



# 改进 UNet ++ 的瓷器文物显微气泡分割

刘阳洋<sup>1,2</sup>, 耿国华<sup>1,2</sup>, 刘鑫达<sup>1,2</sup>, 李展<sup>1,2</sup>, 路正涵<sup>1,2</sup>

(1. 西北大学 文化遗产数字化国家地方联合工程研究中心, 陕西 西安 710127;

2. 西北大学 可视化技术研究所, 陕西 西安 710127)

**摘要** 对瓷器文物显微气泡的分割, 可以更加清晰地观察瓷器表面微观气泡的形态、数量以及分布规律, 进而辅助文物专家进行瓷器碎片分类和文物鉴定等工作。但瓷器显微图像中气泡复杂多变, 大小及分布不均匀, 现有图像分割方法难以适应瓷器显微气泡特征。因此, 该文提出一种基于卷积激活单元的网络 AGUNet ++, 该网络重新设计密集跳跃连接, 节点间采用 Z 字形连接方式, 充分提取图像语义特征, 防止信息丢失; 同时, 在卷积单元的密集跳跃连接处, 结合注意力门控模块 Attention Gate 提出卷积激活单元 CAU, 增强与瓷器文物显微气泡分割任务相关的气泡区域学习, 抑制不相关的区域; 在训练过程中对每一层子网络的输出采用深度监督和交叉熵损失, 有效增强瓷器文物显微气泡特征提取能力, 细化分割结果。该方法在 SD-saliency-900 以及 PRMI 数据集上的实验结果表明, 与经典图像分割网络相比, AGUNet ++ 在 MIoU、Precision、Recall 和 F1 分数中均有一定的提升, 表现出更好的分割效果。

**关键词** 瓷器文物显微图像; 显微图像分割; UNet ++; 注意力门

**中图分类号:** TP391.41 **DOI:** 10.16152/j.cnki.xdxbzr.2025-01-011

## Microscopic bubbles segmentation of porcelain relics based on improved UNet ++

LIU Yangyang<sup>1,2</sup>, GENG Guohua<sup>1,2</sup>, LIU Xinda<sup>1,2</sup>, LI Zhan<sup>1,2</sup>, LU Zhenghan<sup>1,2</sup>

(1. National and Local Joint Engineering Research Center for Cultural Heritage Digitization, Northwest University,

Xi'an 710127, China; 2. Institute of Visualization Technology, Northwest University, Xi'an 710127, China)

**Abstract** The segmentation of microscopic bubbles of porcelain relics can provide a clearer observation of the morphology, quantity, and distribution of micro bubbles on the surface of porcelain, which is of great significance in assisting relic experts in classifying porcelain cultural relic fragments and identifying porcelain cultural relics. However, the bubbles in porcelain microscopic images are complex and varied, with uneven size and distribution. Existing image segmentation methods are difficult to adapt to the characteristics of porcelain microscopic bubbles. Therefore, a network named AGUNet ++ based on convolution attention unit is proposed.

**收稿日期:** 2024-04-20

**基金项目:** 国家自然科学基金(62271393); 陕西省教育厅一般项目(19JK0842); 虚拟现实技术与系统全国重点实验室(北京航空航天大学)开放课题基金(VRLAB2024C02)。

**第一作者:** 刘阳洋, 女, 工程师, 从事智能信息处理、图像处理研究, yyliu@nwu.edu.cn。

**通信作者:** 耿国华, 女, 教授, 博士生导师, 从事智能信息处理、虚拟现实与可视化研究, ghgeng@nwu.edu.cn。

This network utilizes a zigzag connection approach between nodes to fully extract image semantic features and prevent information loss. Meanwhile, a convolution attention unit is introduced by combining the dense skip connection of the convolution unit with the attention gate. The CAU enhances the learning of bubble regions relevant to the task of microscopic bubble segmentation in porcelain artifacts while suppressing irrelevant regions. Deep supervision and cross entropy loss are applied to the output of each sub network layer during the training process, which effectively enhance the ability to extract microscopic bubble features in porcelain artifacts and refine the segmentation results. The experimental results of this method on the SD-saliency-900 and PRMI demonstrate that AGUNet ++ exhibits certain improvements in MIoU, Precision, Recall, and F1\_score, showing better segmentation performance compared to classical image segmentation networks.

**Keywords** porcelain relics micro image; microimage segmentation; U-Net ++; attention gate

在近一个世纪的考古发掘历程中,众多瓷器文物的出土为研究中国古代社会提供了丰富的实物资料。这些瓷器文物不仅是中国古代生产力发展的象征,也是中华文明进步的重要标志。瓷器的器型、釉色、表面纹饰及其微观特征等元素,共同构成了不同历史时期瓷器器物的独特风格和特点。鉴于出土瓷器文物及碎片的数量庞大、种类繁多且分布广泛,对其进行系统的修复和鉴定不仅具有深远的学术价值,也对文化遗产的保护与传承具有不可替代的意义<sup>[1-2]</sup>。

瓷器文物由于其本身精细且易碎,传统的手工复原和拼合存在诸多问题,计算机辅助虚拟复原能够有效避免人为因素的干扰,不受时间和空间的限制,可高效安全地完成文物复原,避免文物的二次破坏<sup>[3-4]</sup>。进行瓷器文物的复原工作之前,对大量瓷器文物碎片进行精确分类是一项技术要求极高的任务,且人工分类费时费力,而计算机辅助瓷器碎片分类方法可识别其类别,辅助碎片拼接和鉴定等后续任务<sup>[5]</sup>。鉴定瓷器文物的真伪、确定其产地(窑口)、断定其制作年代以及评估其历史价值,是瓷器文物鉴定的基本任务<sup>[6]</sup>。这些任务不仅是学术研究的需要,更是国家文化遗产保护与传承的战略需求。

瓷器文物鉴定领域主要采用2种方法:直观鉴定(目鉴)与科学鉴定(科鉴)。直观鉴定依赖于专家的经验 and 直观感受,而科学鉴定则通过科学仪器和技术手段进行分析<sup>[7]</sup>。在传统方法中,直观鉴定占据主导地位,但其准确性往往受限于鉴定者的专业知识和经验。随着现代科技发展,利用电子显微镜观察瓷器釉面各相(玻璃相、气相、晶相)的显微结构,在新的维度上为瓷器文物的鉴定提供了参考<sup>[8]</sup>。在瓷器文物的修复过程中,对大量瓷器碎片进行分类是一项基础且关键

的工作<sup>[9-11]</sup>。然而,传统的人工分类方法不仅耗时耗力,而且由于瓷器碎片外观的相似性,分类的准确性往往难以保证。针对这一问题,瓷器文物显微图像中的气泡提取技术显示出其独特的价值。显微图像分析能够揭示瓷器碎片的微观气泡结构,这些结构特征可以作为辅助分类的重要依据,从而提高分类的效率和准确性<sup>[12-13]</sup>。因此,瓷器文物显微图像中的气泡提取不仅对瓷器文物碎片的分类具有重要的辅助作用,而且对于瓷器文物的鉴定也具有重要的参考意义。通过显微图像分析,可以更全面地理解瓷器的微观特征,为瓷器科学的科学鉴定提供更为坚实的基础。

瓷器显微气泡的形成与烧制过程中的多个环节密切相关,如用料、粉碎、腐泥、炼泥、成型、凉坯、上釉、烧制等,每一个步骤都对气泡的形成有直接影响。由于古代烧制瓷器的工艺多为私人作坊所掌握,生产技术各不相同,因此,即便是同一作坊生产的瓷器,其显微气泡的特征也会存在显著差异<sup>[14]</sup>。这种差异性为瓷器的鉴定和研究提供了重要的线索。

在现代科技的帮助下,研究人员开始利用先进的图像处理技术来分析这些显微气泡的特征。李庆武等提出一种利用小波分析对瓷器显微图像进行去噪处理的方法<sup>[15]</sup>,并通过分水岭分割技术对去噪后的图像进行分割,从而提高了图像分割的效果,这种方法不仅能够清晰地展示瓷器显微图像中的气泡结构,还为后续的气泡特征分析提供了高质量的图像基础;朱顺龙则通过对不同年代耀州瓷显微图像的气泡进行数据曲线分析<sup>[16]</sup>,探究了不同年代瓷器显微图像气泡的区别,为耀州瓷的显微鉴定提供了科学依据,这种分析方法不仅有助于识别不同年代的瓷器,还能够揭示不同生产技术和工艺对瓷器气泡特征的影响;刘国

高进一步提出了一种基于快速曲波变换的方法<sup>[17]</sup>,对瓷器显微图像进行去噪和数据增强,这种方法不仅能有效去除图像中的噪声,还能增强图像的局部特征,使得气泡的边界更加清晰,随后,他们采用分水岭算法对瓷器显微图像中的气泡进行分割,进一步提高了气泡识别的准确性。

传统图像分割算法虽然一定程度上提升了图像的分析效果,但具有对噪声鲁棒性低、难以获取高级语义信息等缺点。与传统图像分割算法相比,基于深度学习的图像语义分割算法对每个像素点类别进行更精确分类,能够充分提取语义特征信息。随着深度学习技术的出现,图像分割领域迎来了新的发展机遇。Long 等提出用于图像语义分割的全卷积网络 (fully convolutional network, FCN) 作为一种典型的深度学习图像分割框架<sup>[18]</sup>,该框架不含全连接层,可适应任意尺寸的输入图像;Ronneberger 等提出经典图像分割网络 U-Net<sup>[19]</sup>,该网络属于一种特殊的 FCN 网络,以更快、更少的图像进行训练,但在下采样过程中容易丢失特征信息;基于 U-Net 网络,Zhou 等提出一种全新的 UNet ++ 网络架构<sup>[20]</sup>,将 U-Net 网络的嵌套密集跳转连接进行拼接,实现更高效的信息传输;为充分学习不同层级相邻节点间的语义信息,Su 等提出了不同于密集跳跃连接的融合上级特征信息的 ZUnet ++ 网络<sup>[21]</sup>,节点的输入既包括上采样的深层特征,也包括下采样的浅层特征,从而更全面地学习不同层级相邻节点间的语义信息;Ardiyanto 提出一种高效的表面缺陷检测 ESD-Net 网络<sup>[22]</sup>,利用 GhostNet 融合块来代替传统编码器解码器网络中的所有卷积运算,降低网络的计算复杂性,取得精确的分割精度,使用增强的跳接连接和全局辅助层 GAL 以融合多尺度上下文语义特征信息,提高效率和边缘计算兼容性。随后,为放大图像关键特征信息而抑制无关特征信息,提高信息传输效率<sup>[23]</sup>,注意力机制被提出,使得模型能够更好地抓住数据中的根本模式,提升泛化能力。注意力机制最早出现在自然语言中,后被成功引入计算机视觉研究<sup>[24]</sup>。Chen 等提出将共享网络应用到语义分割领域<sup>[25]</sup>。常用的 2 种经典注意力类型有乘性注意力<sup>[26]</sup>和加性注意力<sup>[27]</sup>,乘性注意力计算速度快,节约内存,但加性注意力在大维度输入特征上表现更好<sup>[28]</sup>。

尽管这些基于深度学习的图像分割方法在场景理解、工业缺陷检测、医学图像分析、遥感等数

据集上取得了显著的成果<sup>[29-30]</sup>,但目前国内外针对瓷器文物显微数据集上的图像分割提取的研究仍然较少。这表明,将深度学习技术应用于瓷器显微图像的分析 and 鉴定,还有很大的发展空间和潜力。与其他图像分割数据集相比,瓷器显微气泡数据更具有其自身特点。由于瓷器显微图像数据采集难,且气泡标注需要依靠专业人员进行,因此瓷器显微图像数量样本较少。

本文基于 U-Net 网络框架,将基于深度学习的图像分割算法应用在瓷器文物显微图像数据集中,针对瓷器文物显微图像中气泡结构边界模糊、容易黏连等问题,基于 UNet ++ 网络框架,提出基于卷积激活单元的网络 AGUNet ++。该模型在节点间采用 Z 字形连接方式来充分提取语义特征,防止信息丢失。同时,在卷积单元的密集跳跃连接处结合注意力门控模块提出卷积激活单元,增强与瓷器文物显微气泡分割任务相关的气泡区域学习,抑制不相关的区域,以获得更准确、细致的分割结果。最后在工业检测领域的带钢表面缺陷数据集 SD-saliency-900 以及瓷器文物显微气泡图像数据集 PCRI 上验证了分割准确性。

## 1 相关工作

### 1.1 FCN

全卷积网络 FCN 通过显示每个像素的分类结果来实现对任意大小输入图像的像素级分类。FCN 网络将传统卷积神经网络 CNN 架构的全连接层换成卷积层,在得到的特征图上加入反卷积进行上采样,将特征数据恢复到与输入图像大小相同的尺寸,使得网络能够处理任意尺寸的输入图像。FCN 经典架构如图 1 所示。

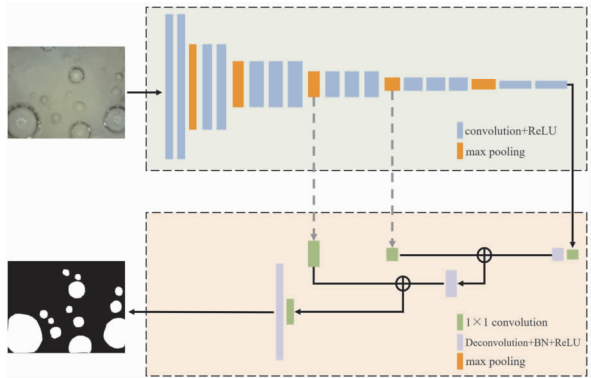


图 1 FCN 网络结构

Fig. 1 Structure of FCN

1.2 U-Net

U-Net 网络由编码器和解码器 2 部分组成(见图 2)。编码器由多个卷积块组成,用于提取图像的低级特征;解码器在结构上相对于编码器是对称的,由多个反卷积块组成,用于学习图像的高级语义信息。此外, U-Net 还通过级联的跳跃连接将编码器和解码器中的相应层连接在一起,能捕捉多尺度特征信息,但是 U-Net 只有一个端到端的路径,在使用过程中不是很灵活,多次卷积和上采样会损失语义信息。其次,跳跃连接非常简单,上采样路径处的节点仅从该相同级别接收跳跃连接。

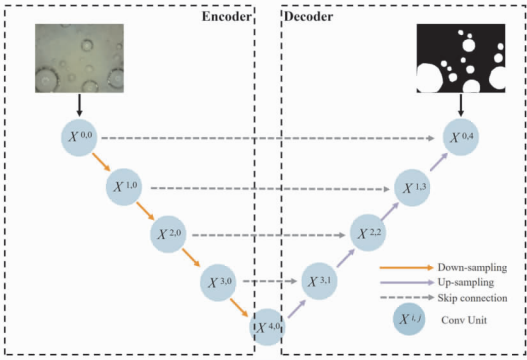


图 2 U-Net 网络结构  
Fig.2 Structure of U-Net

1.3 UNet ++

UNet ++ 引入深度监督、嵌套和密集的跳跃连接来抓取不同层次的特征,提升模型灵活度,增强特征传播和融合能力,从而提高分割的精度和鲁棒性,UNet ++ 网络结构如图 3 所示。

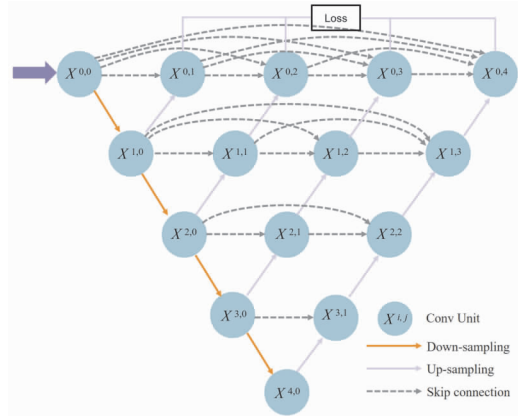


图 3 UNet ++ 网络结构  
Fig.3 Structure of improved UNet ++

图 3 中,  $X^{i,j}$  是 1 个卷积单元,其中  $i$  表示第  $i$  个下采样层,  $j$  代表当前跳跃连接层中第  $j$  个卷积层,  $x^{i,j}$  表示卷积单元  $X^{i,j}$  的输出,  $x^{0,0}$  是原始图像输入的初始节点。图 3 中卷积单元的输出  $x^{i,j}$  可用

公式(1) 表示,

$$x^{i,j} = \begin{cases} H(D(x^{i-1,j})), & j = 0 \\ H([\ (x^{i,k})_{k=0}^{j-1}, U(x^{i+1,j-1}) \ ]), & j > 0 \end{cases} \quad (1)$$

式中:  $x^{i,j}$  表示卷积单元的输出;  $H()$  表示卷积操作和激活函数;  $U()$  表示上采样;  $[\ ]$  表示拼接操作。当  $j = 0$  时,节点只接收来自前 1 个下采样层的输入;当  $j > 0$  时,节点接收  $j + 1$  个输入,包括跳跃连接和上采样层的输入。跳跃连接结合深层和浅层的语义信息,加强信息传输,提高语义分割的准确性。

2 研究方法

2.1 网络结构

UNet ++ 网络结构中的卷积单元  $X^{i,j}$  接收来自下一层的上采样和同层的密集跳跃连接信息,但缺乏多尺度语义信息。本文改进 UNet ++ 网络框架中的同层相邻卷积单元的密集跳跃连接方式,结合注意力门控模块(attention gate, AG)提出基于卷积激活单元(convolution attention unit, CAU)的网络结构 AGUNet ++,如图 4 所示。

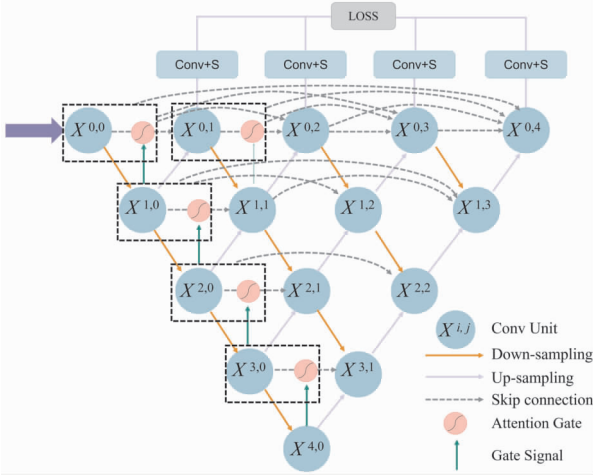


图 4 AGUNet ++ 网络结构  
Fig.4 Structure of AGUNet ++

该网络主要包括 Z 形跳跃连接以及由卷积单元  $X^{i,j}$  和注意力门控模块 AG 组成的卷积激活单元 CAU, Z 形跳跃连接将原本与同层相邻卷积单元的连接方式改为与上一层卷积单元下采样后的输出相连,充分提取图像多尺度语义特征;注意力门控模块可增强与瓷器文物显微气泡分割任务相关的气泡区域特征的学习,抑制非气泡区域特征的学习,自动聚焦不同形状和大小的瓷器文物

显微气泡特征。

## 2.2 卷积激活单元

AGUNet++ 网络中的卷积激活单元 CAU 由卷积单元  $X^{i,j}$  和加性注意力门控模块 AG 组成, CAU 网络结构如图 5 所示。其中,卷积单元  $X^{i,j}$  的基本结构如图 5(a)所示,分别由 2 组  $1 \times 1$  卷积 Conv  $1 \times 1$ 、归一化 BN 和激活函数 ReLU 组成。注意力门控模块如图 5(b)所示,将第  $i$  层的卷积单元输出  $x^{i,j}$  和第  $i+1$  层的卷积单元输出  $x^{i+1,j}$  经过上采样后,分别经过 2 个  $1 \times 1 \times 1$  卷积操作后将其相加,以增强特征区域。接着使用激活函数 ReLU 后再次执行  $1 \times 1 \times 1$  卷积操作得到特征矩阵。再将特征矩阵经过激活函数 Sigmoid 后得到权重图  $\alpha$  ( $\alpha \in [0, 1]$ ), 其中瓷器文物显微气泡特征趋近于 1, 非气泡特征趋近 0。最后对权重图  $\alpha$  经过三线线性插值重采样器与原始特征图进行相乘得到注意力门模块的输出,来增强输入图像中气泡特征表示。注意力门控模块的公式定义如公式(2)和(3)所示。

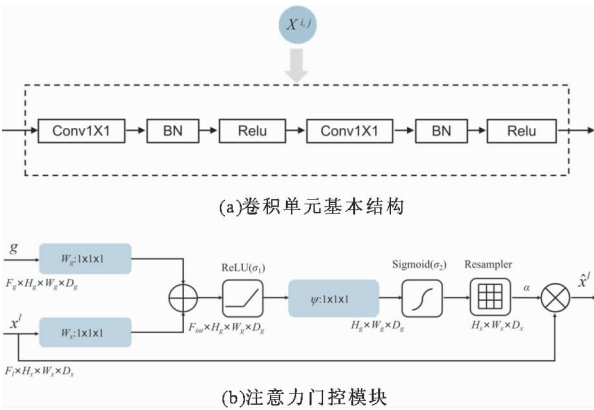


图 5 卷积单元基本结构和注意力门控模块

Fig. 5 Structure of the basic convolution unit and attention gate

$$q_{att,i}^l = \psi^T [\sigma_1 (W_x^T x_i^l + W_g^T g + b_{xg})] + b_{\psi} \quad (2)$$

$$\alpha^l = \sigma_2 [q_{att}^l(x^l, g; \Theta_{att})] \quad (3)$$

式中:  $\sigma_1$  是线性激活函数 ReLU;  $\sigma_2$  是激活函数 Sigmoid;  $W_x$  和  $W_g$  均表示  $1 \times 1$  卷积;  $b_{xg}$  和  $b_{\psi}$  为偏置项;  $\psi$  为卷积操作。

## 2.3 损失函数

损失函数用于描述模型输出值与真实值的差距。模型通过反向传播更新各个参数,降低损失值,使模型输出值更靠近真实值。瓷器显微图像的气泡分割是背景和气泡的二分类问题,实验数据较少,因此需要较为快速的收敛,同时为使该网络

在更多环境下可训练,选择交叉熵损失作为损失函数。

对 AGUNet++ 网络中每个深度监督的子网络得到的输出结果 ( $x^{0,1}, x^{0,2}, x^{0,3}, x^{0,4}$ ) 经过  $1 \times 1$  卷积和 Sigmoid 激活函数进行预测,将这些预测结果的平均值与真实标签进行比较来计算损失,而不仅仅从最后一层输出计算损失,从而加速收敛速度,缓解梯度消失。

将 AGUNet++ 网络的 4 个输出 ( $x^{0,1}, x^{0,2}, x^{0,3}, x^{0,4}$ ) 的损失函数相加求平均,得到的平均损失函数  $L$  如公式(4)所示,其中  $L^k$  是输出节点  $x^{0,k}$  对应的损失函数。公式(5)为二元交叉熵损失函数(binary cross-entropy loss, BCE)。

$$L = \frac{1}{4} \sum_{k=1}^4 L^k \quad (4)$$

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)] \quad (5)$$

式中:  $N$  是样本数;  $y_i$  表示真实值;  $\hat{y}_i$  表示预测值。

## 3 实验设计

### 3.1 数据集

1) SD-saliency-900 数据集是工业检测领域专门针对带钢表面缺陷进行检测的公开数据集。该数据集包含 3 类缺陷,每类 300 张,一共 900 张带钢表面图像及其对应缺陷的标注图像,每张照片的分辨率为  $200 \times 200$ 。

2) PRMI 数据是基于文化遗产数字化国家地方联合工程研究中心收集的大量瓷器文物碎片进行采集的,部分碎片如图 6 所示。针对瓷器碎片样本,使用艾尼提高清显微镜(3R Anyty)对瓷器文物碎片显微结构进行拍摄采集,将放大 200、400、600 倍的成像进行对比,最终选取清晰度更高、显微气泡更明显的 600 倍成像作为标准,构建瓷器文物显微图像数据集。PRMI 数据集可用于考古领域瓷器碎片分类、修复、辨真伪、朝代分析、窑口判断等任务,为瓷器文物修复、分析与鉴定提供了新的角度。实验室采集了 600 张  $640 \times 480$  像素的瓷器文物显微图像,并邀请文物专家对瓷器文物显微图像中的气泡结构使用 labelme 软件进行标注。在做好的成对数据集中,选取 450 张瓷器显微图像对作为训练集,150 张图片对作为测试集,部分数据对如图 7 所示。

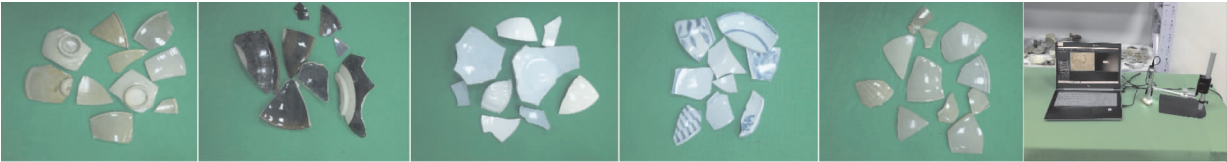


图 6 瓷器碎片实体及数据采集设备

Fig. 6 Porcelain relics fragments entity and data acquisition equipment

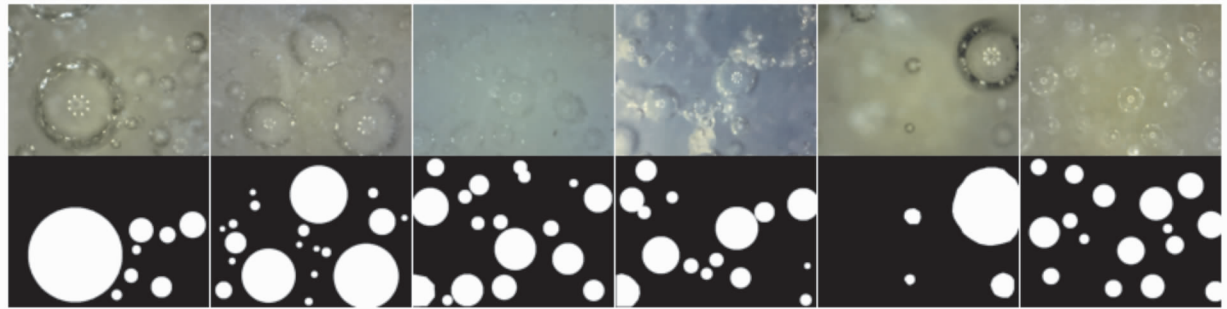


图 7 PRMI 数据集与标签

Fig. 7 PRMI dataset with labels

3.2 实验细节

本文算法基于 Pytorch 框架实现,使用 GeForce GTX 3060 显卡对模型进行训练。本文将输入图像的尺寸大小统一处理为  $256 \times 256$ 。模型训练 epoch 为 200,初始学习率为 0.001,使用 Adma 优化器,参数  $\beta_1$  为 0.9,  $\beta_2$  为 0.999,选择交叉熵作为损失函数对模型进行优化。

3.3 评价指标

为更客观体现本文气泡分割方法在瓷器文物显微图像气泡分割中的优越性,引入混淆矩阵表示每个像素类别分类正确或错误的数量(见表 1)。其中模型将显微图像中的气泡正确预测为气泡,表示为 True Positive(TP),将气泡错误预测为背景,表示为 False Negative(FN);将背景错误预测为气泡,表示为 False Positive(FP);将背景正确预测为背景,则表示为 True Negative(TN)。

表 1 混淆矩阵

Tab. 1 Confusion matrix

| 混淆矩阵 |    | 预测值                 |                    |
|------|----|---------------------|--------------------|
|      |    | 背景                  | 气泡                 |
| 真实值  | 气泡 | True Positive (TP)  | False Negative(FN) |
|      | 背景 | False Positive (FP) | True Negative (TN) |

量化指标采用常用的精确率(Precision)、召回率(Recall)和平均交并比(mean intersection over union)进行评价分割。

精确率(Precision)又叫查准率,表示预测结果为正例的样本中实际为正样本的比例。当反例

被错误预测成正例(FP)的代价很高时,适合用精确率。根据公式可知,精确率越高,FP 越小。计算公式如式(6)所示,

$$\text{Pr} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$

召回率(Recall)又被称为查全率,表示预测结果为正样本中实际正样本数量占全样本中正样本的比例。当正例被错误的预测为反例(FN)产生的代价很高时,适合用召回率。根据公式可知,召回率越高,FN 越小。计算公式如式(7)所示,

$$\text{Re} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$

$F_1$  分数( $F_1$ -Score),又称为平衡  $F$  分数(Balanced Score),它被定义为精确率和召回率的调和平均数。计算公式如式(8)所示,

$$F_1 = \frac{2 \times \text{Pr} \times \text{Re}}{\text{Pr} + \text{Re}} \tag{8}$$

平均交并比体现模型对每一类预测的结果和真实值的交集与并集的比值,求和再平均的结果,为语义分割的标准度量。计算公式如式(9)所示,

$$\text{MIoU} = \frac{1}{k + 1} \sum_{i=0}^k \frac{\text{TP}}{\text{FN} + \text{FP} + \text{TP}} \tag{9}$$

4 实验结果与分析

为验证本文提出的改进的 AGUNet ++ 网络模型的有效性,分别在工业检测领域公共数据集

SD-saliency-900 以及本文提出的数据集 PRMI 上进行图像分割实验,并将实验预测结果分别与其他经典网络模型的实验预测结果进行对比。接着在 PRMI 数据集上对不同层的卷积单元  $X^{i,j}$  处结合不同数量的注意力门控模块 AGs 进行实验,将得到的预测结果进行对比,实验结果分别从可视化定性实验和客观定量指标进行评价。

4.1 定性实验

图 8 展示了 PRMI 数据集在不同网络模型中对测试样本的预测结果,以 PRMI 数据集中 3 组显微气泡图像为例,图 8 为原始气泡图像、真实人工标注图像、U-Net、Attention U-Net、UNet ++、ZU-Net ++ 和 AGU-Net ++ 网络预测出的分割结果,图中红框和绿框标记了预测结果差异较大的气泡部位,并将结果在图像下方进行放大。

从图 8 可以看出,测试样本经过不同网络的预测结果与真实标签图像进行对比,其中 U-Net、Attention U-Net、UNet ++、ZUNet ++ 模型对气泡中心或边缘的预测性能较差,且存在错误预测情况。而本文提出的 AGU-Net ++ 网络能更准确地分割出原瓷器显微图像中的气泡结构。在不同层的卷积单元  $X^{i,j}$  处添加注意力门控模块 AGs 可得到不同数量的卷积激活单元 CAU,图 9 为添加不同数量 AGs 的模型气泡分割可视化结果。图中给出 3 张瓷器文物显微气泡图像的预测结果,可以看出,添加 5 个注意力门控模块 AGs 的预测结果与真实标签值最接近,其他几种情况均存在过度预测或漏预测等情况,AGUNet ++ (w/5AGs) 网络预测结果更准确,细节信息更丰富。

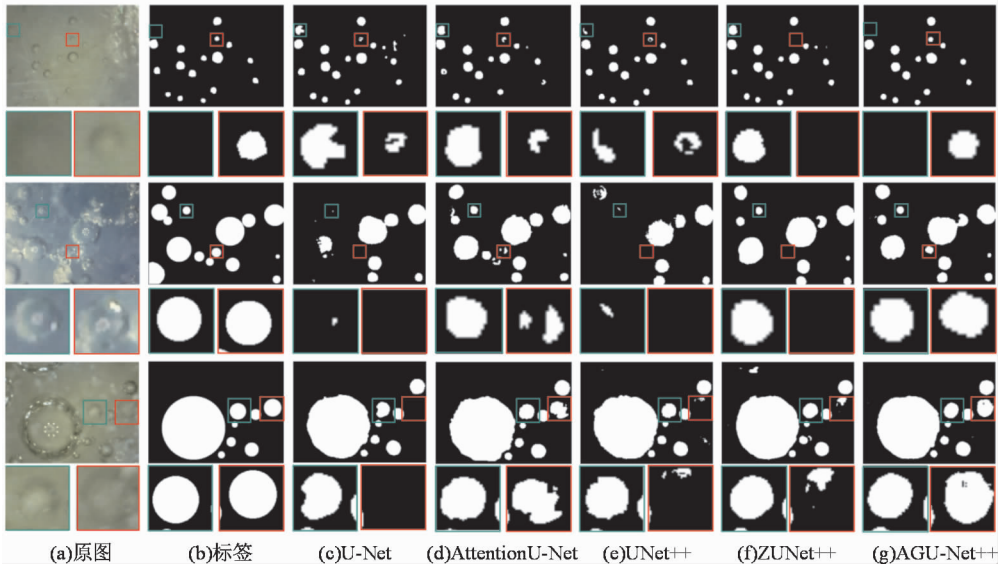


图 8 不同模型气泡分割可视化结果

Fig. 8 Visualization results of bubble segmentation for different models

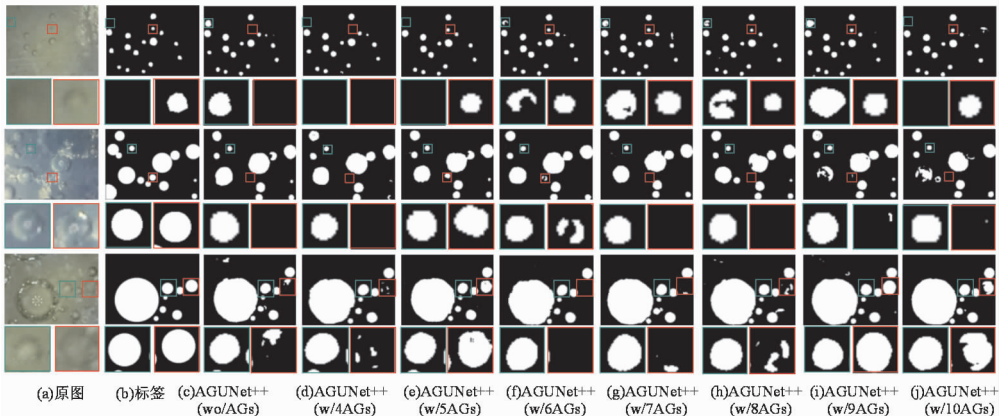


图 9 添加不同数量 AGs 的模型气泡分割可视化结果

Fig. 9 Visualization results of bubble segmentation for models added different numbers of AGs

4.2 定量实验

如表 2 所示为工业检测领域中带钢表面缺陷数据集 SD-saliency-900 的图像分割结果,经典网络的分割结果来自文献[15],可以看出,本文提出的改进的 AGUNet ++ 网络在 SD-saliency-900 数据集上的 MIoU 为 0.856,分割性能最佳。

表 2 SD-saliency-900 数据集不同网络定量比较  
Tab. 2 Quantitative comparisons of different methods on SD-saliency-900

| 方法/指标                  | MIoU  |
|------------------------|-------|
| FCN <sup>[11]</sup>    | 0.793 |
| U-Net <sup>[12]</sup>  | 0.718 |
| ESDNet <sup>[15]</sup> | 0.824 |
| AGUNet ++              | 0.856 |

表 3 PCMI 测试集中各网络评价指标对比  
Tab. 3 Comparison of network evaluation indicators in the test set on PCMI

| 方法                 | MIoU           | Precision      | Recall         | $F_1$          |
|--------------------|----------------|----------------|----------------|----------------|
| U-Net              | 0.853 5        | 0.908 9        | 0.932 2        | 0.920 4        |
| Attention U-Net    | 0.860 9        | 0.913 9        | 0.932 9        | 0.923 3        |
| UNet ++            | 0.861 8        | 0.912 3        | 0.935 5        | 0.924 7        |
| ZUNet ++           | 0.863 9        | 0.917 0        | <b>0.935 6</b> | 0.926 2        |
| AGUNet ++ (w/5AGs) | <b>0.866 2</b> | <b>0.921 8</b> | 0.932 9        | <b>0.927 3</b> |

表 4 添加不同数量注意力门控 AGs 的指标对比  
Tab. 4 Comparison of indicators with different numbers of attention gates

| 方法                  | MIoU           | Precision      | Recall         | $F_1$          |
|---------------------|----------------|----------------|----------------|----------------|
| AGUNet ++ (wo/AG)   | 0.863 9        | 0.917 0        | 0.863 9        | 0.926 2        |
| AGUNet ++ (w/4AGs)  | 0.848 5        | 0.902 6        | 0.932 8        | 0.917 5        |
| AGUNet ++ (w/5AGs)  | <b>0.866 2</b> | <b>0.921 8</b> | <b>0.932 9</b> | <b>0.927 3</b> |
| AGUNet ++ (w/6AGs)  | 0.863 9        | 0.920 6        | 0.930 0        | 0.925 3        |
| AGUNet ++ (w/7AGs)  | 0.853 0        | 0.906 3        | 0.934 9        | 0.920 3        |
| AGUNet ++ (w/8AGs)  | 0.863 5        | 0.925 9        | 0.925 7        | 0.925 7        |
| AGUNet ++ (w/9AGs)  | 0.863 5        | 0.925 9        | 0.925 7        | 0.925 8        |
| AGUNet ++ (w/10AGs) | 0.860 6        | 0.915 6        | 0.932 5        | 0.923 9        |

5 结语

瓷器在烧制过程中随着温度变化会产生不同的气泡结构,分析其气泡结构对瓷器分类和鉴定具有重要的指导意义,瓷器文物显微图像的气泡结构具有体积小、边界模糊等特点。针对瓷器文物显微图像中的气泡结构,提出一种改进的 AGUNet ++ 网络,用于瓷器文物显微图像中气泡分割。该网络在 UNet ++ 基础上采用 Z 字形跳跃连接和不同数量的注意力门控模块,通过增强与瓷器文物显微气泡分割任务相关的气泡区域特征

表 3 展示了在 PRMI 数据集上采用不同方法的各量化指标数值。可以看出,本文改进的 AGUNet ++ 网络模型在 MIoU、Precision、Recall 和  $F_1$  分数达到 0.866 2、0.921 8 和 0.932 9 和 0.927 3。对比 U-Net、Attention U-Net、UNet ++ 和 ZUNet ++ 网络,本文提出的 AGUNet ++ 网络模型在 MIoU、Precision 和  $F_1$  分数均有一定提升。表 4 所示为是否添加以及添加不同数量的注意力门控模块 AGs 的网络预测结果的评价指标,加粗项代表最优结果。可知在卷积单元处添加 5 个注意力门控模块的 AGUNet ++ 网络模型在测试集上的 MIoU、Precision、Recall 和  $F_1$  分数最高,表现最佳。

的学习,抑制非气泡区域特征的学习,自动聚焦不同形状和大小的瓷器文物显微气泡特征,更好地结合上下文语义信息,得到更准确、细致的分割结果。实验结果表明,本文提出的改进 AGUNet ++ 网络在工业检测领域的带钢缺陷数据集 SD-saliency-900 以及瓷器显微图像数据集 PRMI 均表现出更好的分割效果,其 MIoU、Precision、Recall 和  $F_1$  分数值均有一定提升,达到 0.866 2、0.921 8、0.932 9 和 0.927 3。同时,在卷积单元处结合 5 个特征激活模块,所得到的 MIoU、Precision、Recall 和  $F_1$  分数值最优。

## 参考文献

- [1] 耿国华,冯龙,李康,等. 秦陵文物数字化及虚拟复原研究综述[J]. 西北大学学报(自然科学版), 2021, 51(5): 710-721.  
GENG G H, FENG L, LI K, et al. A literature review on the digitization and virtual restoration of cultural relics in the Mausoleum of Emperor Qinshihuang[J]. Journal of Northwest University (Natural Science Edition), 2021, 51(5): 710-721.
- [2] 李春龙,周明全,成欣,等. 轴对称破碎文物的虚拟复原方法[J]. 计算机辅助设计与图形学学报, 2006, 18(5): 620-624.  
LI C L, ZHOU M Q, GENG X, et al. Virtual restoration of axisymmetric relic fragments[J]. Journal of Computer-Aided Design and Computer Graphics, 2006, 18(5): 620-624.
- [3] 王飘,耿国华,张雨禾. 基于表面纹理特征定义的碎片拼接方法[J]. 激光与光电子学进展, 2018, 55(8): 81012.  
WANG P, GENG G H, ZHANG Y H. Fragment splicing method based on surface texture characteristic[J]. Laser & Optoelectronics Progress, 2018, 55(8): 081012.
- [4] 张雨禾,耿国华,魏潇然,等. 基于形状骨架图匹配的文物碎片自动重组方法[J]. 自动化学报, 2017, 43(4): 622-633.  
ZHANG Y H, GENG G H, WEI X R, et al. Reassembly of fractured fragments based on skeleton graphs matching[J]. Journal of Automation, 2017, 43(4): 622-633.
- [5] LI R X, GENG G H, WANG X Z, et al. LBCapsNet: A lightweight balanced capsule framework for image classification of porcelain fragments[J]. Heritage Science, 2024, 12(1): 1-14.
- [6] 律海明. 浅析古陶瓷的几种鉴定方法[J]. 文物鉴定与鉴赏, 2014(4): 90-94.  
LYU H M. A brief analysis of several identification methods for ancient ceramics[J]. Identification and Appreciation to Cultural Relics, 2014(4): 90-94.
- [7] 汤辉. 标准化技术在古陶瓷眼学鉴定中的应用[J]. 景德镇陶瓷, 2020(5): 8-11.  
TANG H. The application of standardization techniques in visual identification of ancient porcelain[J]. Jingdezhen Ceramics, 2020(5): 8-11.
- [8] 金信苗. 微观痕迹在古陶瓷鉴定中的有效性探索[J]. 科技风, 2022(32): 145-147.  
JIN X M. Exploration of the effectiveness of microscopic traces in the authentication of ancient porcelain[J]. Technology Wind, 2022(32): 145-147.
- [9] 王克刚. 基于数字图像特征的古瓷片分类研究[D]. 西安: 西北大学, 2009.
- [10] 鱼跃华. 破损文物碎片的深度学习分类方法研究[D]. 西安: 西北大学, 2022.
- [11] 陆正杰,李纯辉,耿国华,等. 基于多特征描述子自适应权重的文物碎片分类[J]. 激光与光电子学进展, 2020, 57(4): 321-329.  
LU Z J, LI C H, GENG G H, et al. Classification of cultural fragments based on adaptive weights of multi-feature descriptions[J]. Laser & Optoelectronics Progress, 2020, 57(4): 321-329.
- [12] 耿国华,薛米妍,周蓬勃,等. 基于对比学习与多尺度结合的陶瓷显微图像分类方法[J]. 西北大学学报(自然科学版), 2021, 51(5): 734-741.  
GENG G H, XUE M Y, ZHOU P B, et al. Ceramic microscopic image classification based on the combination of contrastive learning and multi-scale methods[J]. Journal of Northwest University (Natural Science Edition), 2021, 51(5): 734-741.
- [13] 刘景怡. 基于显微特征的古代瓷器碎片分类方法的研究与实现[D]. 西安: 西北大学, 2021.
- [14] 张佳宏. 论古瓷器的气泡[J]. 收藏界, 2012(12): 118-119.  
ZHANG J H. On the bubbles of ancient porcelain[J]. Collection World, 2012(12): 118-119.
- [15] 李庆武,王珺,盛惠兴,等. 基于小波包分析的陶瓷显微图像处理[C]//第十二届全国信号处理学术年会论文集. 苏州, 中国: 中国电子学会信号处理分会, 2005: 321-324.
- [16] 朱顺龙. 古陶瓷显微科技气泡测定研究初探: 以越窑青瓷研究为例[J]. 中国文物科学研究, 2008(1): 38-41.  
ZHU S L. Preliminary exploration of microscopic bubble measurement in ancient ceramics: A case study of Yue kiln celadon porcelain[J]. China Cultural Heritage Scientific Research, 2008(1): 38-41.
- [17] 刘国高. 基于快速曲波变换的陶瓷显微图像处理[J]. 中国陶瓷工业, 2008, 15(2): 17-21.  
LIU G G. Ceramic microscopic image procession based on fast curvelet transform[J]. China Ceramic Industry, 2008, 15(2): 17-21.
- [18] LONG J, SHELHAMER E, DARREL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [19] RONNEBERGER O, FISCHER P, BROX T. U-Net;

Convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer International Publishing, 2015:234-241.

[20] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHS N, et al. UNet ++ : A nested U-Net architecture for medical image segmentation[J]. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support, 2018, 11045:3-11.

[21] SU L Z, XIE Q Y, ZHAO F J, et al. Change detection for multispectral images using modified semantic segmentation network[J]. Journal of Applied Remote Sensing, 2022, 16(1):014518.

[22] ARDIYANTO I. Edge devices-oriented surface defect segmentation by ghostNet Fusion block and global auxiliary layer[J]. Journal of Real-Time Image Processing, 2024, 21(1): 13.

[23] LIU X D, WANG L L, HAN X G. Transformer with peak suppression and knowledge guidance for fine-grained image recognition[J]. Neurocomputing, 2022, 492: 137-149.

[24] WANG X L, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake, USA: IEEE, 2018: 7794-7803.

[25] CHEN L C, YANG Y, WANG J, et al. Attention to scale: Scale-aware semantic image segmentation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 3640-3649.

[26] LUONG M T, PHAM H, MANNING C D. Effective approaches to attention-based neural machine translation[EB/OL] (2015-08-17) [2024-02-14]. <https://arxiv.org/abs/1508.04025>.

[27] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[C]//3rd International Conference on Learning Representations. San Diego, USA: ICLR, 2015: 1-15.

[28] BRITZ D, GOLDIE A, LUONG M T, et al. Massive exploration of neural machine translation architectures[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark: USAACL, 2017: 1442.

[29] MINAE 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.

[30] 苏虎,张家斌,张博豪,等. 基于视觉感知的表面缺陷检测综述[J]. 计算机集成制造系统, 2023, 29(1):169-191.

SU H, ZHANG J B, ZHANG B H, et al. A review of research on the inspection of surface defects based on visual perception[J]. Computer Integrated Manufacturing Systems, 2023, 29(1):169-191.

(编辑 李波)