

酵母自体荧光寿命显微图像的自动分析方法

钟佳慧¹, 伍君鑫², 孔亚伟², 苏文华², 马炯^{1,2**}, 糜岚^{1,2*}

¹复旦大学工程与应用技术研究院生物医学工程技术研究所, 上海 200433;

²复旦大学信息科学与工程学院光科学与工程系, 上海 200433

摘要 酿酒酵母是最具有吸引力的微生物之一, 监测其不同生长时期的新陈代谢状态变化对基础生物学和工业研究都具有重要意义。依据酵母的生成曲线规律培养了不同时间的酵母细胞, 基于荧光寿命显微成像 (FLIM) 进行自体荧光寿命图像的采集, 并提出了一种基于机器学习的自动分析方法, 可无标记快速鉴别年轻和衰老的酵母细胞。首先, 采用深度监督 U-Net 实现酵母细胞的自动分割; 然后, 提取每个酵母细胞的荧光寿命特征和形态特征; 最后, 采用无监督聚类方法实现分类。实验结果表明, 酵母的衰老伴随着新陈代谢的变化。FLIM 作为一种无标记成像技术可应用于酵母细胞的代谢分析中, 结合自动化分析流程可快速准确地区分具有不同代谢差异的细胞, 为后续单细胞的筛选奠定了基础。

关键词 生物技术; 荧光寿命显微成像; 机器学习; 图像处理; 酵母细胞分类

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/LOP202259.0617019

Automated Analysis Methods for Autofluorescence Lifetime Microscopic Images of Yeast

Zhong Jiahui¹, Wu Junxin², Kong Yawei², Su Wenhua², Ma Jiong^{1,2**}, Mi Lan^{1,2*}

¹*Institute of Biomedical Engineering and Technology, Academy for Engineer and Technology, Fudan University, Shanghai 200433, China;*

²*Department of Optical Science and Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China*

Abstract *Saccharomyces cerevisiae* is one of the most attractive microorganisms, and monitoring changes in its metabolic state at different growth periods is significant for both basic biology and industrial research. In this paper, yeast cells are cultured for different periods based on the yeast generation curve rule, autofluorescence lifetime images are collected using fluorescence lifetime imaging microscopy (FLIM), and an automatic analysis method based on machine learning is proposed, which can rapidly identify young and senile yeast cells without markers. First, a deep-supervised U-Net is applied to automatically segment yeast cells. Then, the features of fluorescence lifetime and morphology of each yeast cell are extracted. Finally, the classification is achieved using the unsupervised clustering method. The experimental results reveal that yeast senescence is accompanied by changes in metabolism. FLIM, as a label-free imaging technique, can be used for the metabolic analysis of yeast cells. When combined with the automated analysis process, it can swiftly and accurately distinguish cells with different metabolic

收稿日期: 2021-11-22; 修回日期: 2021-12-24; 录用日期: 2021-12-31

基金项目: 上海市自然科学基金(20ZR1405100)、上海市卫生健康委员会重点学科建设计划(2020-2022)(GWV-10.1-XK01)、上海毛发医学工程技术研究中心(19DZ2250500)、复旦大学医工结合项目(yg2021-022)、复旦大学工研院先导项目(gyy2018-001,gyy2018-002)

通信作者: *lanmi@fudan.edu.cn; **jiongma@fudan.edu.cn

differences, laying the foundation for subsequent screening of single cells.

Key words biotechnology; fluorescence lifetime imaging microscopy; machine learning; image processing; yeast cells classification

1 引言

酿酒酵母(SC)是一种简单且能快速生长的单细胞真核微生物。在工业领域中,SC作为最有价值的物种之一,被大规模应用于食品、饮料和乙醇的生产^[1]。同时,SC也是一种重要的模式生物,在现代细胞生物学和医药研究中都起到至关重要的作用^[2]。酵母与复杂生命形式中的衰老机制相关,也是衰老研究中常用的生物模型,揭示了许多重要现象,如营养传感信号通路在调节衰老中的重要性^[3],帮助人们更深入地了解遗传学和生理学^[4-5],为解决组织或高等生物中的多细胞衰老谜题奠定了基础,对人类医学和寿命延长的医学保健问题具有重要意义^[6]。

常用来检测酵母衰老的方法是通过荧光染色或荧光蛋白标记老化的酵母细胞^[7],这种方法难以实现对细胞的原位研究,且细胞在生物标记后通过流式细胞仪进行分类,其准确率也受到了一定的挑战^[8]。近年来,荧光寿命显微成像(FLIM)已被用于多种细胞和组织的自体荧光成像,在内源性荧光分子中烟酰胺腺嘌呤二核苷酸(磷酸)(NAD(P)H)是参与多种代谢的辅酶,根据其荧光寿命的变化可以检测代谢情况^[9]。因此,通过FLIM观测细胞或组织内的NAD(P)H可应用于癌症诊断^[10]、细胞分裂^[11]以及干细胞分化等方面的研究^[12]。FLIM同样也适用于酵母细胞的自体荧光成像,避免了繁琐的染色标记步骤,同时又能检测酵母新陈代谢的状态,为原位研究提供保障。

针对大量FLIM图像,基于机器学习的自动化分析流程有助于研究人员及时采集状态信息,更好地了解细胞的生理状态变化,快速获得鉴别结果。图像的自动分析主要分为图像处理、特征采集和统计分析三个阶段。图像处理阶段主要是进行细胞分割,已有针对酵母细胞的分割方法大多依赖于传统的图像处理技术,如阈值处理或边缘检测^[13-14],但酵母细胞拥挤、形态不规则、聚焦不完美等常见现象导致分割效果不理想。随着神经网络的应用,基于分割网络(SegNet)^[15],掩模-区域卷积神经网络(Mask-RCNN)^[16]以及U-Net^[17]的分割、目标识别网络结构逐渐被应用于酵母的明场、微分干涉对比

(DIC)显微图像中,并取得了优异的成绩。当输入图像不再是明场或DIC图像而是FLIM图像时,分割方法也需要进行相应的改进。在特征采集和统计分析阶段,已有研究主要对酵母细胞的形态进行定量分析,如横截面积、液泡大小、轴长,以观测酵母细胞经历的变化^[18]。Coelho等^[19]发现细胞的形态与其生理状态密切相关,酵母细胞的大小可作为环境压力因素影响的指标^[20],液泡大小特征也可用于酵母的发酵研究中^[21]。但形态特征的获取建立在酵母分布稀疏且图像质量较高的前提下,当细胞过于密集时,很难进行准确测量,且大多数研究停留在了差异性分析上,没有进行后续的分类处理。

针对SC细胞的自体FLIM图像,本文提出了一种分析方法,依次对细胞进行分割、特征提取、聚类分析,以监测SC细胞在不同培养时间下的新陈代谢状态。在细胞分割阶段,结合基于U-Net的网络结构和深度监督(DS)强化分割效果。相比传统的显微图像,FLIM图像可表征细胞的新陈代谢特征,因此在特征提取阶段,不仅提取了细胞的形态特征,还提取了平均荧光寿命、不同成分的荧光寿命及其占比等数值特征。同时,基于迁移学习的卷积神经网络(CNN)提取了细胞图像的抽象特征。最后,采用无监督聚类分析,根据不同的特征输入分析SC细胞的差异性,实现对SC细胞的快速分类。

2 实验方法

2.1 酵母的培养和测量

用SC细胞作为研究对象,将SC细胞接种在酵母培养基(YPD)中,每瓶YPD(10 mL)包含质量浓度分别为20 g/L、20 g/L、10 g/L的葡萄糖、蛋白胨和酵母提取物。将SC放入摇床(ZHTY-50,知楚仪器,中国)中培养,温度设置为30℃,转速为220 r/min。分别在培养6 h、24 h、72 h后取出100 mL含酵母的培养基,在温度为4℃、相对离心力为1000g(g 为重力加速度)下离心处理3 min后弃上清液,然后用磷酸盐缓冲液(PBS)清洗并离心弃上清液。重复上述去除YPD的过程3次,取约10 μ L富集SC细胞的PBS液体滴于载玻片上,准备进行SC细胞的FLIM测量。

在 SC 发酵过程中,常用光密度(OD)值反映 SC 细胞的生长状态。用分光光度计(V5000,上海元析仪器有限公司,中国)在波长为 600 nm 时,测量 SC 细胞在 0~72 h 中不同时间点的 OD 值(X_{OD}),每个时间点对样品测试 5 个 OD 值并取平均值,用于绘制 SC 细胞的生成曲线。

2.2 FLIM 的测量和数据拟合方法

FLIM 系统是在激光扫描共聚焦显微镜(FV300/IX 71,Olympus,日本)的基础上,结合时间相关单光子计数 TCSPC(SPC-150,Becker & Hickl,德国)信号采集系统搭建而成。激发光源为 405 nm 的皮秒脉冲激光器,用 60 倍物镜(数值孔径 $NA=1.2$)进行观测,采用 417~477 nm 的带通滤光片收集 SC 细胞的 NAD(P)H 荧光,共采集 40 个不同的视野区域,每张 FLIM 图像的尺寸为 256 pixel×256 pixel。

用 SPCImage 软件(Becker & Hickl GmbH,德国)分析每个像素的荧光衰减曲线,用双 e 指数对衰减曲线进行拟合,激光发射脉冲后的荧光强度可表示为

$$I(t) = a_1 \exp\left(-\frac{t}{\tau_1}\right) + a_2 \exp\left(-\frac{t}{\tau_2}\right) + C, \quad (1)$$

式中, t 为时间, τ_1 和 τ_2 分别为 NAD(P)H 的游离态和结合态寿命, a_1 和 a_2 分别为其在整个寿命中的占比。每个像素点的平均寿命 t_m 可表示为

$$t_m = a_1 \tau_1 + a_2 \tau_2. \quad (2)$$

针对每张 FLIM 图像数据,荧光强度图像、不同特征成分的荧光寿命(t_m/a_2)伪彩图以及包含每个像

素点 t_m 和 a_2 的数据文件均可通过 SPCImage 软件获得,用于后续的分析处理。

2.3 酵母细胞图像的分割方法

实验获得的 SC 细胞原始荧光强度图有时会出现部分区域亮度低或对比度差的问题,给分割造成了一定困难,因此,需要在分割前对原始图像进行预处理。限制对比度自适应直方图均衡化(CLAHE)方法^[22]能有效提高图像亮度,克服图像噪声干扰,凸显细胞区域。因此,用 CLAHE 方法将整张图像分割成单独的图像块,以块为单位进行直方图均衡化,增强 SC 细胞荧光强度图的对比度。

为了高效准确地分割 SC 细胞,在 U-Net 的基础上添加了深度监督对模型进行优化。U-Net 是一种端到端的分割网络模型,自提出以来,在医学图像以及细胞分割领域展现出了优越的性能^[23-25]。用下采样的编码器部分获取语义特征,通过解码器部分的上采样将特征图恢复到输入尺寸,得到最终的分割结果。为了提升 SC 细胞的分割准确率,使网络能适应 SC 细胞在不同时期的形态大小,基于深度监督策略,将解码器部分的特征图进行常规 3×3 卷积及双线性上采样操作后,进行通道连接,特征图通过最后的卷积层处理得到分割结果,具体的网络结构如图 1 所示。其中,Conv 为卷积层,Up-Conv 为上采样反卷积层,ReLU 为线性激活函数,MaxPool 表示窗口尺寸为 2×2 的最大池化操作。选取 29 张 SC 细胞的荧光强度图作为训练集和验证集,用 11 张 SC 细胞的荧光强度图作为测试图像,分

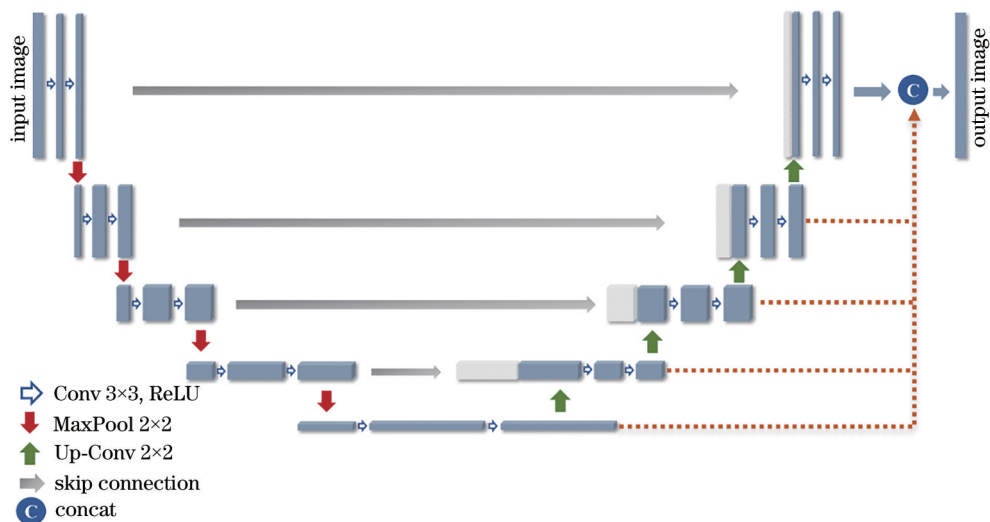


图 1 基于深度监督和 U-Net 的分割网络结构

Fig. 1 Structure of the segmentation network based on deep supervision and U-Net

割标签通过手动标注。输入的图像原始尺寸为 $256 \text{ pixel} \times 256 \text{ pixel}$, 为了减少过拟合情况, 增加图像的泛化能力, 利用随机裁剪、旋转不同的角度、水平和垂直翻转以及平移等数据增强方式扩充数据集。在目标和背景的二分类任务中, 将 Dice loss 作为损失函数, 用随机梯度下降 (SGD) 法作为优化器, 将初始学习率设为 0.001, 批量大小设为 4, 所有模型在该条件下的训练迭代次数均为 100, 最终网络输出尺寸为 $256 \text{ pixel} \times 256 \text{ pixel}$ 的掩模图像。

因 SC 细胞图像具有细胞密集的特性, 对分割预测结果中的粘连区域基于分水岭 (Watershed) 算法^[26]进行二次分割。通过连通域分析法, 找到图像中具有相同像素值且位置相邻的连通区域, 分析每个区域的像素数量、形态特征等相关标记信息, 并根据图像中 SC 细胞粘连区域像素数量多及圆率小的特性, 保留粘连可能性大的区域, 将其余区域的像素值置为 0。对二值图像进行距离变换, 将每个像素的值确定为其与最邻近非 0 像素之间的距离, 再根据分水岭算法得到脊线, 叠加在原始输入图像上, 即可实现粘连区域的分离, 且仅对粘连区域进行分割, 避免了局部极小值点导致的过分割。

2.4 酵母图像特征的提取

分割完成后, 根据连通域算法, 获取单个连通区域的位置和像素数量信息, 以计算单个 SC 细胞的平均寿命 t_m 、NAD(P)H 的蛋白结合态寿命占比 a_2 的均值和横截面积。除此之外, 利用基于迁移学习的 CNN 作为特征提取器, 获取 SC 细胞荧光寿命图的高维抽象特征。基于模型的迁移学习已广泛应用于生物医学领域中^[27-29], 这种方法将在大型数据集上预先训练好的部分网络 (包含网络结构和网络参数) 迁移到目标域中, 虽然目标数据源和训练数据源不相同, 但由于深度神经网络的前几层感受野较小, 可用于表征细粒度特征, 具有一定的通用性^[30]。选取深度卷积网络 (VGG16) 作为网络模型, 将其在 ImageNet 大型数据集上训练后的权重和偏置迁移到目标网络中。相比 VGG16 目标网络的修改: 1) 仅保留了前 4 个卷积块; 2) 第 4 个卷积块中的最大池化核尺寸由原始的 2 改为 24, 以降低输出特征维度。以单个 SC 细胞 t_m 、 a_2 的 FLIM 图像作为网络输入, 尺寸为 $48 \text{ pixel} \times 48 \text{ pixel}$ 。利用分割后的掩模对整张图像上的连通域进行逐个遍历, 裁剪每个连通域对应的原始图像, 然后用 6 h、24 h、72 h 的数据分别得到维度为 128 的特征向量。

为了了解 CNN 提取的特征分布, 采用 t-分布随机邻居嵌入 (t-SNE) 方法进行可视化。t-SNE 方法是一种无监督的降维技术, 可将多维特征向量之间的距离关系映射到二维旋转不变的欧氏空间中并保持数据之间的相对关系, 自 2008 年由 van der Maaten 等^[31]提出后, 在机器学习领域得到了广泛应用。因此, 用 t-SNE 方法可视化荧光寿命图后提取高维特征, 再进行分析观察。

2.5 聚类分析

在不同培养时间下均存在年轻和衰老的 SC 细胞, 无法确定单个 SC 细胞的类别标签, 因此, 采用 K-means 无监督算法对提取到的特征向量进行聚类分析。K-means 算法具有简单、快速等优点, 首先, 将输入数据随机分为 K 组, 将距离簇中心近的点分为同一簇, 且随着同簇数据的增多, 会重新确定簇中心, 最终通过迭代计算直观地展示聚合情况和中心点的位置信息。本方法用欧氏距离作为衡量数据间相似性的指标, 将初始聚类数设为 2, 观察数据的聚类分布。

2.6 软件运行环境

实验基于 Windows 系统进行, 用基于 Pytorch 的深度学习框架, 用 OpenCV 计算机视觉库完成 SC 细胞的分割和特征提取, 用 Matlab 中的数据分析工具箱完成聚类分析。采用 8 核 3.7 GHz 的 CPU, 内存为 32 G, GPU 为 RTX3080 的计算机, 实验语言为 Python、Matlab。

3 分析与讨论

3.1 酿酒酵母的生长曲线

图 2 为 SC 细胞在不同时间点的 X_{OD} 值。用 SGompertz 函数拟合得到生长曲线, 以展示不同时期 SC 细胞数量与其对应的生长速率, 拟合结果可表示为

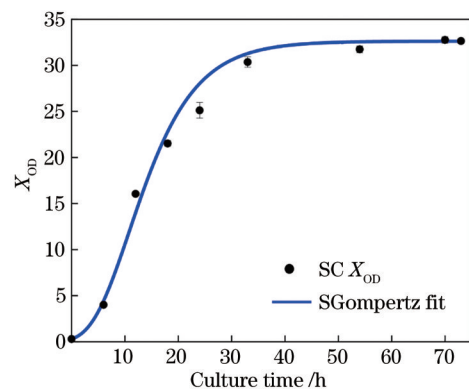


图 2 SC 细胞的生长曲线

Fig. 2 Growth curve of the SC cells

$$y = a \times \exp\left\{-\exp\left[-k \times (x - x_c)\right]\right\}, \quad (3)$$

式中, a 的值为 32.626 ± 1.292 , 表示随着时间的增长, SC 细胞最终达到的饱和 OD 值; k 的值为 0.142 ± 0.005 , 表示增长系数; x_c 的值为 10.827 ± 0.442 , 表示当培养时间处于该值时, SC 细胞的生长曲线趋于缓和。调整后曲线的拟合度 (R^2) 为 0.992, 表明函数的拟合效果较好。因此, 用 6 h, 24 h, 72 h 作为对数生长期、平稳期与衰老死亡期的三个典型时间点, 后续均用这三个时间点的 SC 细胞进行 FLIM 研究。

3.2 细胞的分割结果

为了客观评价不同分割模型的泛化能力和表现性能, 用分割领域中两种比较流行的评价指标, 即交并比 (IOU) 和 Dice 分数评估模型的性能。IOU 为真实标签和预测结果的交集和并集之比, 可表示为

$$X_{\text{IOU}} = \frac{|X \cap Y|}{|X \cup Y|}, \quad (4)$$

式中, X 和 Y 分别为真实值和预测值。Dice 分数为预测区域和标签的重叠部分, Dice 分数越大, 表示

重叠率越高, 预测结果越准确, 可表示为

$$X_{\text{Dice}} = \frac{2|X \cap Y|}{|X| + |Y|}. \quad (5)$$

表 1 为不同分割模型在 11 张测试图像上的 IOU 以及 Dice 分数均值。可以发现, 传统的最大类间方差法 (Otsu) 效果最差, 相比传统的阈值分割方法, U-Net 的分割精度有显著提升。对 U-Net 添加深度监督以及分水岭算法 (DS-UNet+watershed) 后, 模型的分割效果最好, IOU 和 Dice 分数分别为 0.832、0.908。

选取三个具有代表性的 SC 细胞图像对不同模型的分割性能进行可视化, 结果如图 3 所示。可以

表 1 不同分割模型的分割结果

Table 1 Segmentation results of different segmentation models

Model	IOU	Dice score
Otsu	0.614	0.756
U-Net	0.823	0.903
U-Net+watershed	0.825	0.904
DS-UNet	0.830	0.907
DS-UNet+watershed	0.832	0.908

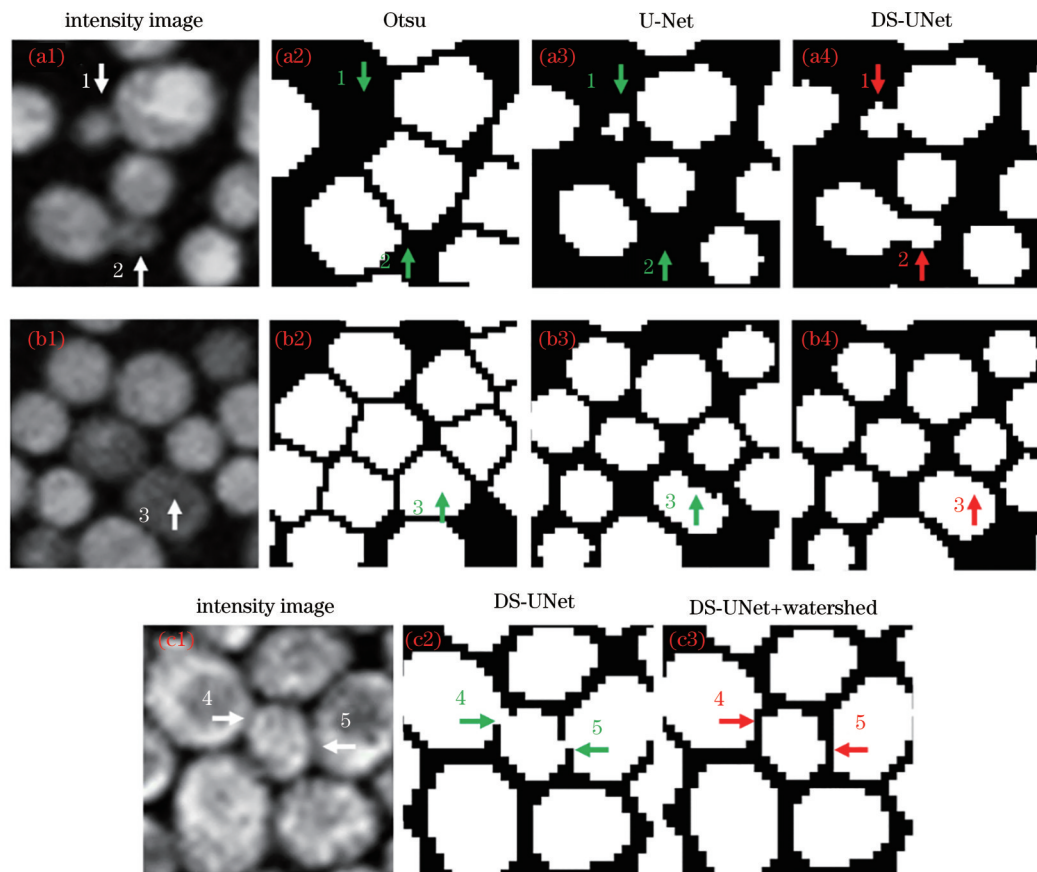


图 3 不同模型的分割结果。(a) 图像 1; (b) 图像 2; (c) 图像 3

Fig. 3 Segmentation results of different models. (a) Image 1; (b) image 2; (c) image 3

发现,基于 Otsu 得到的细胞轮廓多为多边形,而基于深度学习的 U-Net 和 DS-UNet 模型得到的细胞轮廓多为圆形或椭圆形,更符合 SC 细胞的实际情况。原因是 U-Net 的编码器可以借助浅层网络中的卷积操作,提取 SC 细胞的微小结构以及纹理等细节特征,搭配池化层进行下采样操作可扩大卷积核的感受野,从而在深层网络上获得更抽象的高级语义特征;解码器在上采样过程中,采用跨层连接的方式,拼接编码器部分处于同一深度的特征图,以融合浅层细节信息。图 3(a1)中的箭头 1 和箭头 2 分别指向两个正在生长的芽,且尺寸很小,尚未形成子细胞,U-Net 将箭头 1 指向的芽误识别为一个单独的细胞,对于箭头 2 指向的芽则丢失了相关信息,而 DS-UNet 准确识别分割了箭头 1、2 所指的芽。图 3(b1)为细胞亮度不均匀的荧光强度图,U-Net 的分割结果出现了细胞图像像素缺失的情况(如图 3(b3)中的箭头 3),而 DS-UNet 能保证 SC 细胞形态上分割的完整性,相比 U-Net,分割效果更好。原因是在 U-Net 的网络结构上添加深度监督后,DS-UNet 模型同时将多尺度特征进行融合,更加关注细节信息,使网络得到了更充分的训练。将 DS-UNet 模型分割后的图像用分水岭算法进行后处理,结果如图 3(c2)和图 3(c3)所示。在箭头 4、箭头 5 指向的区域中,DS-UNet+watershed 模型可以区分细胞粘连区域,使单细胞的分割结果更准确。综上所述,从图像分割效果来看,DS-UNet+watershed 模型的效果最理想。

除 SC 细胞图像的分割效果外,模型的运行时间对大量图像的处理也至关重要。Otsu 模型所需的时间最少,约为 1 min;U-Net 和 DS-UNet 模型在设置相同参数时,网络模型的训练时长均为 30 min 左右,11 张图像的预测总时长约为 2 s,无明显差异。虽然 Otsu 模型的分割速度最快,但分割结果较差,不利于后续进一步的分析。相比其他模型,DS-UNet+watershed 模型在评价指标以及分割细节上具有较好的鲁棒性和适应性。综合考虑客观评价指标、图像整体分割性能以及运行时间,后续实验均选用 DS-UNet+watershed 模型对图像进行分割。

3.3 特征提取结果

针对分割结果中存在的误差,在特征统计之前,先进行连通域筛选。观察多张 SC 细胞图像的荧光寿命分布曲线发现:SC 细胞的平均荧光寿命 t_m

主要分布在 700~1500 ps 范围内,NAD(P)H 的蛋白结合态寿命占比 a_2 主要分布在 10%~30% 范围内,将不在该范围内的值判定为背景;在形态上,实验拍摄的 FLIM 图像和已报道的图像都有类似规律,即单个 SC 细胞的横截面积在 10~100 μm^2 范围内^[21],圆度大于 0.6^[32]。因此仅保留 t_m 、 a_2 、横截面积、圆度值在合理范围内的连通域,以提高统计对象的精准性。

图 4 为 SC 细胞在培育 6 h、24 h、72 h 后的 NAD(P)H 荧光寿命图像、分布曲线以及统计均值。图 4(a)和图 4(b)为 SC 细胞的 t_m 和 a_2 图像,可以发现,在 6 h 时,大部分 SC 细胞呈黄色,24 h 时大部分 SC 细胞呈绿色,少量的 SC 细胞呈蓝色,72 h 时大部分 SC 细胞呈蓝色,对应于 SC 细胞平均荧光寿命和蛋白结合态寿命占比 a_2 的不断增长。图 4(c)为图 4(a)和图 4(b)对应的 t_m 、 a_2 、横截面积分布曲线,通过计算图中所有单细胞 SC 在不同荧光寿命区间的数量获得。可以发现:SC 细胞在 6 h 的 t_m 峰值为 880 ps,在 24 h 的 t_m 峰值为 963 ps,在 72 h 的 t_m 峰值 1114 ps;在 6 h 的 a_2 峰值为 16.1%,在 24 h 的 a_2 峰值为 19.4%,在 72 h 的 a_2 峰值为 22.6%。SC 细胞的横截面积在三个不同时刻下多分布在 23~25 μm^2 范围内,无明显差异,但相对于 6 h 和 24 h 的 SC 细胞,72 h 时横截面积在 40~60 μm^2 范围内的细胞数量占比更高。图 4(d)为不同视野区域内所有 SC 细胞的统计均值,统计得到培育 6 h、24 h、72 h 后的 SC 细胞数量分别为 947、1166、1696, t_m 的统计平均值分别为 (865±100) ps、(943±126) ps、(1076±143) ps, a_2 的统计平均值分别为 16.6%±3.0%、19.1%±3.7%、21.7%±3.6%,横截面积的统计平均值分别为 29.4%±14.2%、29.3%±12.9%、31.6%±14.1%。可以发现, t_m 和 a_2 的统计平均值都随培育时间的增加而增加,横截面积的统计平均值差异不大,但标准差相差较大,这表明相同条件下 SC 细胞的个体差异很大。FLIM 图像与统计分布结果表明,SC 细胞从对数生长期到稳定期再到衰老期,代谢状态逐渐发生变化,NAD(P)H 结合态寿命的占比也逐渐增加。实验采用的统计分析方法可以直接定位到单个 SC 细胞,进行荧光寿命信息或形态大小特征的统计,相比 SPCImage 软件直接对整张图像进行分析,可以有效排除背景干扰,在 SC 细胞密集且形态大小相似的情况下,荧光寿命的变化更直观和有效。

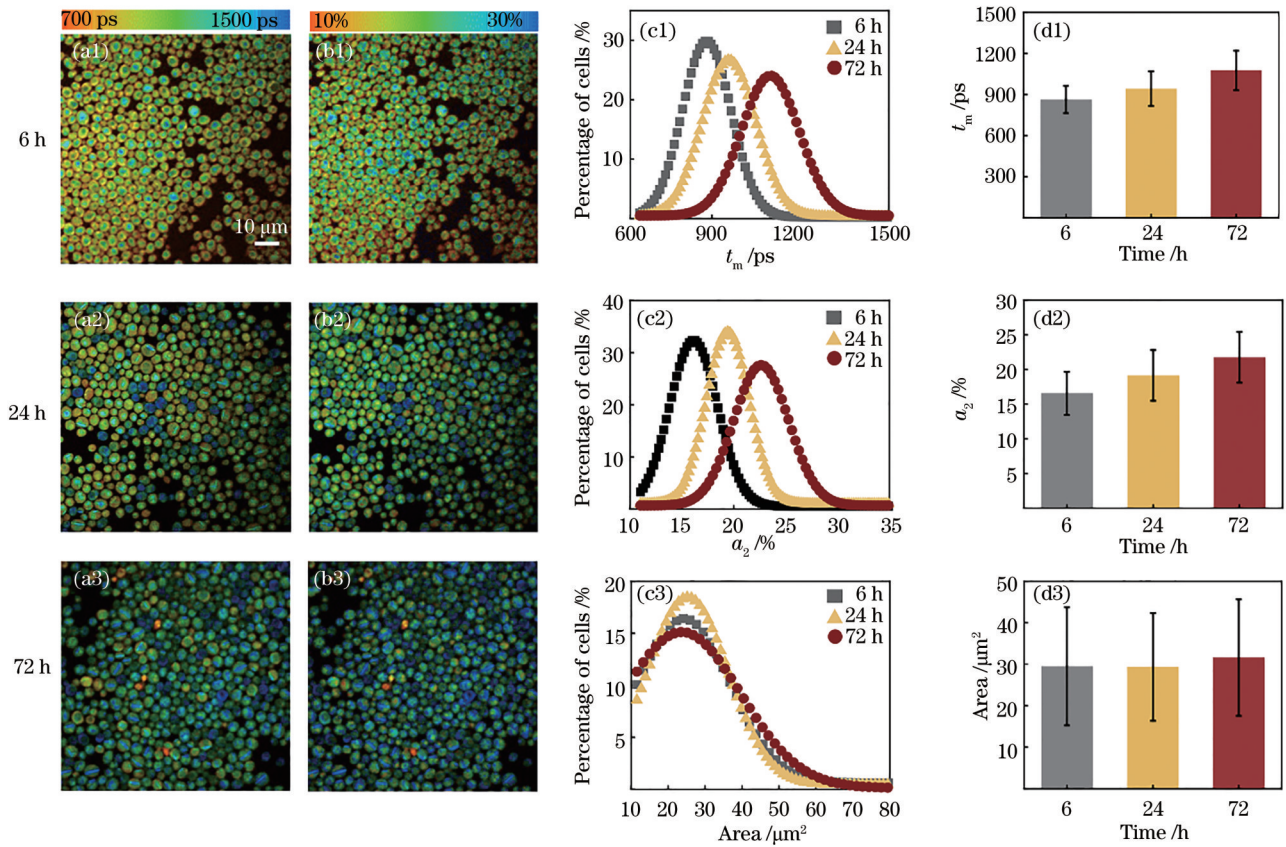


图 4 不同时长酵母细胞的 FLIM 图、分布曲线和统计平均值。(a) t_m 的 FLIM 图; (b) a_2 的 FLIM 图; (c) t_m 、 a_2 、横截面积的分布曲线; (d) t_m 、 a_2 、横截面积的统计平均值

Fig. 4 FLIM images, distribution curves and statistical values of yeast cells at different ages. (a) FLIM images of t_m ; (b) FLIM images of a_2 ; (c) distribution curve of t_m , a_2 and cross-sectional area; (d) statistical average value of t_m , a_2 and cross-sectional area

利用 VGG16 网络模型, 将单细胞酵母的 t_m 荧光寿命图、单细胞酵母 a_2 的荧光寿命图以及同一个细胞 t_m 和 a_2 的 RGB (Red, Green, Blue) 图像组成 6 通道数据分别作为输入时, 每张图像通过特征提取后可获得 128 维的特征向量。图 5 为采用 t-SNE 方法对高维特征的可视化展示。可以发现, 不同培养时间

下 SC 细胞的荧光寿命图像存在视觉差异特征, 培养 6 h 和 72 h 的 SC 细胞数据在空间位置上的距离相对较远, 培养 24 h 的 SC 细胞数据位于中间, 呈现一个逐步变化的过程。t-SNE 图通过点与点之间的距离表征数据之间的相似性, 点的分布无数值化表示。

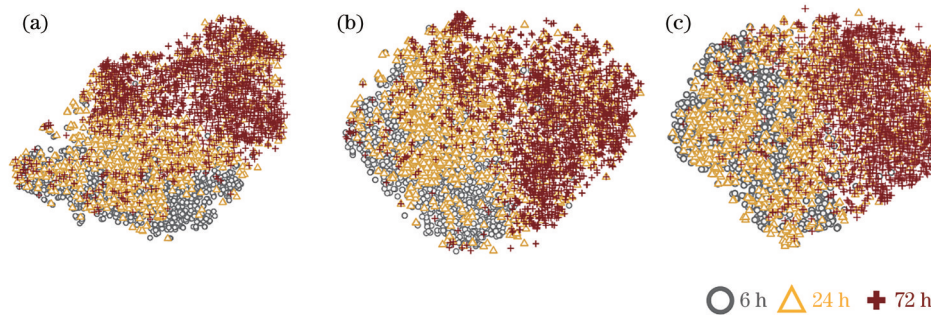


图 5 t-SNE 方法的可视化结果。(a) t_m 图; (b) a_2 图; (c) t_m 和 a_2 图

Fig. 5 Visualization results of t-SNE method. (a) t_m map; (b) a_2 map; (c) t_m and a_2 maps

3.4 聚类结果

选取不同的特征组合作为聚类分析的输入, 利

用 K-means 算法将数据分为两个类别。图 6(a) 为将 t_m 和 a_2 进行标准化后作为特征输入的聚类结果,

类中心分别为 $(-0.654, -0.656)$ 和 $(0.838, 0.815)$ 。图 6(b)~图 6(d) 分别为培养 6 h、24 h、72 h 的 SC 细胞数据分布, 每个数据点表示一个 SC 细胞, 十字标记为密集中心, 6 h、24 h、72 h 的密集中心坐标分别为 $(-0.757, -0.755)$ 、 $(-0.256, -0.126)$ 、 $(0.598, 0.509)$ 。从数据分布中可以发现, 随着培养时间的增加, 三组数据的密集中心均

在不断变化。6 h 后绝大部分 SC 细胞的二维特征集中分布在第一簇中, 72 h 后绝大部分 SC 细胞的二维特征值集中分布在第二簇中, 24 h 后 SC 细胞在两簇中均有分布。图 6(e)~图 6(h) 为输入特征为 t_m 、 a_2 和酵母横截面积时的三维聚类分析结果, 和上述规律的一致性较好。

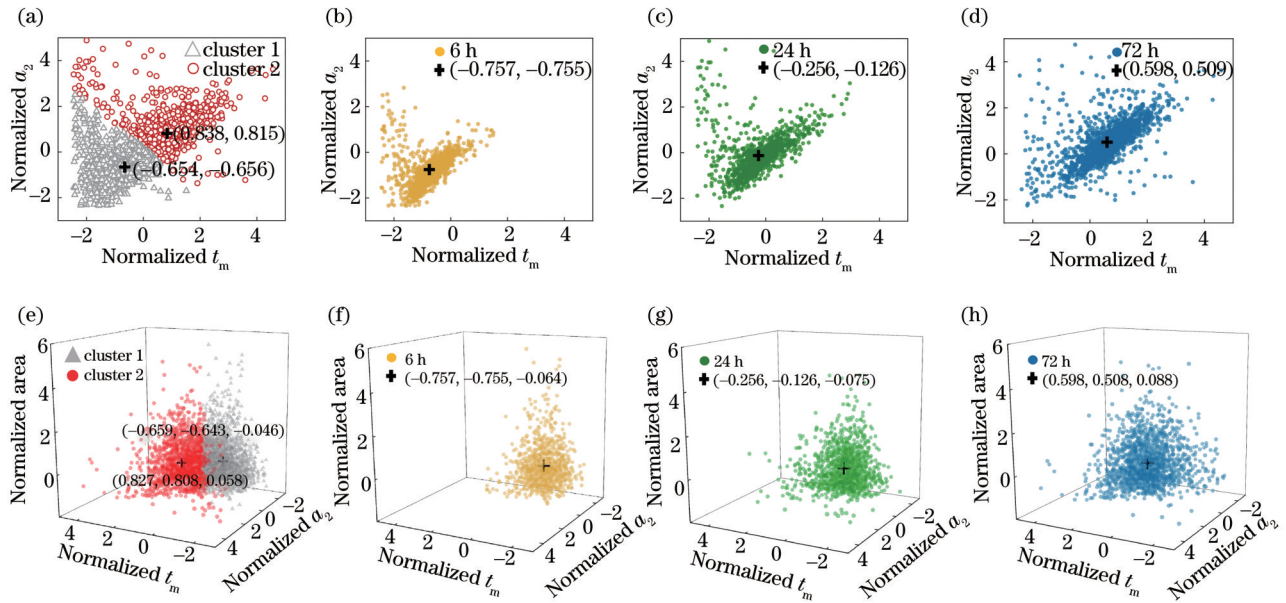


图 6 不同特征输入的聚类结果和数据分布。(a) 输入特征为 t_m 和 a_2 ; (b) 6 h 的二维特征分布; (c) 24 h 的二维特征分布; (d) 72 h 的二维特征分布; (e) 输入特征为 t_m 、 a_2 和横截面积; (f) 6 h 的三维特征分布; (g) 24 h 的三维特征分布; (h) 72 h 的三维特征分布
 Fig. 6 Clustering results and data distribution for difference feature input. (a) Input features are t_m and a_2 ; (b) two-dimensional feature distribution at 6 h; (c) two-dimensional feature distribution at 24 h; (d) two-dimensional feature distribution at 72 h; (e) input feature is t_m , a_2 and cross-sectional area; (f) three-dimensional feature distribution at 6 h; (g) three-dimensional feature distribution at 24 h; (h) three-dimensional feature distribution at 72 h

为了更好地量化输入数据判定的类别, 表 2 分别给出了 K-means 算法对不同特征的聚类结果。其中, CNN- t_m 、CNN- a_2 、CNN- $t_m \& a_2$ (Area) 分别为 CNN 从 t_m 、 a_2 、 $t_m \& a_2$ (Area) 荧光寿命图像中提取的特征。需要说明的是, 利用 CNN 提取的特征有 128 维, 无法对实际聚类情况进行可视化。可以发现, 无论选取哪种特征作为输入, 6 h 时有 79% 以上的

SC 细胞位于第一簇中, 72 h 时有 74% 以上的 SC 细胞位于第二簇, 而 24 h 的数据则居于两者之间, 这与 SC 细胞的实际生长情况一致, 即在 6 h 时, 绝大部分为生长旺盛的年轻细胞; 24 h 时, 有一部分细胞表现出衰老特征; 72 h 时, 大部分细胞处于衰老状态。因此可以推测距离第一簇中心较近的为年轻酵母, 距离第二簇中心较近的为衰老酵母。计入 SC

表 2 不同输入特征中不同簇的细胞数量占比

Table 2 Proportion of the number of cells in different clusters in different input characteristics unit: %

Feature	6 h		24 h		72 h	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
$t_m \& a_2$ (Area)	93	7	71	29	25	75
CNN- t_m	89	11	72	28	26	74
CNN- a_2	82	18	61	39	14	86
CNN- $t_m \& a_2$ (Area)	79	21	58	42	13	87

细胞的横截面积时未对分类效果产生影响,这表明相比于形态上的特征,荧光寿命信息对酵母代谢差异分析起到了关键作用。将 CNN 作为特征提取器,通过多个卷积对单个 SC 细胞的输入图像进行操作,最终输出从局部细节逐渐抽象为高维的语义信息,相比手工提取的特征,可以从全局图像中进行特征抽取,更关注细节信息,得到更接近真实情况的聚类结果。

目前实验尚无法获得单个 SC 细胞的真实标签,且目前能对单个 SC 细胞生长状态进行无标记表征的方法较少,但 Atanasova 等^[33]利用 DNA 染料 PO-TEDM-1 对大量酵母细胞进行了平均活性的统计,结果表明,酵母培养 72 h 后,细胞活性下降了 $82\% \pm 2\%$,与表 2 中的结果基本一致。此外,利用 CNN 作为特征提取器时,输入数据直接关系到特征提取的有效性,大量的研究结果也表明,融合特征能有效提高分类精度,如:Zahedi 等^[34]发现融合人类胚胎干细胞的多种形态特征,相比单一特征输入的分类准确性更高;Yang 等^[35]发现整合乳腺癌钙化点的化学和形态信息可以显著提高乳腺癌恶性肿瘤的评估效率;Chen 等^[36]将荧光寿命特征和信息熵作为输入特征参数,用于皮肤肿瘤的诊断,取得了较高的灵敏度。实验中利用 CNN 对 t_m 和 a_2 组成的 6 通道图像进行特征提取,相比将单独的 t_m 或 a_2 作为输入图像时,可能更容易提高分类的准确率。综上所述,用无标记 FLIM 探测酵母细胞生长状态时,选取 CNN- t_m &- a_2 特征作为输入可能获得更准确的分类结果,验证了将该技术应用于原位区分酵母的代谢差异、年轻和衰老酵母细胞分类中的可行性。

4 结 论

提出了一种针对 SC 细胞自体荧光寿命显微图像的自动分析方法,通过对 SC 细胞图像进行分割、特征提取和聚类分析实现不同代谢状态下酵母的快速分类。首先,通过深度监督 U-Net 实现 SC 细胞的分割,相比传统 U-Net 的分割准确性更高;然后,利用连通域算法,提取 SC 细胞的荧光寿命信息和形态大小特征,并将基于迁移学习的 CNN 作为特征提取器,从 SC 细胞荧光寿命图上获取高维抽象特征;最后,通过输入不同的特征组合,采用 K-means 聚类算法进行分类。SC 细胞从生长期到衰老期,其代谢状态存在明显的差异,实验结果表明,利用无标记成像技术和机器学习方法,无需对 SC

细胞进行繁琐的染色步骤,就能对大量细胞进行自动分割和分类,快速区分年轻和衰老的酵母。之后还将利用单细胞分选技术进行 SC 细胞的准确区分,完善无标记成像技术在区分年轻和衰老 SC 细胞上的应用,并将该方法扩展到其他细胞的荧光寿命图像分析中。

参 考 文 献

- [1] Parapouli M, Vasileiadis A, Afendra A S, et al. *Saccharomyces cerevisiae* and its industrial applications [J]. *AIMS Microbiology*, 2020, 6(1): 1-31.
- [2] Rahmat E, Kang Y. Yeast metabolic engineering for the production of pharmaceutically important secondary metabolites [J]. *Applied Microbiology and Biotechnology*, 2020, 104(11): 4659-4674.
- [3] Longo V D, Shadel G S, Kaerberlein M, et al. Replicative and chronological aging in *saccharomyces cerevisiae* [J]. *Cell Metabolism*, 2012, 16(1): 18-31.
- [4] Botstein D, Fink G R. Yeast: an experimental organism for 21st century biology [J]. *Genetics*, 2011, 189(3): 695-704.
- [5] Duina A A, Miller M E, Keeney J B. Budding yeast for budding geneticists: a primer on the *saccharomyces cerevisiae* model system [J]. *Genetics*, 2014, 197(1): 33-48.
- [6] Moreau K, Surand J, le Dantec A, et al. Recombinant yeast and human cells as screening tools to search for antibacterial agents targeting the transcription termination factor Rho [J]. *The Journal of Antibiotics*, 2018, 71(4): 447-455.
- [7] Trevors J T, Merrick R L, Russell I, et al. A comparison of methods for assessing yeast viability [J]. *Biotechnology Letters*, 1983, 5(2): 131-134.
- [8] Williams T C, Xu X, Ostrowski M, et al. Positive-feedback, ratiometric biosensor expression improves high-throughput metabolite-producer screening efficiency in yeast [J]. *Synthetic Biology*, 2017, 2(1): ysw002.
- [9] Datta R, Gillette A, Stefely M, et al. Recent innovations in fluorescence lifetime imaging microscopy for biology and medicine [J]. *Journal of Biomedical Optics*, 2021, 26: 070603.
- [10] Okkelman I A, Dmitriev R I, Foley T, et al. Use of fluorescence lifetime imaging microscopy (FLIM) as a timer of cell cycle S phase [J]. *PLoS One*, 2016, 11(12): e0167385.
- [11] Wang M Y, Tang F, Pan X B, et al. Rapid diagnosis and intraoperative margin assessment of

- human lung cancer with fluorescence lifetime imaging microscopy[J]. *BBA Clinical*, 2017, 8: 7-13.
- [12] Meleshina A V, Dudenkova V V, Shirmanova M V, et al. Probing metabolic states of differentiating stem cells using two-photon FLIM[J]. *Scientific Reports*, 2016, 6: 21853.
- [13] Carpenter A E, Jones T R, Lamprecht M R, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes[J]. *Genome Biology*, 2006, 7(10): R100.
- [14] Bredies K, Wolinski H. An active-contour based algorithm for the automated segmentation of dense yeast populations on transmission microscopy images [J]. *Computing and Visualization in Science*, 2011, 14(7): 341-352.
- [15] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [16] Lu A X, Zarin T, Hsu I S, et al. YeastSpotter: accurate and parameter-free web segmentation for microscopy images of yeast cells[J]. *Bioinformatics*, 2019, 35(21): 4525-4527.
- [17] Kong Y, Li H, Ren Y Y, et al. Automated yeast cells segmentation and counting using a parallel U-Net based two-stage framework[J]. *OSA Continuum*, 2020, 3(4): 982-992.
- [18] Buchanan R L, Whiting R C, Damert W C. When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves[J]. *Food Microbiology*, 1997, 14(4): 313-326.
- [19] Coelho M A Z, Belo I, Pinheiro R, et al. Effect of hyperbaric stress on yeast morphology: study by automated image analysis[J]. *Applied Microbiology and Biotechnology*, 2004, 66(3): 318-324.
- [20] Portell X, Ginovart M, Carbó R, et al. Differences in stationary-phase cells of a commercial *saccharomyces cerevisiae* wine yeast grown in aerobic and microaerophilic batch cultures assessed by electric particle analysis, light diffraction and flow cytometry [J]. *Journal of Industrial Microbiology & Biotechnology*, 2011, 38(1): 141-151.
- [21] Guadalupe-Daqui M, Chen M D, Thompson-Witrick K A, et al. Yeast morphology assessment through automated image analysis during fermentation [J]. *Fermentation*, 2021, 7(2): 44.
- [22] Pizer S M, Amburn E P, Austin J D, et al. Adaptive histogram equalization and its variations[J]. *Computer Vision, Graphics, and Image Processing*, 1987, 39(3): 355-368.
- [23] Ulman V, Maška M, Magnusson K E G, et al. An objective comparison of cell-tracking algorithms[J]. *Nature Methods*, 2017, 14(12): 1141-1152.
- [24] Mu H W, Guo Y, Quan X H, et al. Magnetic resonance imaging brain tumor image segmentation based on improved U-Net[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410022.
牟海维, 郭颖, 金星慧, 等. 基于改进 U-Net 的磁共振成像脑肿瘤图像分割[J]. *激光与光电子学进展*, 2021, 58(4): 0410022.
- [25] Luo W J, Han G Q, Tian X D. Retinal vessel segmentation method based on multi-scale attention analytic network[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 2017001.
罗文劫, 韩国庆, 田学东. 多尺度注意力解析网络的视网膜血管分割方法[J]. *激光与光电子学进展*, 2021, 58(20): 2017001.
- [26] Beucher S, Meyer F. The morphological approach to segmentation: the watershed transformation[M]// *Mathematical morphology in image processing*. Boca Raton: CRC Press, 2018: 433-481.
- [27] Xu W H, Pei Y J, Gao D L, et al. Mass classification of breast mammogram based on attention mechanism and transfer learning[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410007.
许文慧, 裴以建, 郜冬林, 等. 基于注意力机制与迁移学习的乳腺钼靶肿块分类[J]. *激光与光电子学进展*, 2021, 58(4): 0410007.
- [28] Alzubaidi L, Fadhel M A, Al-Shamma O, et al. Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis[J]. *Electronics*, 2020, 9(3): 427.
- [29] Alzubaidi L, Al-Shamma O, Fadhel M A, et al. Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model[J]. *Electronics*, 2020, 9(3): 445.
- [30] Oquab M, Bottou L, Laptev I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C]// *2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 1717-1724.
- [31] van der Maaten L, Hinton G. Visualizing data using

- t-SNE[J]. *Journal of Machine Learning Research*, 2008, 9(11): 2579-2605.
- [32] Ginovart M, Carbó R, Blanco M, et al. Digital image analysis of yeast single cells growing in two different oxygen concentrations to analyze the population growth and to assist individual-based modeling[J]. *Frontiers in Microbiology*, 2018, 8: 2628.
- [33] Atanasova M, Yordanova G, Nenkova R, et al. Brewing yeast viability measured using a novel fluorescent dye and image cytometer[J]. *Biotechnology & Biotechnological Equipment*, 2019, 33(1): 548-558.
- [34] Zahedi A, On V, Lin S C, et al. Evaluating cell processes, quality, and biomarkers in pluripotent stem cells using video bioinformatics[J]. *PLoS One*, 2016, 11(2): e0148642.
- [35] Yang Y F, Yang Y L, Liu Z J, et al. Microcalcification-based tumor malignancy evaluation in fresh breast biopsies with hyperspectral stimulated Raman scattering[J]. *Analytical Chemistry*, 2021, 93(15): 6223-6231.
- [36] Chen B L, Lu Y, Pan W H, et al. Support vector machine classification of nonmelanoma skin lesions based on fluorescence lifetime imaging microscopy [J]. *Analytical Chemistry*, 2019, 91(16): 10640-10647.