力^[3],是浮选性能的关键驱动因素,有助于确定浮选过程的品位和回收率^[4]。因此,对浮选泡沫图像中合并及破裂气泡的准确分割能客观描述浮选气泡的稳定度,对浮选生产指标和工艺改善具有重要意义。

目前,国内外学者提出了多种不稳定区域的检测方法:Jahedsaravani等^[5]先对相邻图像帧进行位移矫正再求差,将差值图像中的高亮区域判为相应不稳定区域,但该方法的检测准确度会受到气泡变形和光照的影响,导致气泡内部亮斑或边缘高亮区域被误判为不稳定区域;Zhang等^[6]对相邻图像帧差值图像中高亮区域对应的原始

相邻图像区域计算相似度,并对相似度低的区域利用分水岭进行分割,此方法依赖于分水岭的分割结果,受光照影响大;魏等^[7]利用改进后的Mask2Former对不稳定区域分割,从而计算出稳定度,对于视觉特征不明显的区域容易发生漏检测;刘等^[8]对相邻两帧泡沫图像应用改进后的SuperGlue进行特征点检测及匹配,根据匹配结果估算稳定度,该方法检测精度在很大程度上依赖于匹配阈值的设置,存在一定的误差敏感性。同时,在光照改变的条件下,特征点匹配率发生较大变化,影响了稳定度估算的准确性;Liao等^[9]首次对泡沫进行红外热成像研究,发现气泡产生崩塌或合并后释放出热量,相比周围气泡的温度高,热成像后呈现出高亮的显著区域,对不稳定气泡具有直观的展现效果。这一方法有效避免了可见光图像受光照变化影响而导致浮选泡沫稳定度视觉检测不准确的问题。因此,本文通过对泡沫红外视频图像进行显著区域检测以实现合并及破裂气泡的分割,根据分割结果进行泡沫稳定度的评估。

近年来,对显著性目标分割方法的研究较多,尤其是基于Transformer架构的研究取得了显著进展。Chen等^[10]将Transformer的全局自注

注意力机制嵌入 U-Net 编码器构建 TransUNet 网络,以增强全局信息建模能力,实现显著目标的精确定位和分割,在一定程度上提升了分割精度,但也带来了较高的计算开销;Cao 等^[11]利用 Swin Transformer 结合 U-Net 结构的对称跳跃连接构建医学图像分割网络 Swin-Unet,该方法能够在降低计算复杂度的同时,充分保留细节信息,使分割结果更加精细,然而分割性能受噪声干扰的影响大;Xie 等^[12]提出了利用 Mix Transformer 主干提取多尺度特征的分割网络 Segformer,该方法效率高、易于扩展,可在不同分辨率的输入下保持鲁棒性,但是解码器设计较简单,在分割精细边界和处理小目标方面可能表现不足;Yuan 等^[13]提出了一种结合多层次、多模式、多阶段注意力机制的网络 M³Net,通过逐层特征融合机制,能够提取更丰富的显著性信息,从而实现复杂场景中的显著目标分割。尽管上述方法在显著性目标分割方面取得了一定成果,但仍存在边缘细节处理不足或对噪声鲁棒性差等问题,尚有进一步优化空间。

鉴于上述分析,本文将浮选泡沫稳定度评估转化为红外显著目标分割问题,构建基于 U-Net 架构的红外显著目标分割模型。采用 Swin Transformer 模块构建编码器网络,解码器通过双线性插值实现上采样,在跳跃连接中嵌入双重交叉注意力模块,然后设计一种残差细化模块学习分割图与真值之间的残差以改善红外显著区域的边缘细节,实现显著目标的精细化分割。最后根据显著目标分割结果计算泡沫稳定度,并自建数据集进行实验测试和结果比较分析。

2 泡沫表面红外热成像

针对可见光图像易受光照强弱影响导致浮选泡沫稳定度视觉检测不准确的问题,对浮选槽表面流动的泡沫进行红外热成像。如图 1 所示,将红外相机架设在浮选槽表面上方,以获取浮选表面的气泡图像,并将图像数据传输到计算机进行处理。采集的红外与可见光双模态图像如图 2 所示,气泡产生崩塌或合并后释放出热量,相比其他气泡的温度高,热成像后呈现出高亮的黄色区域。相较于可见光泡沫图像,经过红外热成像

后的泡沫红外图像能够较为直观展现合并、崩塌等不稳定泡沫。因此,可通过对红外泡沫图像进行显著目标分割统计进而实现泡沫稳定度估计。

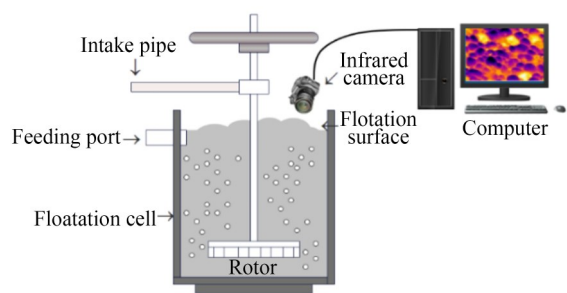


图 1 泡沫表面红外热成像系统

Fig. 1 Foam surface infrared thermal imaging system

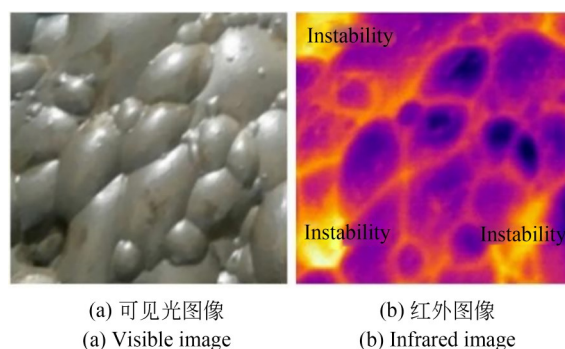


图 2 浮选泡沫双模态图像

Fig. 2 Flotation froth dual-modality image

3 红外显著目标分割网络模型

本文采用 U-Net 架构构建泡沫红外显著目标分割网络 (Foam Infrared Salient Object Segmentation Network, FIRSegNet), 总体架构如图 3 所示, 主要包含编码器、解码器和跳跃连接。采用 Swin Transformer 模块设计编码器, 为了融合不同尺度的跳跃连接特征, 引入双重交叉注意力模块 (Dual Cross Attention, DCA), 使模型在分割不同尺寸红外显著目标时具备更高的适应性和鲁棒性。解码器由三组 CBR (3×3Conv、Batch Norm 和 ReLU) 组成, 并通过双线性插值 (Bilinear Upsampling) 实现上采样。针对经过第一阶段检测的显著区域存在边界模糊问题, 设计了残差细化模块 (Residual Refinement Module, RRM) 学习真实图像与预测显著

区域的差值,从而对分割图像边缘进一步优化。

3.1 基于 Swin Transformer 的编码器网络设计

本文采用 Swin Transformer 模块构建编码器网络实现了长距离依赖关系的捕捉,从而提升了模型对图像全局上下文的理解能力,有效克服了经典卷积网络仅能处理局部信息的局限性。设计的编码器结构如图 3 所示,在初始阶段,输入图像首先通过 Patch Partition 层得到图像块,然后通过 Linear Embedding 层得到 tokens

(图像块的向量表示)。然后,将得到的 tokens 传入 Stage1 的 Swin Transformer block 中,获取深层语义特征信息。将 Stage1 得到的特征输入至 Stage2 的 Patch Merging 层来减少 tokens 的数量。在 Stage2 合并操作中,网络将每组 2×2 邻近图像块的特征拼接在一起,并通过一个线性层处理这些拼接后的特征,从而将四个图像块合并为一,将合并后的特征再传入 Swin Transformer block 中。Stage3 和 Stage4 将重复这一处理过程。

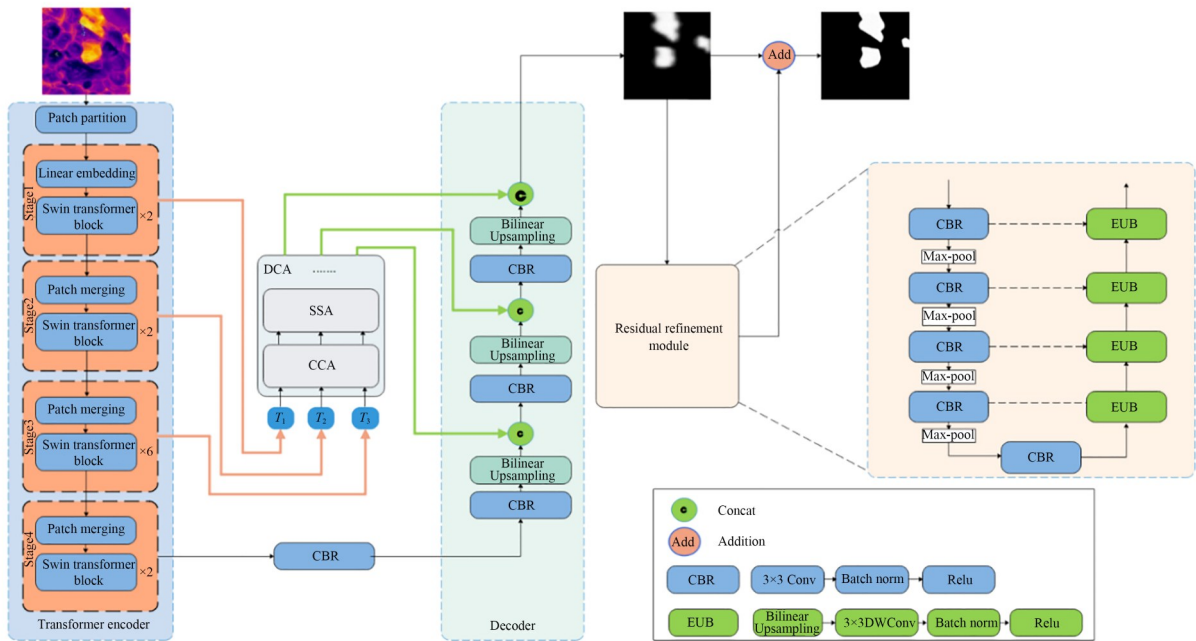


图3 分割网络整体框架

Fig.3 Segmentation framework

相比于 Transformer block, Swin Transformer block 用基于滑动窗口的多头自注意力机制模块 (SW-MSA) 替换标准的多头自注意力模块 (MSA), 其余的网络结构保持不变, 网络结构如图 4 所示。LN 表示 Layer Norm 层; SW-MSA 表示基于滑动窗口的多头自注意力模块; W-MSA 表示基于窗口的多头自注意力模块; MLP 表示多层感知机。Swin Transformer block 的计算过程可以通过以下公式来表示:

$$\hat{z}^l = W - MSA(\text{LN}(z^{l-1})) + z^{l-1}, \quad (1)$$

$$z^l = \text{MLP}(\text{LN}(\hat{z}^l)) + \hat{z}^l, \quad (2)$$

$$\hat{z}^{l+1} = \text{SW-MSA}(\text{LN}(z^l)) + z^l, \quad (3)$$

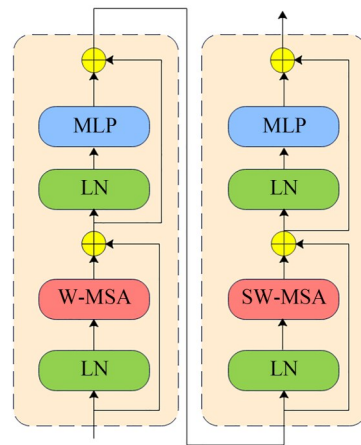


图4 Swin Transformer 基本结构

Fig.4 Basic structure of Swin Transformer block

$$\hat{z}^{l+1} = \text{MLP}\left(\text{LN}\left(\hat{z}^{l+1}\right)\right) + \hat{z}^{l+1}, \quad (4)$$

其中: $\hat{z}^l$ 表示经过 W-MSA 输出的特征信息; $\hat{z}^{l+1}$ 表示经过 SW-MSA 输出的特征信息; $\hat{z}^l$ 和 $\hat{z}^{l+1}$ 表示经过 MLP 模块输出的特征信息; $\hat{z}^{l-1}$ 表示经过上一个 Swin Transformer block 的 MLP 模块输出的特征信息。

3.2 双重交叉注意力模块(DCA)

U-Net 结构中简单跳跃连接不足以建模全局多尺度上下文,为了在编码器特征连接到其对应的解码器部分之前,有效地融合这些特征。本文采用双重交叉注意力模块(DCA)^[14],用于提取编码器中多尺度特征的通道维度和空间维度的相互依赖性,从而增强了模型分割不同大小红外显著性区域的能力。DCA 模块将编码器网络前三个阶段的输出作为输入,生成增强特征,并将它们连接到对应的三个解码器阶段。如图 5 所示,DCA 模块分为两个主要阶段:第一阶段包含多尺度补丁嵌入(Multi-Scale Patch Embedding),用于获取编码器多尺度特征的向量表示(tokens)。在第二阶段中,对这些 tokens 使用双重交叉注意力机制,通过通道交叉注意力(Channel Cross-Attention, CCA)模块和空间交叉注意力(Spatial Cross-Attention, SSA)模块来捕获不同尺度特征之间的关系;最后,依次对这些 tokens 使用 Layer Norm 和 GeLU 函数^[15]处理,并通过上采样操作将其与解码器中对应尺度的特征进行连接。

3.2.1 多尺度补丁嵌入

从 3 个多尺度编码器阶段(即跳跃连接)中提取 Patch。设第 i 个编码器阶段的特征为 $E_i \in \mathbb{R}^{C_i \times \frac{H}{2^{i+1}} \times \frac{W}{2^{i+1}}}$,其中 C_i 表示通道数, H 与 W 分别为输入的高与宽。对于每个阶段,设 Patch 大小为 $P_i^s = \frac{P^s}{2^{i+1}}, i=1,2,3$ 。使用二维平均池化操作从各尺度特征图中提取 Patch,池化核大小与步长均设为 P_i^s 。然后,在展平后的二维 Patch 上使用 1×1 深度卷积进行通道投影。其计算过程可表示为:

$$T_i = \text{DConv1D}_{E_i}(\text{Reshape}(\text{AvgPool2D}_{E_i}((E_i)))), \quad (5)$$

其中: $T_i \in \mathbb{R}^{P \times C_i}$ 表示第 i 个编码器阶段生成的展平 Patch, P 为 Patch 的数量。

3.2.2 双重交叉注意力机制

CCA 模块将来自上一阶段的 tokens $T_i, i=1,2,3$ 沿着通道维度拼接创建 T_c 作为 K 和 V ,使用 T_i 作为 Q_i ,使用 1×1 深度卷积对 tokens 进行变换:

$$Q_i = \text{DConv1D}_{Q_i}(T_i), \quad (6)$$

$$K = \text{DConv1D}_K(T_c), \quad (7)$$

$$V = \text{DConv1D}_V(T_c), \quad (8)$$

$$\text{CCA}(Q_i, K, V) = \text{Softmax}\left(\frac{Q_i^T K}{\sqrt{C_c}}\right) V^T, \quad (9)$$

其中: Q_i, K, V 分别代表查询、键和值矩阵, $\frac{1}{\sqrt{C_c}}$

是缩放因子, $\text{DConv1D}_x()$ 代表 1×1 深度卷积。SCA 模块将来自 CCA 模块的输出 $\bar{T}_i, i=1,2,3$ 沿通道维度拼接后的 tokens $\bar{T}_c$ 作为 Q 和 K ,而使用每个 tokens $\bar{T}_i$ 作为 V ,使用 1×1 深度卷积对 tokens 进行变换。

$$Q = \text{DConv1D}_Q(\bar{T}_c), \quad (10)$$

$$K = \text{DConv1D}_K(\bar{T}_c), \quad (11)$$

$$V_i = \text{DConv1D}_{V_i}(\bar{T}_i), \quad (12)$$

$$\text{SCA}(Q, K, V_i) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V_i, \quad (13)$$

其中: Q_i, K, V 分别代表查询、键、和值矩阵, $\frac{1}{\sqrt{d_k}}$

是缩放因子, $\text{DConv1D}_x()$ 代表 1×1 深度卷积。

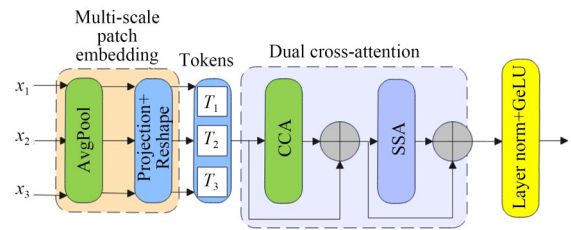


图 5 双重交叉注意力模块

Fig. 5 Dual cross attention

3.3 残差细化模块(RRM)

本文构建残差细化模块对显著区域边缘进一步优化。通过学习显著图 S_{coarse} 与 GT 间的残差 S_{residual} 进一步优化显著边缘得到精细化分割图:

$$S_{\text{refined}} = S_{\text{coarse}} + S_{\text{residual}}. \quad (14)$$

残差细化模块结构如图 3 所示,采用 U 型编码器-解码器的结构。编码器网络中每个阶段都含有一个 CBR 模块和 Max-pool 层,其具体步骤:先通过 64 个卷积核的 3×3 卷积层进行特征提取,然后将特征输入至 Batch Norm 层和一个 RE-LU 激活函数,最后利用 Maxpool 进行下采样。其公式表示为:

$$x_{l+1} = \text{Maxpool}(\text{ReLU}(\text{BN}(\text{Conv}(x_l)))) \quad (15)$$

解码器网络是由高效上采样卷积块(Efficient Upsampling Block, EUB)组成,其具体步骤:先通过双线性上采样函数放大特征图的尺寸,然后利用 3×3 深度卷积增强上采样后的特征,最后将特征先后输入至 Batch Norm 层和 RE-LU 激活函数。EUB 模块通过深度卷积代替标准卷积操作,有效降低了计算成本,同时保持了特征质量。其公式表示为:

$$x_{l+1} = \text{ReLU}(\text{BN}(\text{DWC}(\text{Up}(x_l)))) \quad (16)$$

4 泡沫稳定度估计

4.1 泡沫稳定度估计

本文利用红外显著目标分割结果估算泡沫稳定度 S ,将红外泡沫图像中的黄色高亮部分即破裂与合并气泡分割出来后,统计分割目标区域像素点的个数从而进行稳定度评估。 S 的计算公式定义如式所示(17)(其中 n 取 100)。该公式由两部分组成:前半部分表示单帧图像中泡沫的合并及破裂区域大小信息,后半部分为段内调整因子,用于考虑段内波动对结果的影响,使得最终的稳定度 S 不仅反映单帧瞬时状态,也反映泡沫在一段时间内的整体稳定性。当段内波动较大时,单帧图像指标会被相应压低;当段内波动较小时,单帧图像指标基本保持原值。

$$S = \left(\frac{X_i - x_i}{X_i} \right) \times \left(\frac{\bar{x} - \sigma}{\bar{x}} \right), \quad (17)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (18)$$

其中: X_i 表示每一帧浮选泡沫图像的像素数目; x_i 为红外浮选视频帧中分割的不稳定泡沫区域的像素点数目; $\bar{x}$ 为当前阶段泡沫序列分割区域的平均像素点数目; σ 为当前阶段泡沫序列分割区域像素点总数的标准差; $i = 1, 2, \dots, n$ 。

4.2 稳定度估计流程及步骤

本文所提出的泡沫稳定度评估流程如图 6 所示,其具体流程步骤如下:

Step1:采集铅浮选红外泡沫图像,构建包含多种工况的数据集,并将其划分为训练集、验证集、测试集。

Step2:构建泡沫红外成像显著目标分割模型 FIRSegNet,并使用训练集进行模型训练,通过验证集监控训练过程中的性能表现,从而获得最优参数。

Step3:将训练完成的 FIRSegNet 模型应用于测试集图像中,进行红外显著目标分割任务。

Step4:基于模型输出的分割图像,利用式(17)计算稳定度 S 。

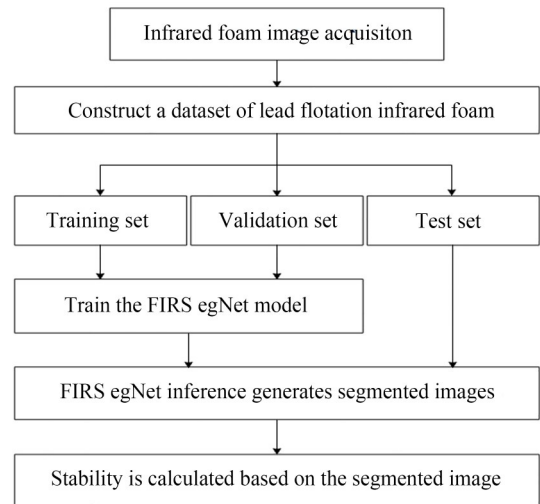


图 6 泡沫稳定度估计流程

Fig. 6 Foam stability estimation process

5 实验结果与讨论

5.1 实验环境配置与数据集

本文实验环境:软件平台为 Windows 10, Matlab 2018, Python3.8 和 Pytorch1.8.0。硬件平台为 Intel (R) Core (TM) i7-9700 CPU@3.00 GH, NVIDIA RTX3090 GPU, 16.00 GB RAM。在训练阶段,模型使用 AdamW 优化器进行优化,学习率为 2×10^{-4} 。批量大小设置为 16, 训练 200 个 epoch。

铅浮选红外泡沫数据集:该数据集在福建金东矿业股份有限公司的铅矿浮选厂,使用菲力尔

T620 红外热像仪采集,图像分辨率为 256×256 pixels,详细记录了铅浮选回路中粗选、扫选和精选状态下的红外泡沫视频(各占三分之一)。其中 60% 的 3 978 张图像用作训练数据,20% 的 1 326 张图像作为验证数据,剩余 20% 的 1 326 张图像作为测试数据。标注由三名浮选领域专家使用 Labelme 工具完成,主要标注显著泡沫区域的边界和连通特征。为评估标注一致性,随机抽取 10% 的样本由另一名专家进行复核,并采用 Cohen's κ 系数进行一致性分析^[16]。一致性达到 93.6% (Cohen's $\kappa=0.89$)。

铅浮选双模态泡沫数据集:该数据集包含正常浮选、过浮选和欠浮选状态下 10 分钟运行的泡沫双模态图像,图像分辨率为 256×256 pixels。为确保可见光图像中不稳定泡沫区域的标注准确性,在人工标注过程中通过来回切换前后两帧图像进行比对。红外图像仅用于后续实验输入,无需标注。该数据集用于验证基于红外图像的泡沫稳定度估计的准确性

为了全面且准确地量化评估本文算法,在常

用的分割指标中选取包括平均交并比(mIoU)、平均 Dice 系数(mDice)以及平均绝对误差(MAE)对实验结果进行评价。

5.2 泡沫红外显著目标分割效果

为验证本文红外显著目标分割模型在不同工况下、不同大小类型的气泡具有普遍性,利用过浮选、正常浮选、欠浮选三组泡沫红外视频序列对提出的分割方法(Ours)进行实验并与其他 5 种算法进行定性和定量对比,以此来验证本文提出算法的有效性,如图 7,图 8 和图 9 所示,为本文方法与其他比较方法的定性实验结果,其中 Ground Truth(GT)为人工标注的真实值。从图中可以看出,Segformer 和 effisegnet 在过浮选图像中的分割结果不够完整,会把泡沫内部区域像素错误识别为背景;TransUNet 容易将边缘黄亮区域识别为合并或破裂区域而造成过分割。同时,过浮选图像分割结果不够完整,会把泡沫内部区域像素错误识别为背景。本文方法分割效果最接近 GT,边缘平滑且分割得较为细致。

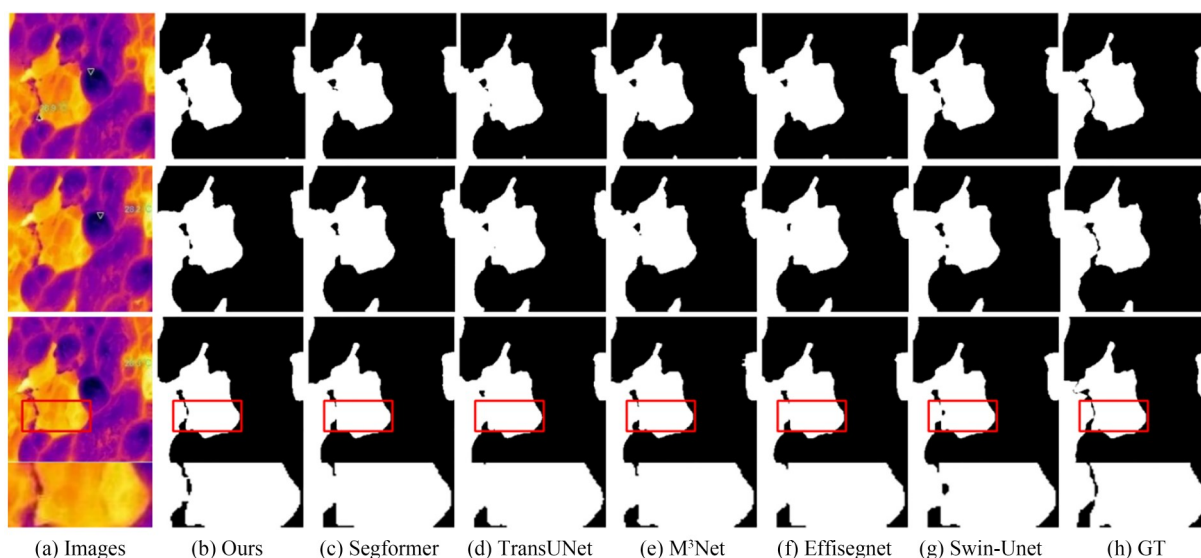


图 7 过浮选红外泡沫分割结果

Fig. 7 Over-footation infrared froth segmentation results

为了客观定量地衡量各种算法的性能,利用铅浮选红外泡沫数据集集中的测试集作为实验对象进行定量评估。各算法分割结果定量比较如表 1 所示,其中 IoU 和 Dice 指标以“均值±标准差”的形式给出。为进一步验证本文方法的显著

性,将其与 Swin-Unet 和 TransUNet 的结果进行配对 t 检验。其中,“+”表示与 TransUNet 相比具有显著性差异($P<0.05$);“*”表示与 Swin-Unet 相比具有显著性差异($P<0.05$);“+*”,表示与两者相比均具有显著性差异($P<0.05$)。本文方法

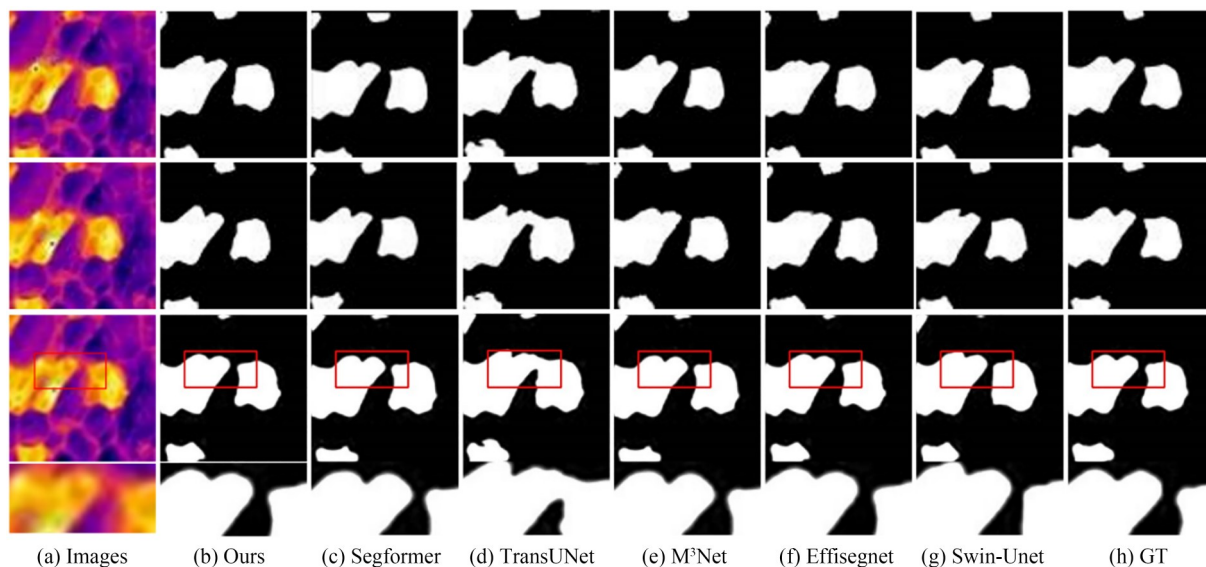


图 8 正常浮选泡沫红外图像分割结果

Fig. 8 Normal flotation infrared froth segmentation results

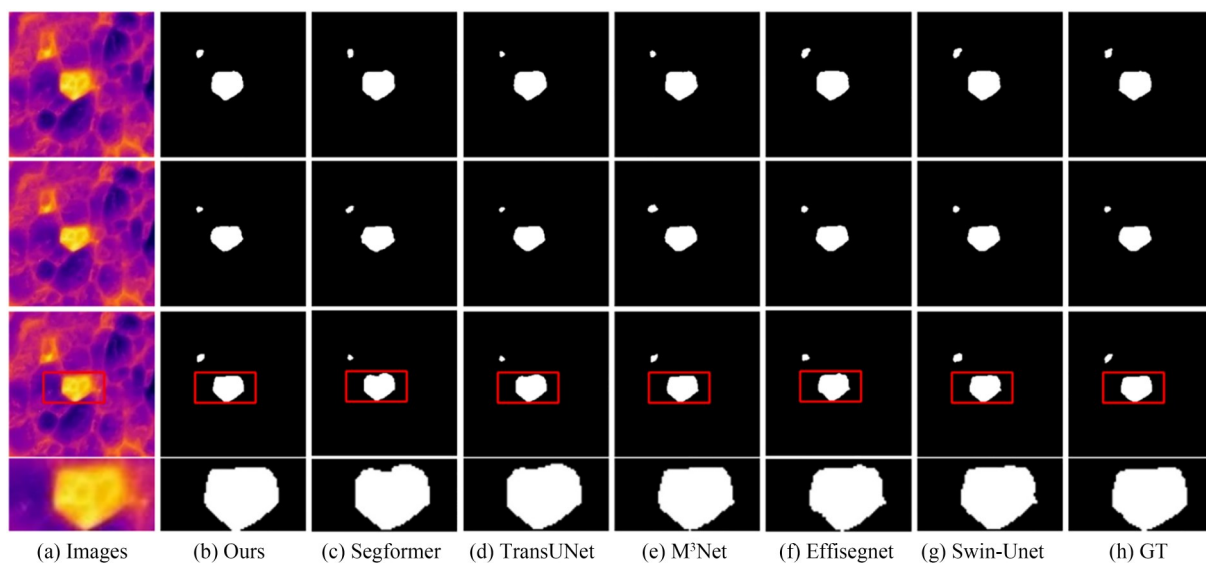


图 9 欠浮选泡沫红外图像分割结果

Fig. 9 Under flotation infrared froth segmentation results

在三种不同工况下, mIoU 分别达到了 0.926 ± 0.035 (正常浮选), 0.894 ± 0.045 (欠浮选) 和 0.909 ± 0.043 (过浮选), 均要优于其他模型。值得注意的是, 基于 U-Net 结构优化的模型, 如 Swin-Unet 和 TransUNet 其 MAE 指标在所有测试场景中均比本文方法差。这归因于本文方法更复杂的网络结构及其对细节信息的更好捕捉能力, 特别是在泡沫红外显著目标分割这种细节丰富且轮廓模糊的任务中。综上所述, 本文方法

表现出较高的分割精度, 不同工况下均表现出良好的精度和鲁棒性, 满足浮选生产工况动态变化的需求。

5.3 分割网络各个模块消融实验

为分析 FIRSegNet 中各模块的有效性和贡献。在铅浮选红外泡沫数据集上进行消融研究, 实验结果如图 10 和表 2 所示。图 10 展示了 DCA 模块对模型特征关注区域的影响, 采用 Grad-CAM 热力图可视化进行特征响应区域的对比展

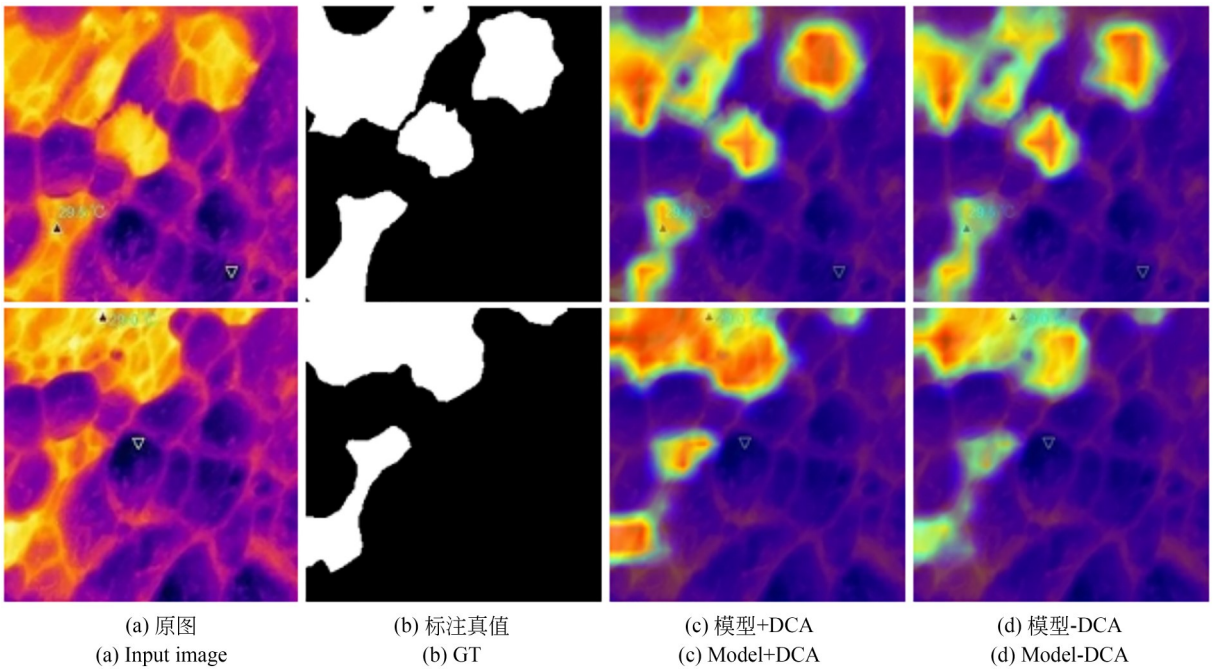


图 10 DCA 模块对特征关注区域的影响可视化
Fig. 10 Visualization of DCA's Effect on Feature Attention

表 1 几种主流方法分割结果比较
Tab. 1 Comparison of segmentation results of several mainstream methods

Method	Under-flotation			Normal flotation			Over-flotation		
	mIoU $\uparrow$	mDice $\uparrow$	MAE $\downarrow$	mIoU $\uparrow$	mDice $\uparrow$	MAE $\downarrow$	mIoU $\uparrow$	mDice $\uparrow$	MAE $\downarrow$
Segformer	0.882 $\pm$ 0.047	0.931 $\pm$ 0.042	0.026	0.912 $\pm$ 0.039	0.950 $\pm$ 0.035	0.020	0.893 $\pm$ 0.045	0.941 $\pm$ 0.045	0.023 2
TransUNet	0.873 $\pm$ 0.051	0.924 $\pm$ 0.048	0.028	0.905 $\pm$ 0.043	0.944 $\pm$ 0.037	0.021	0.884 $\pm$ 0.049	0.934 $\pm$ 0.041	0.024
Swin-Unet	0.878 $\pm$ 0.049	0.927 $\pm$ 0.045	0.027	0.909 $\pm$ 0.041	0.953$\pm$0.036	0.020	0.889 $\pm$ 0.047	0.937 $\pm$ 0.046	0.023
M ³ Net	0.891 $\pm$ 0.046	0.939$\pm$0.042	0.021	0.917 $\pm$ 0.038	0.951 $\pm$ 0.036	0.019	0.898 $\pm$ 0.044	0.944 $\pm$ 0.044	0.021
effisegnet ^[17]	0.880 $\pm$ 0.050	0.929 $\pm$ 0.046	0.027	0.911 $\pm$ 0.039	0.951 $\pm$ 0.037	0.020	0.887 $\pm$ 0.048	0.932 $\pm$ 0.040	0.024
Ours	0.894$\pm$0.045⁺*	0.938$\pm$0.041⁺*	0.024	0.926$\pm$0.035⁺*	0.953$\pm$0.035⁺	0.019	0.909$\pm$0.043⁺*	0.946$\pm$0.038⁺*	0.020

示。相比未使用 DCA 模型的图 10(d)热力图,引入 DCA 后的效果如图 10(c)热力图所示,模型在目标区域的响应更为显著且集中。

对比表 2 中的第一行与第二、三行,可观察到在网络中添加 DCA 模块或 RRM 模块能提升分

割性能。如表 2 第四行所示,与基线网络(第一行)相比,同时包含 DCA 和 RRM 的网络在 mIoU 和 mDice 方面均有改进:mIoU 从 0.865 0 提升至 0.925; mDice 从 0.910 提升至 0.952; MAE 从 0.030 降到 0.021。值得注意的是,在加入 DCA

表 2 消融实验
Tab. 2 Ablation experiment

Backbone	DCA	RRM	mIoU $\uparrow$	mDice $\uparrow$	MAE $\downarrow$	Param/M	GFLOPs/G
✓			0.865	0.910	0.030	28.321	20.365
✓	✓		0.892	0.930	0.028	28.433	21.879
✓		✓	0.901	0.936	0.025	28.478	20.784
✓	✓	✓	0.925	0.952	0.021	28.589	22.298

与 RRM 模块后,模型的参数量仅从 28.321 M 增加至 28.589 M,增幅不到 1%,而 GFLOPs 从 20.365 G 增加至 22.298 G。结果表明,所提出模块在计算开销增加不大的情况下,显著提升了分割精度。总体而言, FIRSegNet 中的 DCA 和 RRM 模块对图像分割任务有益,能有效捕获全局和局部语义特征,尤其在浮选图像上效果显著,验证了本文方法的有效性。

5.4 浮选泡沫稳定度估计及比较分析

本文分别选取三段正常浮选、过浮选、欠浮选泡沫各 10 个红外视频样本,每个视频样本为 2 s,对合并及破裂泡沫进行红外显著目标分割,对分割出来的区域进行像素数目统计,然后利用式(17)计算出相应稳定度 S ,统计的稳定度曲线如图 11 所示。可以看出图 11(c)过浮选泡沫的稳定度最高,图 11(b)欠浮选泡沫的稳定度最低,而图 11(a)正常浮选泡沫的稳定度适中,与实际

工况下泡沫的稳定度吻合。

为了进一步验证本文稳定度估计的准确性,在铅浮选双模态泡沫数据集上进行实验。每隔 3 秒计算一次泡沫稳定度,将采集的实验数据输入搭建好的分割模型,通过式(17)计算得到稳定度的预测值。随后,利用人工标注的可见光泡沫图像,同样根据式(17)计算稳定度,将其作为参考值,用于评估模型预测的准确性。三种状态下稳定度参考值与预测值的对比结果如图 11(d)所示,并选择 RMSE, MAE, MAPE 和 AMAPE 作为评价指标,并计算三种工况下的平均值,实验统计结果如表 3 所示。三种工况下 RMSE, MAE, MAPE 和 AMAPE 的平均值分别为 0.014 5, 0.013 3, 2.384 8% 和 0.025 6,可以看出预测结果与实际值较接近,预测准确性较高。进一步分析发现,误差在欠浮选工况下表现最大,其次为正常浮选,过浮选工况下误差最低。造成

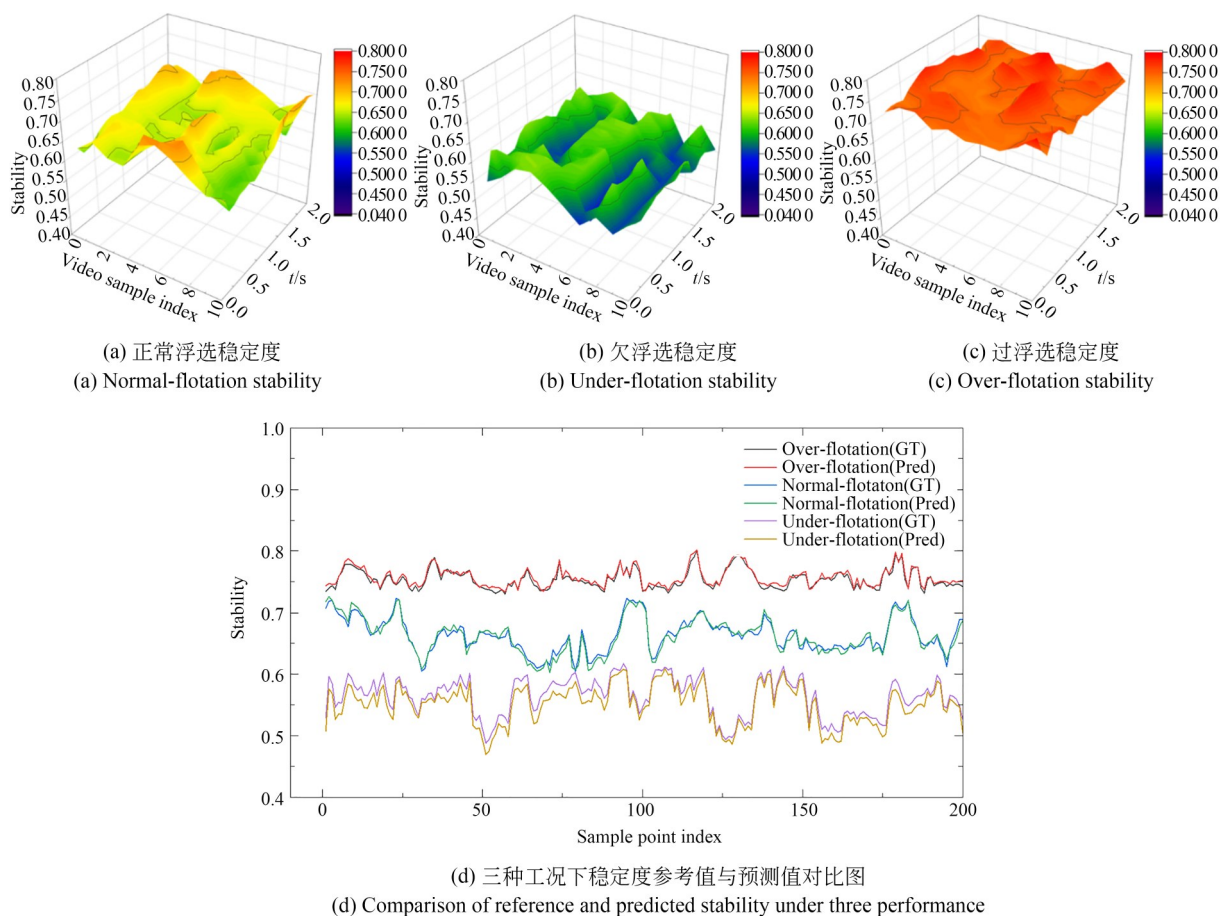


图 11 不同工况下泡沫稳定度估计展示

Fig. 11 Froth stability of different performance display

这一差异的主要原因是, 欠浮选状态下泡沫运动剧烈, 存在大量小气泡, 释放热量不明显, 导致红外热成像显著性不足, 造成不稳定小气泡存在欠分割, 从而影响预测精度。

表 3 三种工况下泡沫稳定度预测性能
Tab.3 Prediction performance of foam stability under three working conditions

工况	RMSE ↓	MAE ↓	MAPE/% ↓	AMAPE ↓
过浮选	0.010 8	0.010 2	1.853 4	0.019 2
正常浮选	0.014 6	0.012 9	2.115 4	0.022 8
欠浮选	0.018 3	0.016 8	3.185 6	0.034 8
平均值	0.014 5	0.013 3	2.384 8	0.025 6

为验证现有泡沫稳定度视觉检测方法抗光照变化影响的性能, 选取包含不稳定气泡的前后两帧 256×256 图像 I_k 和 $I_{k+\Delta T}$ 进行实验。图 12 给出了不同光照下(从上往下分别是光照不足, 过

曝光, 正常光照)的实验结果。提取的前后两帧图像 I_k 和 $I_{k+\Delta T}$ 分别如图 12(a) 和图 12(b)所示。文献[5]使用图像配准后的两个连续图像之间的差值图像来识别破裂的气泡, 图 12(c)为该方法的检测结果, 光照改变对泡沫颜色影响较大, 该方法容易出现误检测和漏检测。文献[6]首先对不稳定气泡进行初步检测, 然后再对初步检测区域与 I_k 对应的区域计算相似度, 相似度低的被识别为不稳定区域, 最后利用分水岭进行分割, 结果如图 12(d)所示, 该方法依赖于分水岭进行分割, 受光照影响大, 大气泡区域过分割、小气泡区域欠分割也会降低检测精度。文献[7]采用改进 Mask2Former 对不稳定区域进行分割, 分割结果如图 12(e)所示, 光照不足情况下无法有效提取不稳定区域的特征导致漏检测。文献[8]基于深度学习特征点匹配方法估算稳定度, 结果如图 12(f), 在光照改变的条件下, 特征点匹配率发生较大变化。

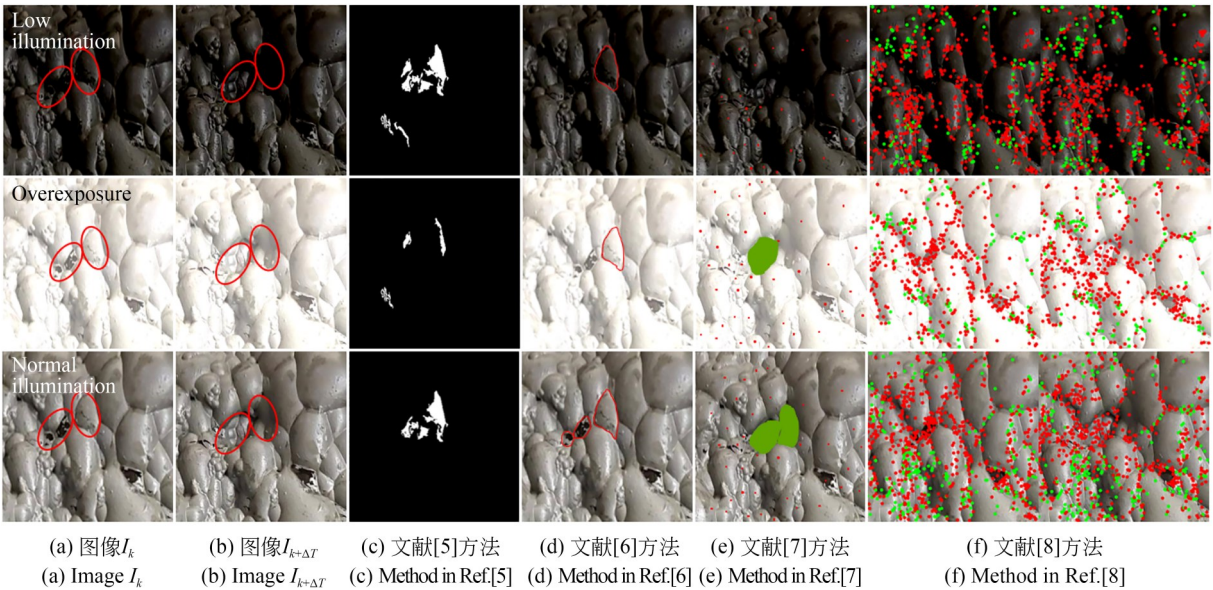


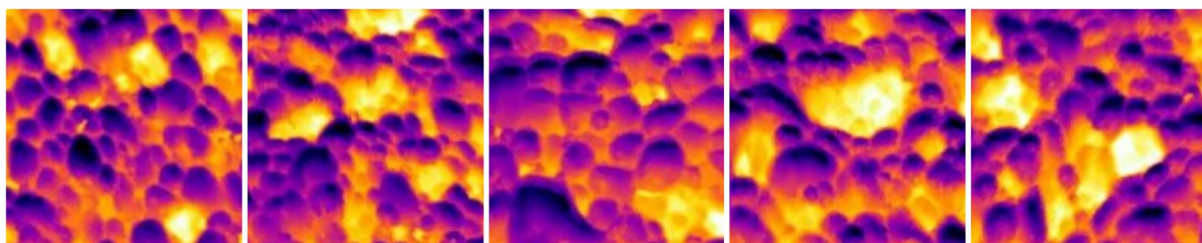
图 12 现有文献方法实验结果
Fig. 12 Experimental results of existing literature methods

将本文方法进行工业现场试验, 在正常浮选条件下采集光照正常、过曝光和光照不足 3 种工况下的泡沫视频图像进行实验, 结果如图 13 所示。以专家人工标注可见光图像并根据式 (17) 计算的稳定度为参考值, 对实验结果进行评估, 并与现有文献方法进行比较, 结果如表 4

所示。在正常光照条件下, 本文方法的平均误差为 0.013 2, 与文献[6]的误差(0.012 8)基本持平。但在过曝光和低照度光照条件下, 本文方法的平均误差分别为 0.013 5 和 0.012 8, 显著低于其他文献方法。与误差最高的文献[5]方法相比, 本文方法的误差分别降低了约 54.2%



(a) 不同光照强度下的可见光图像
(a) Visible light images at different light intensities



(b) 可见光对应的红外图像
(b) Infrared image corresponding to visible light



(c) 检测结果
(c) Test results

图 13 不同光照下泡沫红外成像实验结果

Fig. 13 Foam infrared imaging experimental results under different lighting

表 4 不同光照下稳定度检测结果比较(MAE)

Tab. 4 Stability results under different lighting (MAE)

光照条件	文献[5]方法	文献[6]方法	文献[7]方法	文献[8]方法	本文方法
正常光照	0.017 2	0.012 8	0.016 3	0.018 5	0.013 2
过曝光	0.029 5	0.017 1	0.018 5	0.025 5	0.013 5
光照不足	0.020 5	0.016 0	0.167	0.017 9	0.012 8

和 37.6%，充分体现了其在复杂光照环境下的稳定性。

6 结 论

本文对泡沫表面进行红外热成像,将不稳定气泡检测转换为红外显著目标分割,提出一种泡沫红外显著目标分割的 FIRSegNet 模型及稳定度估计方法。基于 U-Net 架构进行 FIRSegNet

模型构建,采用 Swin Transformer 模块构建编码器网络捕捉长距离依赖关系,提升模型对图像全局上下文的理解能力。在跳跃连接中嵌入双重交叉注意力模块(DCA),以更好融合不同尺度红外泡沫图像的特征信息,增强模型分割不同大小红外显著性区域的能力。设计一种残差细化模块(RRM),实现显著目标的精细化分割,根据分割结果计算泡沫稳定度。实验结果表明:本文方法在自建数据集能取得较好的分割效果,分割精

度较现有方法有一定提升,mIoU平均提升1.73%,mDice平均提高0.63%,MAE平均降低11.6%;正常光照下不同工况泡沫稳定度预测的RMSE,MAE,MAPE和 AMAPE的平均值分别为0.014 5,0.013 3,2.384 8%和0.025 6,预测准确性较高;在过曝光和低照度光照条件下的平均预测误差分别为0.013 5和0.012 8,显著低于现有文献方法。本文方法具有较强的抗光照变化影响能力,实现了不同工况下泡沫稳定度的准确定量描述,在复杂光照变化环境下保持较高的稳定性,为后续的浮选生产工况与指标预测奠定

基础。

作者贡献声明:

廖一鹏:研究方法的提出、论文构思、实验设计、模型构思、结果验证、论文修改和审核;

廖玉华:模型构建与训练、实验与测试、初稿撰写;

董则希:数据集构建、实验数据整理;

白森杰:实验结果的可视化编程;

何建军:现场数据采集、原始数据整理与分类。

参考文献:

- [1] 孙磊,唐倩,廖一鹏,等. 浮选泡沫低照度图像颜色深度编解码校正及多尺度增强[J]. 光学精密工程, 2025, 33(10): 1609-1626.
SUN L, TANG Q, LIAO Y P, *et al.* Multi-scale enhancement and color depth codec correction of flotation foam low illumination images[J]. *Opt. Precision Eng.*, 2025, 33(10): 1609-1626. (in Chinese)
- [2] CHEN X, LIU D, YU L Z, *et al.* Recent advances in flotation froth image analysis via deep learning [J]. *Engineering Applications of Artificial Intelligence*, 2025, 147: 110283.
- [3] ZHAO L N, ZHANG Q. A significant review of froth stability in mineral flotation[J]. *Chemical Engineering Science*, 2025, 302: 120738.
- [4] NEETHLING S J, BRITO-PARADA P R. Predicting flotation behaviour - The interaction between froth stability and performance[J]. *Minerals Engineering*, 2018, 120: 60-65.
- [5] JAHEDSARAVANI A, MARHABAN M H, MASSINAEI M. Prediction of the metallurgical performances of a batch flotation system by image analysis and neural networks[J]. *Minerals Engineering*, 2014, 69: 137-145.
- [6] ZHANG H, TANG Z H, XIE Y F, *et al.* A similarity-based burst bubble recognition using weighted normalized cross correlation and chamfer distance [J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(6): 4077-4089.
- [7] 魏凯,王然风,王珺,等. 基于质心-凸包-自适应聚类法的浮选泡沫动态特征提取[J]. 工矿自动化, 2024, 50(8): 151-160.
WEI K, WANG R F, WANG J, *et al.* Dynamic feature extraction for flotation froth based on centroid-convex hull-adaptive clustering[J]. *Journal of Mine Automation*, 2024, 50(8): 151-160. (in Chinese)
- [8] 刘惠中,阮怡晖,闻成钰,等. 基于改进的SuperGlue模型的浮选泡沫稳定度检测方法研究[J]. 有色金属(选矿部分), 2024(4): 97-104.
LIU H Z, RUAN Y H, WEN C Y, *et al.* Study on flotation froth stability detection method based on improved SuperGlue model [J]. *Nonferrous Metals (Mineral Processing Section)*, 2024(4): 97-104. (in Chinese)
- [9] 廖一鹏,张进,王志刚,等. 结合双模多尺度CNN特征及自适应深度KELM的浮选工况识别[J]. 光学精密工程, 2020, 28(8): 1785-1798.
LIAO Y P, ZHANG J, WANG Z G, *et al.* Flotation performance recognition based on dual-modality multiscale CNN features and adaptive deep learning KELM [J]. *Opt. Precision Eng.*, 2020, 28(8): 1785-1798. (in Chinese)
- [10] CHEN J N, MEI J R, LI X H, *et al.* TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers[J]. *Medical Image Analysis*, 2024, 97: 103280.
- [11] CAO H, WANG Y Y, CHEN J, *et al.* Swin-unet: unet-like pure transformer for Medical image segmentation [C]. *Computer Vision-ECCV 2022 Workshops. Cham: Springer Nature Switzerland*, 2023: 205-218.
- [12] XIE E, WANG W, YU Z, *et al.* SegFormer: Simple and efficient design for semantic segmentation with transformers[J]. *Advances in Neural Informa-*

- tion Processing Systems, 2021, 34: 12077-12090.
- [13] YUAN Y, GAO P, TAN X Y. M³Net: Multilevel, Mixed and multistage attention network for salient object detection [J]. *arXiv preprint arXiv: 2309.08365*, 2023.
- [14] ATES G C, MOHAN P, CELIK E. Dual Cross-Attention for medical image segmentation[J]. *Engineering Applications of Artificial Intelligence*, 2023, 126: 107139.
- [15] HENDRYCKS D, GIMPEL K. Gaussian error linear units (gelus) [J]. *arXiv preprint arXiv: 1606.08415*, 2016.
- [16] YANG F, ZAMZMI G, ANGARA S, *et al.* Assessing inter-annotator agreement for medical image segmentation [J]. *IEEE Access*, 2023, 11: 21300-21312.
- [17] VEZAKIS I A, GEORGAS K, FOTIADIS D, *et al.* EffiSegNet: gastrointestinal polyp segmentation through A pre-trained efficientnet-based network with a simplified decoder[C]. *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. July 15-19, 2024, Orlando, FL, USA. IEEE, 2024: 1-4.

通讯作者:

廖一鹏(1982—),男,福建泉州人,博士,副教授,硕士生导师。主要研究方向为图像处理与分析、工业视觉检测与机器学习、嵌入式边缘计算等。E-mail: fzu_lyp@163.com



廖玉华(2002—),男,江西赣州人,硕士研究生,2023年于太原理工大学获得学士学位,主要从事机器视觉以及人工智能方面的研究。E-mail: 3109074757@qq.com