

# 基于改进 U-Net 的活性污泥显微图像分割方法

赵立杰, 路星奎, 陈斌\*

沈阳化工大学信息工程学院, 辽宁 沈阳 110020

**摘要** 针对活性污泥相差显微图像存在伪影且现有图像分割方法对丝状菌的分割精度不高等问题, 提出一种以 U-Net 网络为基础, 结合残差网络、通道注意力机制、空洞空间金字塔模块的活性污泥显微图像分割模型。使用带有通道注意力机制的 ResNet 网络作为编码器, 通道注意力机制显式地建立了特征通道间的依赖关系, 分析了残差网络强化模型的特征提取能力。通过在编码器的最后加入空洞空间金字塔池化, 可在不增加参数数量的同时获得丝状菌与絮体的多尺度信息。通过在解码器中使用跳跃连接来补充特征信息, 强化网络的重建能力。实验结果表明, 与 U-Net、DeepLabV3+ 等算法相比, 所提模型具有更好的分割性能与效果。

**关键词** 图像处理; 活性污泥; 图像分割; 空洞空间金字塔池化; 注意力机制; ResNet

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.1210008

## Activated Sludge Microscopic Image Segmentation Method Based on Improved U-Net

Zhao Lijie, Lu Xingkui, Chen Bin\*

School of Information Engineering, Shenyang University of Chemical Technology, Shenyang, Liaoning 110020, China

**Abstract** Aiming at the problems of artifact existing in activated sludge phase contrast microscopic images and low segmentation accuracy of existing image segmentation methods for filamentous bacteria, a segmentation model of an activated sludge microscopic image based on the U-Net network, residual network, channel attention mechanism, and atrous spatial pyramid module is proposed. The ResNet network based on channel attention mechanism is used as an encoder. Channel attention mechanism explicitly establishes the dependence among feature channels, and analyzes the feature extraction ability of the residual network reinforcement model. At the end of the encoder, the atrous spatial pyramid pooling is added, which can obtain the multi-scale information of filaments and flocs without increasing the parameters. In order to enhance the ability of network reconstruction, the feature information is supplemented by using jump connection in the decoder. Experimental results show that the proposed model has better segmentation performance and effect than U-Net and DeepLabV3+.

**Key words** image processing; activated sludge; image segmentation; atrous spatial pyramid pooling; attention mechanism; ResNet

**OCIS codes** 100.3008; 100.4996; 200.4260

## 1 引言

目前, 活性污泥法是污水处理厂应用最广泛的污水生化处理工艺<sup>[1]</sup>。污泥膨胀<sup>[2]</sup>是活性污泥污水

处理厂最棘手和最普遍的问题。一般使用污泥质量分数 (MLSS)、沉降速率 (SV)、污泥容积指数 (SVI) 等指标衡量活性污泥的数量和性能好坏<sup>[3]</sup>。以上理化指标的测定成本较高、操作繁琐且周期较长。随

收稿日期: 2020-08-27; 修回日期: 2020-09-24; 录用日期: 2020-10-14

基金项目: 国家重点研发计划(2018YFB1700200)、2020 年辽宁省教育厅创新人才支持计划

\* E-mail: chenbin79056300@163.com

着数字图像处理和显微镜技术的发展,基于数字图像的活性污泥图像处理和分析是监测污水处理厂微生物状态和污泥膨胀的有效潜在手段。

活性污泥相差显微图像中絮体和丝状菌的正确分割是提取活性污泥形态学参数的关键。活性污泥相差显微图像分割主要包括阈值分割法<sup>[4]</sup>、基于梯度边缘的阈值处理方法<sup>[5]</sup>、基于方差算子和拉普拉斯算子的分割方法<sup>[6]</sup>、基于区域的分割方法<sup>[7]</sup>,以及基于分水岭算法的分割方法<sup>[8]</sup>。Perez 等<sup>[9]</sup>提出按照背景校正、直方图均衡、中值滤波和形态学操作的顺序进行絮体分割。Khan 等<sup>[10]</sup>提出使用 Otsu 阈值法、K-means 和模糊 C-means 分割法分割和检测活性污泥显微图像中的絮体。Ang 等<sup>[11]</sup>提出基于相位拉伸变换的相衬图像细丝分割算法。Nisar 等<sup>[12]</sup>提出采用基于高斯混合模型的图像处理算法分割废水显微图像。传统的图像分割方法对于参数敏感,对复杂环境下絮体与丝状菌难以得到精准的分割结果。

近年来,因卷积神经网络的迅速发展,部分学者将卷积神经网络应用到生物与医学方向的图像分割任务中。Shelhamer 等<sup>[13]</sup>提出全卷积神经网络之后,更多的学者提出效果更佳的分割网络模型,如 U-Net<sup>[14]</sup>、DeepLabV3+<sup>[15]</sup>、hourglass<sup>[16]</sup>等。李大湘等<sup>[17]</sup>提出使用 Inception 模块与注意力机制改进的 U-Net 结构分割视网膜血管图像。艾玲梅等<sup>[18]</sup>通过在 U-Net 结构中引入注意力机制与残差紧密模块分割脑肿瘤磁共振图像。张文秀等<sup>[19]</sup>为解决相衬显微镜细胞图像亮度不均匀且目标与背景的对比度较低等问题,提出使用残差块与注意力机制分割细胞图像的方法。目前使用基于深度学习的方式进行活性污泥图像丝状菌分割的案例较少,Boztoprak 等<sup>[20]</sup>使用 Levenberg-MarQuardt 算法、遗传算法、人工蜂群算法训练神经网络的权重,首次使用卷积神经网络对絮体与丝状菌进行分割,获得较精准的形态学参数。Zhao 等<sup>[21]</sup>提出加权二元交叉熵函数和 Dice 系数组合的损失函数,使用双 U-Net 结构实现实验室规模活性污泥显微图像中絮体与丝状菌的自动分割。与 U-Net 模型、FCN-VGG16 分割模型和传统分割方法相比,该方法能够克服相差显微图像光晕和阴影等伪影现象,在一定程度上提高对絮体与丝状菌的分割准确率,但双模型结构存在模型规模大、计算复杂等问题。

针对现有图像分割方法对活性污泥相差显微图像中丝状菌的分割精度不高且存在光晕等伪影的问

题,提出基于 U-Net 网络结构的活性污泥显微图像分割方法,使用带有通道注意力机制<sup>[22]</sup>的 ResNet<sup>[23]</sup>网络提取图像特征,其中通道注意力机制显式地建立通道间的依赖关系,赋予各特征通道不同的权重系数,利用 ResNet 网络中的残差结构强化特征能力,缓解梯度消失等问题;最后使用空洞空间金字塔池化<sup>[24]</sup>,在不增加额外参数量的同时扩大感受野,获得丝状菌与絮体的多尺度信息以及图像级特征。

## 2 基于改进 U-Net 的活性污泥显微图像分割方法

### 2.1 模型框架

针对活性污泥显微图像中丝状菌细长、絮体碎片多、显微镜本身的伪影等因素导致的现有图像分割方法分割精度低的问题,提出基于 U-Net 网络结构的活性污泥显微图像分割方法。基于 U-Net 网络结构的活性污泥显微图像分割算法模型如图 1 所示。该方法将注意力机制引入到 ResNet 网络,建立通道间的依赖关系,根据各特征通道的重要程度生成不同的权重系数。带注意力机制的残差模块 (ARes) 能够增强网络的特征提取能力;使用空洞空间金字塔池化 (ASPP) 捕获不同感受野下絮体、丝状菌的特征信息,在不增加额外参数量的同时扩大感受野,这可使网络保留更多图像的细节特征。在解码阶段,使用跳跃连接将编码器中保存的各阶段特征图与解码器的输入相连,补充特征信息并强化网络的重建能力。

### 2.2 ResNet 中的残差结构

使用 ResNet 网络作为编码器,从活性污泥显微图像中提取更多的细节信息,其中残差结构有效地缓解随着网络结构加深引起的梯度消失、网络退化等问题。残差结构可表示为

$$y = x_{l+1} = F(x_l, \mathbf{W}) + x_l, \quad (1)$$

$$y = x_{l+1} = F(x_l, \mathbf{W}) + \mathbf{w}_s x_l, \quad (2)$$

式中: $y$  与  $x_{l+1}$  为第  $l$  层网络的输出; $x_l$  为第  $l$  层网络的输入; $\mathbf{W}$  为第  $l$  层的权值矩阵; $F$  为残差函数; $\mathbf{w}_s$  为平方矩阵。如果  $F(x, \mathbf{W})$  与输入  $x_l$  的维度不匹配,则使用  $\mathbf{w}_s$  进行线性映射,完成维度匹配。

ResNet 的下采样操作是在残差结构中使用步长为 2 的  $1 \times 1$  卷积实现的。输入数据经过步长为 2 的  $1 \times 1$  卷积后,特征图的尺寸减小为原来的一半。步长为 2 的  $1 \times 1$  卷积会导致  $3/4$  的输入特征

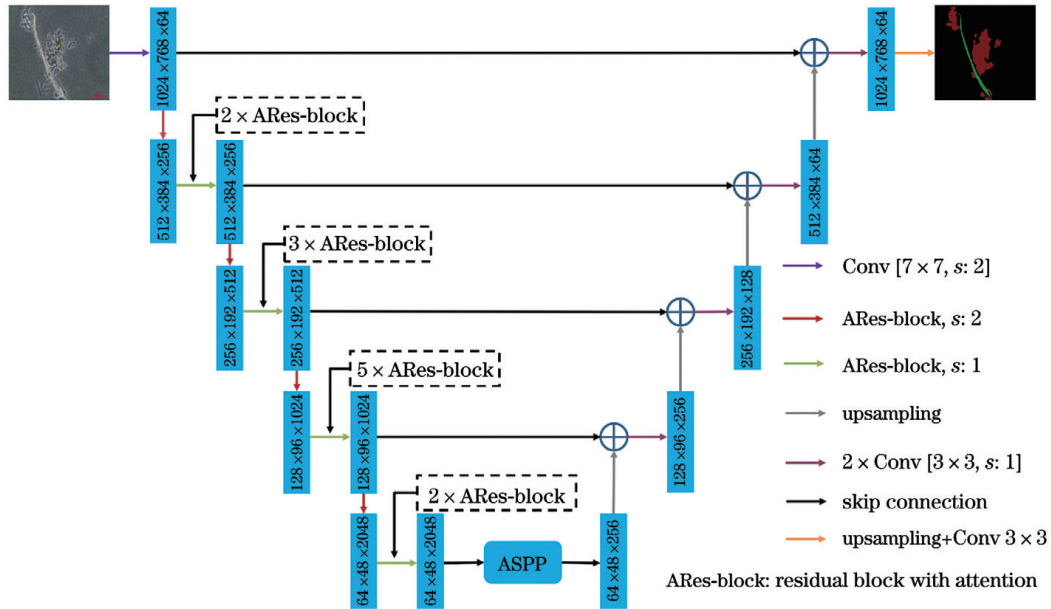


图 1 改进 U-Net 的网络结构图

Fig. 1 Network architecture of improved U-Net

不被利用<sup>[25]</sup>。文献[25]在卷积层前增加一个步长为 2 的平均池化层并将卷积步长修改为 1, 该方法在实际应用中的效果更好, 因此选择图 2 所示的残差结构。

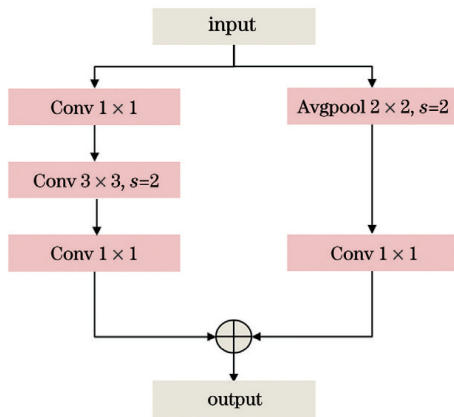


图 2 残差块结构图

Fig. 2 Architecture of residual block

### 2.3 通道注意力机制

考虑到活性污泥显微图像中存在伪影, 通过使用通道注意力机制捕获重要的语义信息, 可减小无关特征的影响, 使模型学习到真正有效的特征信息。

通道注意力机制通过输入特征判断各特征通道的重要程度, 自动生成通道注意力权重系数以作用于各特征通道。带有通道注意力机制的残差结构如图 3 所示, 首先对输入特征进行全局平均池化, 完成压缩过程, 再经过 2 个全连接层 (FC) 完成激励过程, 最后通过 sigmoid 激活函数生成 0 到 1 之间的

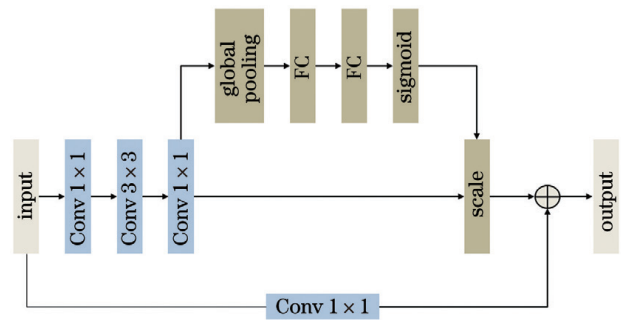


图 3 带有通道注意力机制的残差模块

Fig. 3 Residual block with channel attention mechanism

通道注意力权重系数, 以作用于输入的特征。压缩与激励可分别表示为

$$z_c = F_{sq}(U_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H u_c(i, j), \quad (3)$$

$$s = F_{ex}(Z, W) = \sigma \{g [W_2 \delta(W_1, Z)]\}, \quad (4)$$

式中:  $z_c$  与  $F_{sq}(U_c)$  表示特征通道  $c \in (1, C)$  的特征平均值;  $U_c$  表示特征通道  $c$  上的全部特征值;  $W$  与  $H$  分别表示特征图的宽和高;  $s$  与  $F_{ex}(Z, W)$  表示激励结果;  $Z$  为压缩结果;  $W_1$  和  $W_2$  表示 2 个全连接层的权值矩阵;  $\sigma$  表示 sigmoid 激活函数;  $g$  表示 ReLU 激活函数。

### 2.4 空洞空间金字塔池化

在卷积神经网络中, 大的卷积核代表大的感受野, 也代表大的参数量。因此, 为在不增加额外参数的同时获得更大的感受野, 使用带有空洞的卷积核即空洞卷积代替常规卷积。

使用 ASPP 获取丝状菌与絮体的多尺度信息, 提取不同感受野下活性污泥图像的细节特征。ASPP 结构如图 4 所示, 包括卷积核的大小为  $3 \times 3$ 、空洞率分别为 2, 3, 5 的空洞卷积, 全局平均池化, 以及  $1 \times 1$  卷积。首先, 针对输入特征使用全局平均池化、 $1 \times 1$  卷积获得图像级特征, 使用双线性插值将

图像恢复至输入大小。然后分别对输入特征进行  $1 \times 1$  卷积、不同空洞率的空洞卷积。最后拼接各输出特征, 使用  $1 \times 1$  卷积整合信息, 同时完成降维处理。通过 ASPP 结构可以学习到不同感受野下活性污泥显微图像中丝状菌与絮体的特征信息, 使最后的特征重建更为准确。

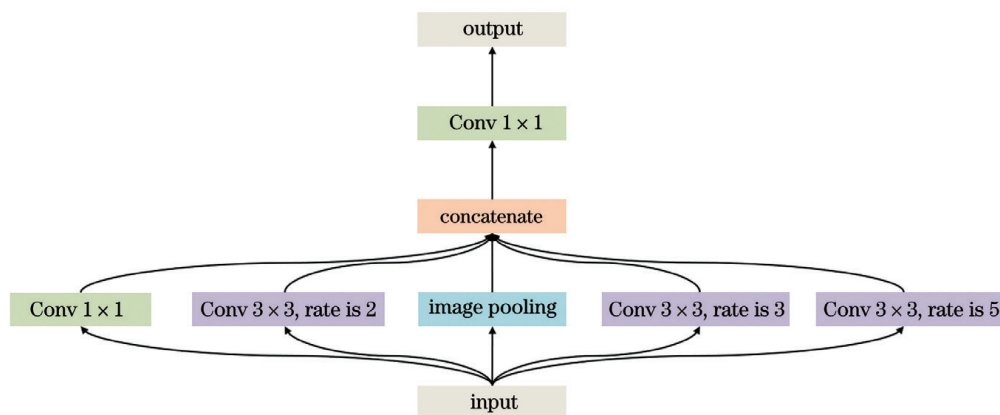


图 4 空洞空间金字塔池化结构图

Fig. 4 Structure of atrous spatial pyramid pooling

### 3 实 验

#### 3.1 活性污泥显微图像数据集

活性污泥样本来自沈阳某污水处理厂曝气池。污泥图像采集系统使用配有彩色 CCD 相机的相差显微镜, CCD 相机输出  $2048 \times 1536$  RGB 彩色图像。图像采集系统如图 5 所示。



图 5 带数码相机的光学显微镜

Fig. 5 Optical microscope with digital camera

2019–2020 年共采集 5472 张活性污泥相差显微图像, 选取 320 张典型图像样本进行丝状菌和絮体标注。为了测试分割模型的性能, 将 320 张图像作为数据集, 其中 257 张图像作为训练集, 63 张作为测试集。活性污泥显微图像中包含的三种目标分别为絮体、丝状菌、背景。图 6 为活性污泥相差显微图像 (PCM), 其中图 6(a) 为原始图像, 图 6(b) 为标注后的图像。

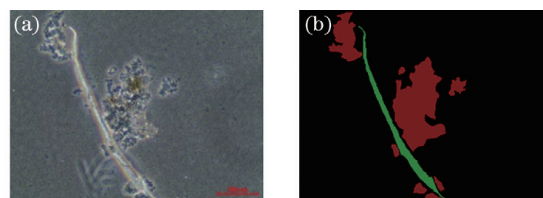


图 6 PCM 图像。(a) 原始图像; (b) 标注的图像

Fig. 6 PCM images. (a) Original image; (b) labeled image

#### 3.2 数据增强

使用 Rich Crop 数据增强策略对图像进行增强处理, 以保证训练过程中数据的多样性, 从而使模型具有更好的鲁棒性。Rich Crop 数据增强策略的参数设置如下: 以 0.2 的最大概率进行高斯模糊, 以 0.5 的概率随机翻转, 最大旋转角度设置为  $90^\circ$ , 剪裁图像面积与原图面积比为 0.5, 宽高比为 0.5, 亮度、饱和度、对比度的调节因子均设置为 0.5, 数据增强样例如图 7 所示。

#### 3.3 评价指标

本文使用交并比 (IoU, 可用  $R_{IoU}$  表示)、精确度 ( $P$ )、召回率 ( $R$ ) 3 个评价指标作为判断图像分割模型好坏的标准。

1) 交并比为模型对前景的预测结果与真实结果的交集与并集之比, 通常被作为图像分割领域的标准度量, 可表示为

$$R_{IoU} = \frac{T_p}{T_p + F_p + F_n}, \quad (5)$$



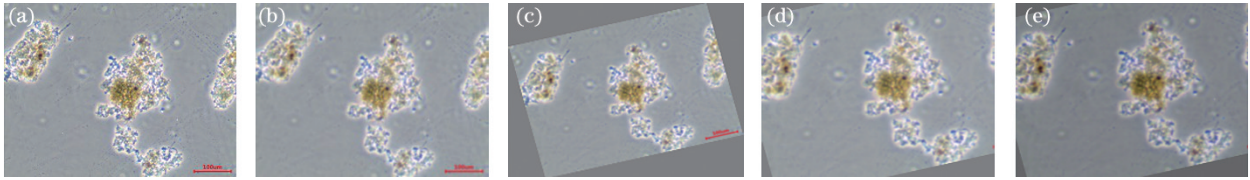


图 7 数据增强结果。(a)训练图像;(b)图像模糊效果图;(c)图像旋转效果图;(d)图像长宽比调整效果图;(e)图像颜色空间扰动效果图

Fig. 7 Data augmentation results. (a) Training image; (b) image with blur effect; (c) image with rotation effect; (d) adjustment effect of image aspect ratio; (e) image with color space perturbation effect

式中: $T_P$  为真实类别为  $c$  且预测为  $c$  的样本数; $F_P$  为真实类别不为  $c$  却预测为  $c$  类的样本数; $F_N$  为真实类别为  $c$  却预测为其他类别的样本数。

2) 精确度是模型对前景预测正确的结果与模型全部预测结果的比值,也称查准率,可表示为

$$P = \frac{T_P}{T_P + F_P} \quad (6)$$

3) 召回率是模型对前景预测正确的结果与真实样本量的比值,可表示为

$$R = \frac{T_P}{T_P + F_N} \quad (7)$$

### 3.4 模型训练

本文算法基于 PaddlePaddle 框架,实验硬件配置为 32 GB 内存的 CPU,显存为 32 GB 的 Tesla V100 显卡,操作系统为 Ubuntu16.04。选择 SGD 优化器进行模型参数优化;学习率呈余弦式衰减,其初始值为 0.001;批量大小为 8;训练 Epoch 为 200。引入早停机制,若连续 5 次迭代絮体和丝状菌的 IoU 不低于 0.92 和 0.74,则停止训练,以防止模型出现过拟合现象。图 8 分别展示了训练时损失值、准确率、交并比随迭代次数的变化曲线。

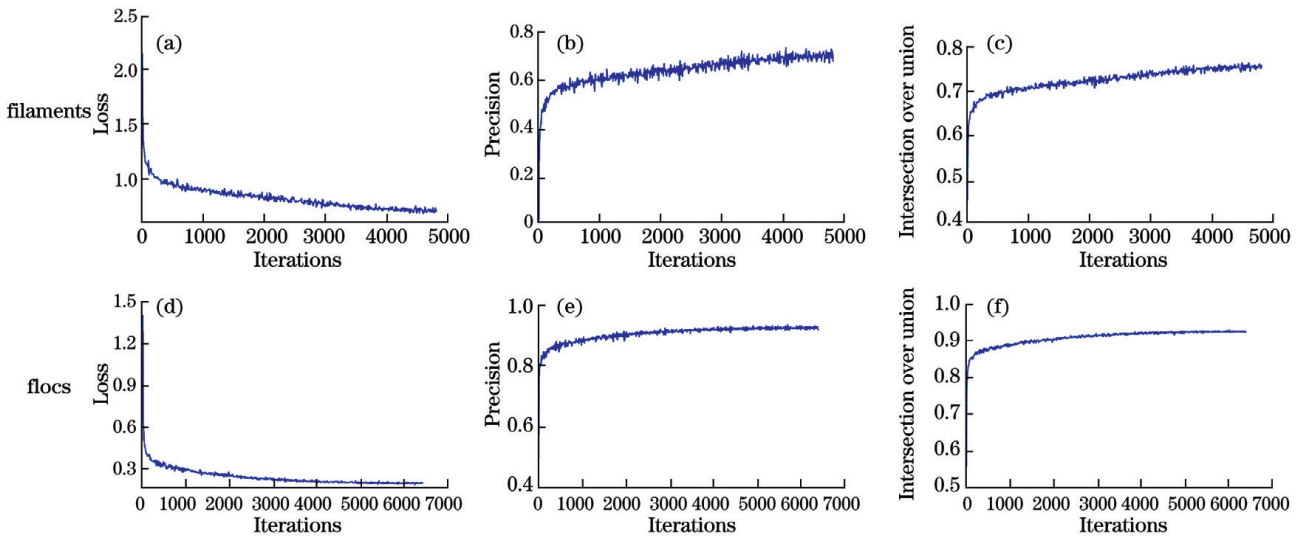


图 8 训练时损失值、准确率、交并比随迭代次数的变化曲线。(a)(d)损失值;(b)(e)准确率;(c)(f)交并比

Fig. 8 Curves of loss, accuracy, and intersection over union varying with iterations. (a)(d) Loss; (b)(e) accuracy; (c)(f) intersection over union

考虑到数据集中的背景与丝状菌、背景与絮体的像素占比差异较大,为提升模型稳定性,本次实验将二元交叉熵损失函数( $L_{bce}$ )与 Dice 系数损失函数( $L_{Dice}$ )组合使用。

二元交叉熵损失函数仅适用于二分类,可表示为

$$L_{bce} = -y \log \hat{y} - (1 - y) \log(1 - \hat{y}), \quad (8)$$

式中: $y$  表示正确结果; $\hat{y}$  表示预测结果。

Dice 系数损失函数是利用 Dice 系数作为损失监督网络,以解决样本不均衡的问题,加速模型收敛。Dice 系数损失函数可表示为

$$L_{Dice} = 1 - \frac{2|y \cap \hat{y}|}{|y| + |\hat{y}|} \quad (9)$$

### 3.5 实验结果与分析

分别使用 DeepLabv3+、U-Net 以及本文算法进行分割,分割结果如图 9 所示。图 9 中从左至右

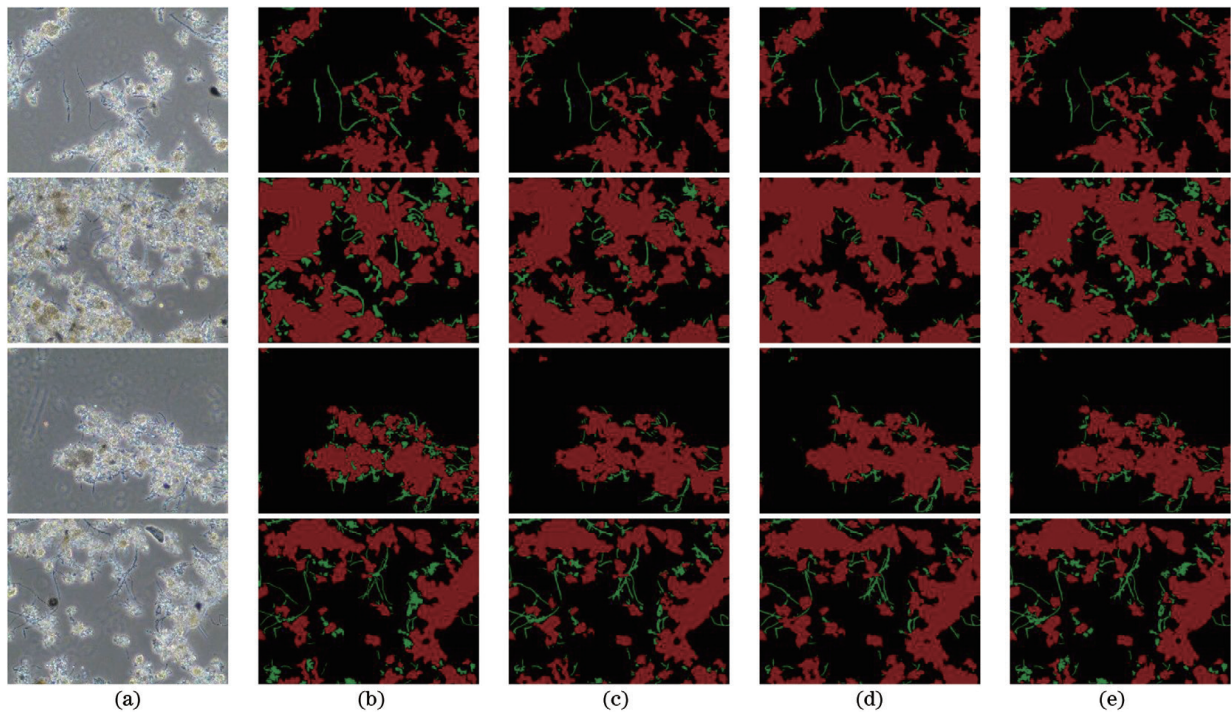


图 9 不同算法分割结果。(a)测试图像;(b)标签图像;(c) DeepLabv3+分割结果;(d) U-Net 分割结果;(e)本文算法分割结果

Fig. 9 Segmentation results of different algorithms. (a) Test images; (b) labeled images; (c) segmentation results of DeepLabv3+; (d) segmentation results of U-Net; (e) segmentation results of proposed algorithm

分别为测试图像、标签图像、DeepLabv3+ 分割结果、U-Net 分割结果以及本文算法分割结果。从整体来看,各模型对于较大的丝状菌与絮体均有较好的分割效果,本文算法对于边缘恢复具有更好的效果。

测试集分割结果的评价指标如表 1 所示。从精确度指标来看,本文方法相对 U-Net、DeepLabv3+ 有所提升。从交并比指标来看,本文算法相较于 U-Net、DeepLabv3+ 在絮体的分割指标上分别有

表 1 评价指标

Table 1 Evaluation indexes

Method	Target	$P$	$R$	$R_{IoU}$
U-Net	Floc	0.8085	0.9720	0.7898
	Filamentous	0.5115	0.7460	0.4355
	Mean	0.6600	0.8590	0.6172
DeepLabv3+	Floc	0.8767	0.9023	0.8007
	Filamentous	0.5974	0.6628	0.4581
	Mean	0.7371	0.7826	0.6296
Our algorithm	Floc	0.8836	0.9280	0.8269
	Filamentous	0.6379	0.6850	0.4932
	Mean	0.7608	0.8065	0.6601

0.0371、0.0262 的提升,在丝状菌的分割指标上也有一定程度的提高,分别为 0.0577、0.0351,这充分体现通道注意力机制与空洞空间金字塔池化在分割任务上的优势。综合来看,本文算法较 U-Net、DeepLabv3+ 有更好的分割指标。

虽然本文算法的分割效果较准确,但是其针对丝状菌的分割结果与絮体相比存在明显的差距,这说明活性污泥显微图像虽然类别数量不多,但是对其进行分割的难度较大。其原因可能是:1)丝状菌的样本量相对絮体较少,存在样本不均衡现象;2)网络下采样导致小目标的语义信息丢失。因此后续工作需要重新思考模型的结构特点与数据的关系以提高分割对象样本间的均衡性。

所提模型的规模大小介于 U-Net 与 DeepLabv3+ 之间。虽然所提模型的评价指标均高于其他模型,但却带来额外的参数量,因此下一步的目标是保证分割指标的同时对模型压缩进行优化。

为验证各模块对 U-Net 网络的性能提升,使用相同的数据分别进行模型训练,通过测试得到表 2 所示的各模块分割指标。U-Net 网络的 IoU 为 0.6172;使用带有通道注意力机制的 ResNet 之后, IoU 提升 0.0370,这证明了残差结构与通道注意力机制的有效性;在 U-Net 结构中添加 ASPP 结构,

表 2 各模块对整体模型的影响

Table 2 Influence of each module on whole performance

Method	IoU
U-Net	0.6172
U-Net+ASPP	0.6263
U-Net+ResNet+attention	0.6542
U-Net+ResNet+attention+ASPP	<b>0.6601</b>

得到的 IoU 提升 0.0091,这说明 U-Net 的感受野包括大部分图像中的目标,仅少部分的大目标需要在空洞空间金字塔池化中获得更大的感受野,以提取特征,因此在 U-Net 结构中添加 ASPP 结构时性能提升相对有限。使用全部模块搭建的网络相比 U-Net, IoU 提升 0.0429,这充分体现了本文所提算法的有效性。

## 4 结 论

针对活性污泥显微图像中丝状菌细长、絮体碎片多、显微镜存在伪影等因素导致图像分割精度低的问题,提出一种基于改进 U-Net 网络结构的活性污泥显微图像分割方法。该方法将 ResNet、通道注意力机制、多空洞空间金字塔池化与 U-Net 结构相结合,在特征提取阶段加入通道注意力机制与空间金字塔池化,使模型更关注重要的特征通道,获得了丝状菌与絮体在不同感受野下的特征信息,从而使特征重建更为准确。实验结果表明,相较于经典的 U-Net、DeepLabv3+ 算法,所提模型对活性污泥显微图像具有较高的分割精确度和交并比。活性污泥显微图像分割中,所提模型对丝状菌的分割效果低于絮体,其原因可能是编码器-解码器结构中多次下采样操作引起丝状菌特征丢失。后续将侧重研究提升丝状菌的分割效果的方法。

## 参 考 文 献

- [1] Zhang Q Z, Xu Y T, Zhou Y, et al. Wastewater treatment and water environment protection [J]. *Water Purification Technology*, 2010, 29(4): 1-4. 张秋卓, 徐亚同, 周扬, 等. 废水处理与水环境保护 [J]. *净水技术*, 2010, 29(4): 1-4.
- [2] Molina M A, Pérez C A A, Leiva C A. Characterization of filamentous flocs to predict sedimentation parameters using image analysis [J]. *Journal of Sensors*, 2020, 2020: 1-8.
- [3] Mesquita D P, Amaral A L, Ferreira E C. Identifying different types of bulking in an activated

sludge system through quantitative image analysis [J]. *Chemosphere*, 2011, 85(4): 643-652.

- [4] Jenné R, Banadda E N, Philips N, et al. Image analysis as a monitoring tool for activated sludge properties in lab-scale installations [J]. *Journal of Environmental Science and Health, Part A*, 2003, 38(10): 2009-2018.
- [5] Heine W, Sekoulov I, Burkhardt H, et al. Early warning-system for operation-failures in biological stages of WWTPs by on-line image analysis [J]. *Water Science and Technology: a Journal of the International Association on Water Pollution Research*, 2002, 46(4/5): 117-124.
- [6] Sikora M, Smolka B. Feature analysis of activated sludge based on microscopic images [C] // *Canadian Conference on Electrical and Computer Engineering 2001 Conference Proceedings (Cat. No. 01TH8555)*, May 13-16, 2001, Toronto, ON, Canada. New York: IEEE Press, 2001: 1309-1314.
- [7] Lee X Y, Khan M B, Nisar H, et al. Morphological analysis of activated sludge flocs and filaments [C] // *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, May 12-15, 2014, Montevideo, Uruguay. New York: IEEE Press, 2014: 1449-1453.
- [8] Yong L X, Nisar H, Ho Y K. An approach for the segmentation and quantification of activated sludge floc blobs [J]. *Advanced Science Letters*, 2013, 19(5): 1372-1376.
- [9] Perez Y G, Leite S G F, Coelho M A Z. Activated sludge morphology characterization through an image analysis procedure [J]. *Brazilian Journal of Chemical Engineering*, 2006, 23(3): 319-330.
- [10] Khan M B, Nisar H, Aun N C. Segmentation and quantification of activated sludge flocs for wastewater treatment [C] // *2014 IEEE Conference on Open Systems (ICOS)*, October 26-28, 2014, Subang, Malaysia. New York: IEEE Press, 2014: 18-23.
- [11] Ang R B Q, Nisar H, Khan M B, et al. Image segmentation of activated sludge phase contrast images using phase stretch transform [J]. *Microscopy*, 2019, 68(2): 144-158.
- [12] Nisar H, Herng G J, Chiong T P. Image segmentation of bright field activated sludge microscopic images using Gaussian mixture model [C] // *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, November 3-7, 2019, Abu Dhabi, United Arab Emirates. New York: IEEE Press, 2019: 1-7.
- [13] Shelhamer E, Long J, Darrell T. Fully convolutional



- networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [14] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [15] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 833-815.
- [16] Newell A, Yang K Y, Deng J. Stacked hourglass networks for human pose estimation[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9912: 483-499.
- [17] Li D X, Zhang Z. Improved U-net segmentation algorithm for the retinal blood vessel images [J]. Acta Optica Sinica, 2020, 40(10): 1010001.  
李大湘, 张振. 基于改进 U-Net 视网膜血管图像分割算法[J]. 光学学报, 2020, 40(10): 1010001.
- [18] Ai L M, Li T D, Liao F Y, et al. Magnetic resonance brain tumor image segmentation based on attention U-net[J]. Laser & Optoelectronics Progress, 2020, 57(14): 141030.  
艾玲梅, 李天东, 廖福元, 等. 基于注意力 U-Net 的脑肿瘤磁共振图像分割[J]. 激光与光电子学进展, 2020, 57(14): 141030.
- [19] Zhang W X, Zhu Z C, Zhang Y H, et al. Cell image segmentation method based on residual block and attention mechanism [J]. Acta Optica Sinica, 2020, 40(17): 1710001.  
张文秀, 朱振才, 张永合, 等. 基于残差块和注意力机制的细胞图像分割方法[J]. 光学学报, 2020, 40(17): 1710001.
- [20] Boztoprak H, Özbay Y, Güçlü D, et al. Prediction of sludge volume index bulking using image analysis and neural network at a full-scale activated sludge plant [J]. Desalination and Water Treatment, 2016, 57(37): 17195-17205.
- [21] Zhao L J, Zou S D, Zhang Y H, et al. Segmentation of activated sludge phase contrast microscopy images using U-net deep learning model [J]. Sensors and Materials, 2019, 31(6): 2013-2028.
- [22] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [23] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [24] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 833-851.
- [25] He T, Zhang Z, Zhang H, et al. Bag of tricks for image classification with convolutional neural networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 558-567.