

## 无监督掩码循环对抗网络实现细胞虚拟染色

林俊豪 张云飞 陈少伟 张国勋 谢浩

### Unsupervised masked cycle-adversarial network for cellular virtual staining

LIN Jun-hao, ZHANG Yun-fei, CHEN Shao-wei, ZHANG Guo-xun, XIE Hao

#### 引用本文:

林俊豪, 张云飞, 陈少伟, 张国勋, 谢浩. 无监督掩码循环对抗网络实现细胞虚拟染色[J]. *中国光学*, 优先发表. doi: 10.37188/CO.2026-0021

LIN Jun-hao, ZHANG Yun-fei, CHEN Shao-wei, ZHANG Guo-xun, XIE Hao. Unsupervised masked cycle-adversarial network for cellular virtual staining[J]. *Chinese Optics*, In press. doi: 10.37188/CO.2026-0021

在线阅读 View online: <https://doi.org/10.37188/CO.2026-0021>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于深度学习的Fano共振超材料设计

Fano resonances design of metamaterials based on deep learning

中国光学 (中英文). 2023, 16(4): 816 <https://doi.org/10.37188/CO.2022-0208>

#### 基于深度学习的空间脉冲位置调制多分类检测器

Spatial pulse position modulation multi-classification detector based on deep learning

中国光学 (中英文). 2023, 16(2): 415 <https://doi.org/10.37188/CO.2022-0106>

#### 夜间动物图像自监督学习增强与检测方法

Self-supervised learning enhancement and detection methods for nocturnal animal images

中国光学 (中英文). 2024, 17(5): 1087 <https://doi.org/10.37188/CO.2024-0011>

#### 难点注意力感知红外小目标检测网络

Indistinguishable points attention-aware network for infrared small object detection

中国光学 (中英文). 2024, 17(3): 538 <https://doi.org/10.37188/CO.2023-0178>

#### 基于轻型自限制注意力的结构光相位及深度估计混合网络

A hybrid network based on light self-limited attention for structured light phase and depth estimation

中国光学 (中英文). 2024, 17(1): 118 <https://doi.org/10.37188/CO.2023-0066>

#### 基于注意力残差网络的快照式多光谱相机图像重构

Image reconstruction of snapshot multispectral camera based on an attention residual network

中国光学 (中英文). 2024, 17(6): 1316 <https://doi.org/10.37188/CO.2023-0196>

文章编号 2097-1842(xxxx)x-0001-14

# 无监督掩码循环对抗网络实现细胞虚拟染色

林俊豪<sup>1,2,3</sup>, 张云飞<sup>3</sup>, 陈少伟<sup>3</sup>, 张国勋<sup>4\*</sup>, 谢浩<sup>3,5,2\*</sup>

(1. 南京信息工程大学 计算机学院, 江苏 南京 210044;

2. 中国科学院大学南京学院, 江苏 南京 211135;

3. 长三角物理研究中心, 江苏 溧阳 213300;

4. 加州大学旧金山分校, 美国 旧金山 94158;

5. 中国科学院物理研究所, 北京, 100190)

**摘要:** 虚拟染色技术通过深度学习实现无标记成像到荧光特异性成像的转换, 能够显著降低活细胞成像的复杂性和光毒性, 从而实现多通道、高通量、长时程的高分辨率成像, 对生物医学研究具有重要意义。现有方法多依赖配对数据的有监督学习, 为降低虚拟染色对配对数据的依赖, 并进一步提升生成图像的质量, 本文提出一种融合掩码自监督机制的无监督虚拟染色框架 MVS-CycleGAN。该方法无需配对图像, 通过引入随机掩码重建任务, 遮挡输入图像的部分区域并强制网络利用语义信息进行补充, 使模型能够同时捕捉目标域的全局形态和局部纹理, 有效施加语义约束, 从而缓解传统无监督模型在跨域转换中常见的语义漂移问题。在三类细胞数据集上的实验表明, MVS-CycleGAN 整体优于传统方法: FSIM 在 BJ-5ta 细胞膜/细胞核分别为 0.784 和 0.565, HEK293T 为 0.854/0.830, Neuromast 为 0.657/0.740(分别提升了 1.03%、9.50%、1.07%、0.85%、1.08%、5.56%)。此外, 下游分割实验进一步证实了虚拟染色图像在定量分析中的有效性。研究表明, 该方法为虚拟染色技术在多样化生物医学场景中的应用提供一种可行的解决思路。

**关键词:** 虚拟荧光染色; 无监督学习; 掩码自监督; 深度学习

中图分类号: TP394.1; TH691.9 文献标志码: A doi: 10.37188/CO.2026-0021 CSTR: 32171.14.CO.2026-0021

## Unsupervised masked cycle-adversarial network for cellular virtual staining

LIN Jun-hao<sup>1,2,3</sup>, ZHANG Yun-fei<sup>3</sup>, CHEN Shao-wei<sup>3</sup>, ZHANG Guo-xun<sup>4\*</sup>, XIE Hao<sup>3,5,2\*</sup>

(1. College of Computer and Software, Nanjing University of Information Technology, Nanjing 210044, China;

2. University of Chinese Academy of Sciences, Nanjing, Nanjing 211135, China;

3. Yangtze River Delta Physics Research Center, Liyang 213300, China;

4. University of California, San Francisco, San Francisco 94158, USA;

5. Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China)

\* Corresponding author, E-mail: Guoxun.Zhang@ucsf.edu; xiehao@iphy.ac.cn

**Abstract:** Virtual staining leverages deep learning to transform label-free images into fluorescence-specific images, markedly reducing the complexity and phototoxicity of live-cell imaging and enabling high-resolu-

收稿日期: 2026-02-12; 修订日期: xxxx-xx-xx

基金项目: 国家自然科学基金委青年科学基金 B 类 (No. 62422515)

Supported by

tion, multi-channel, high-throughput, and long-term acquisition, which is of great significance for biomedical research. Existing methods mostly rely on supervised learning with paired data. To reduce the dependence of virtual staining on paired data and further improve the quality of generated images, this paper proposes an unsupervised virtual staining framework, MVS-CycleGAN, which integrates a masked self-supervised mechanism. Without requiring paired images, MVS-CycleGAN introduces a random masked reconstruction task that occludes parts of the input and forces the network to complete the missing regions using semantic context. This design allows the model to capture both global morphology and local texture in the target domain, imposing effective semantic constraints and alleviating the semantic drift commonly observed in cross-domain translation with conventional unsupervised models. Experiments on three cell datasets demonstrate that MVS-CycleGAN consistently outperforms traditional approaches: FSIM reaches 0.784/0.565 on BJ-5ta membrane/nuclei, 0.854/0.830 on HEK293T, and 0.657/0.740 on Neuromast (corresponding improvements of 1.03%, 9.50%, 1.07%, 0.85%, 1.08%, and 5.56%, respectively). In addition, downstream segmentation experiments further confirm the effectiveness of the virtually stained images for quantitative analysis. These results indicate that the proposed method provides a feasible solution for extending virtual staining to diverse biomedical scenarios.

**Key words:** virtual fluorescence staining; unsupervised learning; masked self-supervised learning; deep learning

## 1 引 言

在生命科学研究中, 光学显微成像技术广泛应用于观察细胞形态结构与动态功能, 基于是否对样本进行特异性标记演化出了荧光显微成像和无标记显微成像两大技术分支<sup>[1]</sup>。荧光成像通过使用特定的荧光染料、荧光蛋白或其他标记物与目标生物结构或分子结合, 在特定波长激光激发下产生可被检测到的光子信号, 经过传感器采集得到具有分子特异性的细胞和组织图像<sup>[2-4]</sup>, 进而分辨生物细胞样本中的微小信号或高分辨率结构<sup>[5]</sup>。同时, 相关图像增强与反卷积方法的引入能够进一步提升荧光显微成像的质量与保真度<sup>[6]</sup>。相比之下, 无标记显微成像通过直接使用光采集样品的固有物理化学属性, 如样品的折射率、自发荧光、分子振动、散射或吸收特性, 来获取样品的结构和成分信息<sup>[7-8]</sup>。

然而传统荧光显微成像技术存在固有的局限性。该方法通常依赖于外源性荧光标记指示剂, 涉及多步标记流程, 不仅实验周期长、成本高, 且对活细胞存在潜在的光毒性与光漂白风险, 限制了其在动态、长时程观测中的应用<sup>[9-10]</sup>。此外, 以荧光显微成像对生物样本进行多通道成像非常困

难, 其一成像系统依赖滤光片二向色镜系统采集多通道图像, 每多一个通道便意味着硬件成本和系统复杂度的增加; 其二需要对样本进行多重标记, 这将干扰活体样本的原始生理状态, 甚至对其自身系统造成较大负担; 其三目前的生物化学标记方法光谱较宽, 不同标记方法的光谱之间无法避免产生混叠, 随着标记种类的增加混叠愈发严重; 其四多通道成像需要对样本进行反复照明和采集, 这不仅降低采集速度, 也增加了成像过程的光毒性<sup>[11]</sup>。相较于荧光显微成像方法, 无标记显微成像主要通过样品内在属性获取形态与化学信息, 分子特异性较低, 难以直接分辨特定的亚细胞高分辨率结构, 从而一定程度上限制了对复杂生物过程的发现和研究<sup>[7, 12-13]</sup>。

为了解决上述问题, 进而实现灵活高效的多通道长时程高通量高分辨率成像, 基于无标记显微成像的虚拟染色技术应运而生。虚拟染色技术通过人工智能算法学习无标记图像和荧光图像之间的映射关系, 在无需实际荧光标记的情况下生成真实荧光成像的预测结果, 从而可以在保证细胞活性的基础上, 快速、低成本地获得可供分析的荧光成像结果<sup>[14-15]</sup>。此外, 由于无需经历真实染色和荧光信号采集环节, 虚拟染色在一定程度上避免了光毒性、光漂白以及样本制备流程复杂

等问题<sup>[16]</sup>。近年来, 基于深度学习的虚拟染色方法得到了快速发展<sup>[17]</sup>。Christiansen 等人首次提出了 In Silico Labeling (ISL) 框架, 开创性地利用基于 U-Net 的深度卷积神经网络, 从透射光 z-stack 图像中预测多种荧光标签, 为无标记图像的计算标记提供了新的思路<sup>[5]</sup>。Ounkomol 等人进一步将该思想扩展至三维荧光图像预测, 实现了对多种亚细胞结构的高保真三维重建<sup>[18]</sup>。然而, 这些方法高度依赖大量高质量的配对训练样本, 获取像素级配对数据不仅耗时费力, 而且成本高昂<sup>[19]</sup>。配对训练样本的获得需要搭建多模态成像系统, 并对系统的模态之间进行仔细地标定, 然后对采集得到的不同模态之间的图像进行像素级配准。此外在某些情况下, 由于感兴趣的生命活动的动态过程快速且不可逆地发生, 难以获得满足严格对齐要求的训练对<sup>[20]</sup>。因此, 对标注数据的强需求与高质量配对数据难以获取之间的矛盾, 阻碍了深度学习在虚拟染色成像领域中进一步的研究<sup>[21]</sup>。无监督深度学习框架 CycleGAN 最初由 Zhu 等人提出, 通过引入循环一致性损失, 在无配对训练数据情况下学习源域到目标域的映射, 从而实现了无监督的图像模态转换, 使得无监督的图像模态转换成为可能<sup>[22]</sup>。在此基础上, Li 等人构建了基于 CycleGAN 的 UTOM 框架, 该方法通过添加显著性约束以保持源图像内容不丢失, 实现了包括自发荧光图像到标准 H&E (Hematoxylin and Eosin) 的虚拟染色、荧光图像恢复以及虚拟荧光标记等多项任务<sup>[23]</sup>。

尽管上述方法展现了虚拟染色技术在实际应用中的巨大潜力, 但现有方法仍存在一定局限。多数性能较优的方法依赖配对训练数据, 而明场图像与荧光图像的精确配准获取成本较高, 限制了方法的实际应用范围。相比之下, 无监督虚拟染色方法虽然能够降低对配对数据的依赖, 但在图像转换过程中仍容易出现结构信息保持不足、局部细节失真以及生成结果不稳定等问题。针对这些问题, 近年来已有研究者进行了相关探索。Liu 等人提出的 Cytoland 方法通过结合自监督与监督学习策略, 实现了从无标记明场或相位图像到细胞核与细胞膜荧光图像的预测, 并取得了较好的成像效果<sup>[24]</sup>。然而, 此类方法仍在一定程度上依赖配对数据或复杂训练策略。

为减少对大规模配对数据的依赖, 本文提出了一种基于无监督的虚拟染色框架 MVS-CycleGAN。该方法不仅不需要配对训练样本, 而且引入自监督图像重建任务, 能够引导网络在训练过程中更充分地学习细胞的高分辨率形态学特征, 从而提升特征表示的稳健性。在此基础上, 模型实现了明场图像向荧光图像的无监督转换, 并在不同生物样本上取得了较好的虚拟染色结果。实验中, 我们在人成纤维细胞、人胚肾细胞以及斑马鱼神经丘上对所提出方法进行了验证。结果表明, 该方法不仅在虚拟荧光预测质量上具有良好表现, 并且显著提升了下游图像分割等分析任务的性能, 为虚拟染色技术在多样化生物学场景中的应用提供了可能。

## 2 方法及原理

### 2.1 用于无监督虚拟染色的 MVS-CycleGAN 框架

为训练对抗生成网络实现明场图像到荧光图像的虚拟染色, 我们选取了三种不同类型的细胞图像, 以用于验证模型在不同细胞类型和成像条件下的有效性 (图 1a)。将原始数据集图像随机裁剪后传入生成器中, 得到与真实图像相似的虚拟染色图像, 再将其与非配对的荧光图像传入判别器中, 训练判别器对真实图像和生成图像的判断能力, 最后将参数传入优化器中对生成器权重进行更新, 使生成器与判别器交替训练, 判别器通过分辨真实图像和生成图像的差异来给生成器提供反馈, 生成器在训练过程中不断调整其参数, 以产生能够欺骗判别器的“真实”图像。最终生成器通过对抗性训练使得生成图像更加真实, 从而实现无监督的图像到图像转换 (图 1b)。本研究利用 CycleGAN 的循环一致性约束, 使模型能够在不依赖配对样本的情况下进行学习。它约束模型在转换前后保持一致, 重点捕捉明场图像与荧光图像之间与内容无关的风格差异, 从而实现无监督的虚拟染色效果 (图 1c)。在训练过程中, 我们采用了掩码自编码器损失等自监督策略, 通过在生成器中加入掩码重建任务, 加强模型对细胞形态学特征的提取能力, 使得生成器能够专注于提取与细胞形态相关的特征, 从而提高图像转换效果 (图 1d)。

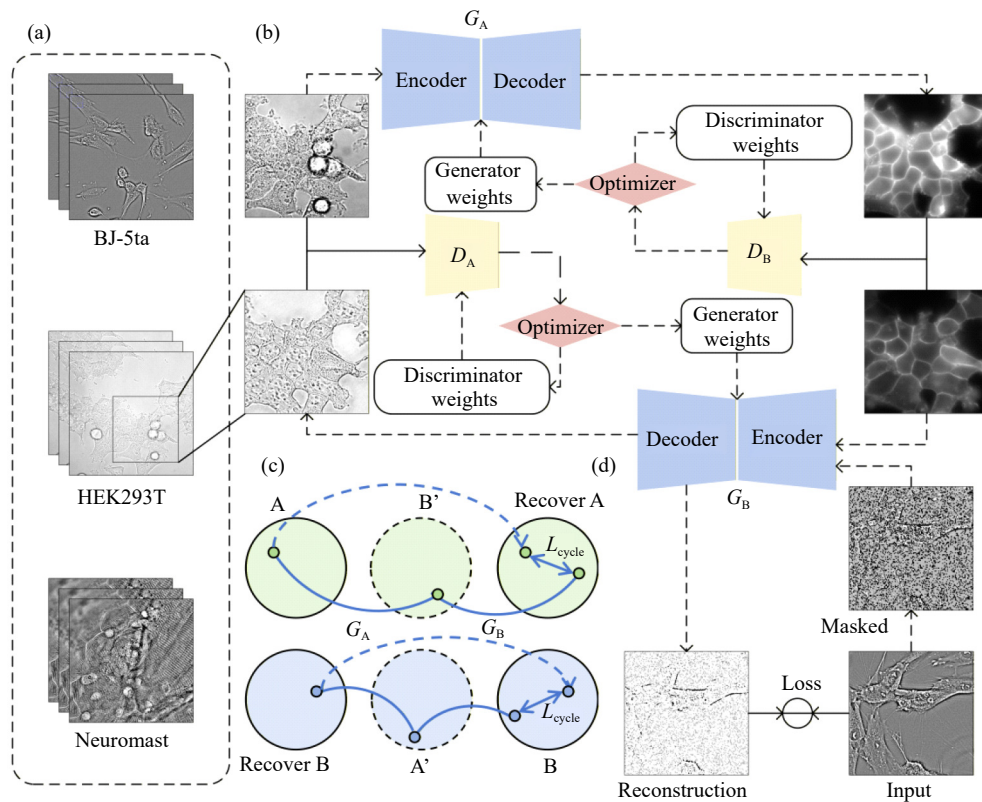


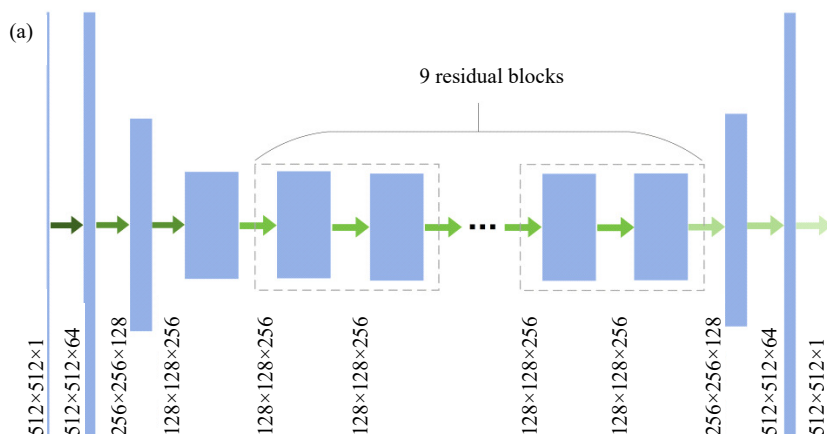
图 1 用于非配对虚拟染色任务的 MVS-CycleGAN 总体架构和训练框架示意图 (a) 不同细胞类型明场图像数据集, 包括 BJ-5ta 人成纤维细胞、HEK293T 人胚肾细胞和 Neuromast 斑马鱼神经丘; (b) 训练流程框架示意图; (c) 循环一致性约束机制; (d) 掩码自监督学习模块

Fig. 1 Overall architecture and training framework of the proposed MVS-CycleGAN for unpaired virtual staining (a) bright-field image datasets from different cell types, including BJ-5ta human fibroblast cells, HEK293T human embryonic kidney cells and Zebrafish Neuromast; (b) Schematic of training process framework; (c) Cycle consistency constraint mechanism; (d) masked self-supervised learning module

## 2.2 网络结构与框架

本文网络模型的总体架构与 CycleGAN 相似<sup>[22]</sup>, 由两个生成器  $G_A$ 、 $G_B$  和两个判别器  $D_A$ 、 $D_B$  构成。生成器采用基于残差学习的编码器-解码器结构, 如图 2(a) 所示; 判别器采用全卷积神经网络结构, 如图 2(b) 所示。以  $G_A$ 、 $G_B$  为例, 该模

型的输入是经过随机裁剪尺寸为  $512 \times 512$  的细胞明场图像, 在生成器  $G_A$  的编码器网络中, 图像通过反射填充和卷积进行初步的特征提取以获得低层的视觉特征, 经过 2 次下采样, 逐步降低空间分辨率并提升通道维度从而扩大感受野并获得更高层次的信息。





器  $G_A$  和  $G_B$  分别用于学习从  $A$  域到  $B$  域和  $B$  域到  $A$  域的映射关系。循环一致性损失约束生成图像在前向和反向映射后仍能恢复到原始的输入图像, 从而避免生成器在转换过程中引入与内容无关的变化, 只对图像进行风格转换。在训练中, 来自  $A$  域的图像  $A$  先通过生成器  $G_A$  得到生成图像  $G_A(A)$ , 再将生成图像传入生成器  $G_B$  将其映射回  $A$  域得到重建图像  $G_B(G_A(A))$ 。重建图像与原始输入之间的差异通过 L1 损失进行度量, 如公式 (1) 所示:

$$L_{\text{cycle}}^A = \|G_B(G_A(A)) - A\|_1, \quad (1)$$

同样的, 对于来自  $B$  域的输入图像  $B$ , 对应的向环一致性损失为:

$$L_{\text{cycle}}^B = \|G_A(G_B(B)) - B\|_1, \quad (2)$$

循环一致性总损失为:

$$L_{\text{cycle}} = L_{\text{cycle}}^A + L_{\text{cycle}}^B, \quad (3)$$

受掩码自编码(masked auto-encoding)启发<sup>[25]</sup>, 我们提出了一种新的损失用于约束生成对抗网络模型, 与依赖成对的前向-反向映射约束的传统循环一致性损失不同, 掩码自编码损失在特征推理时直接提供了自监督信息。在训练过程中, 该损失函数通过在生成器中引入掩码重建任务, 主动遮挡输入图像的部分区域, 从而强制网络必须依据可见上下文推断补全缺失信息。掩码重建已被广泛用于图像自监督学习中, 已有研究表明, 掩码比例会直接影响学习任务的难度: 如果掩码比例过低, 模型仍可依赖大量局部可见信息完成恢复, 难以形成充分的重建约束; 而掩码比例过高时, 又会导致有效信息缺失过多, 使恢复任务过于困难并影响训练稳定性。在高分辨率遥感图像的掩码特征建模研究中, 作者进一步对 15%–85% 等多种掩码比例进行了系统比较, 结果表明 25%–65% 为较优区间, 过低或过高的掩码比例都会导致性能下降, 其原因在于低掩码率提供的学习挑战不足, 而高掩码率又会造成重建任务过难<sup>[26]</sup>。结合细胞显微图像的特点, 本研究最终将掩码比例设定为 25%。与自然图像不同, 细胞显微图像通常存在前景区域占比相对较小、背景区域较大、且细胞轮廓与细胞器结构连续性要求较高等特点。在这一场景下, 若直接采用过高掩码率, 容易破坏

有限的有效结构区域, 不利于细胞边界、细胞核形态及空间组织关系的恢复; 而若掩码率过低, 则又不足以迫使模型跳出局部纹理拟合, 难以充分学习更稳定的形态学特征。因此, 25% 的掩码比例可视为一种在细胞显微图像场景下较为合理且有效的平衡选择。它能够在适度增加重建难度的同时, 促使网络更多地利用上下文信息和整体形态推断被遮挡区域; 另一方面, 它也保留了足够的结构线索, 确保了训练过程具有可行的恢复目标与良好的稳定性。此外, 为避免随机掩码退化为像素级噪声、仅进行简单的纹理修补, 本文将基本掩码单元设置为  $4 \times 4$  的像素块。块状掩码能够在局部范围内连续破坏纹理信息, 从而进一步抑制模型对邻域像素直接复制的依赖, 引导其将注意力更多地集中在细胞轮廓、细胞核形态以及细胞间空间组织关系等结构特征上。基于上述分析, 并在预实验中进行验证后, 25% 的掩码比例与  $4 \times 4$  的块状掩码单元这一组合能够较好地兼顾训练过程的稳定性与对细胞关键结构的恢复效果, 故被确定为本文方法的最终参数配置。将掩码后得到的图像  $\tilde{A}$  输入到生成器  $G_B$  中得到重建图像  $G_B(\tilde{A})$ 。我们选取 L1 损失对重建结果与原始图像之间的差异进行度量, 公式 (4) 如下:

$$L_{\text{mask}}^A = \|G_B(\tilde{A}) - A\|_1, \quad (4)$$

同理, 对  $B$  域采用相同的掩码与重建策略进行施加约束, 损失函数公式如下:

$$L_{\text{mask}}^B = \|G_A(\tilde{B}) - B\|_1, \quad (5)$$

掩码自编码总损失为:

$$L_{\text{mask}} = L_{\text{mask}}^A + L_{\text{mask}}^B, \quad (6)$$

这样, 模型能更好地理解细胞形态的结构与内在规律, 而不是做逐像素的拟合, 从而提高图像转换效果。

为提升生成器生成的图像质量, 我们采用最小二乘生成对抗损失(LSGAN)。对  $B$  域的判别器  $D_B$  而言, 其目标是将真实图像  $B$  的判别输出接近 1, 并将生成图像  $G_A(A)$  的判别输出接近 0, 其中  $E$  表示在对应数据分布下的期望; 对  $A$  域的判别器  $D_A$  同理。判别器损失定义为

$$L_{D_B} = \frac{1}{2} (E_B[(D_B(B) - 1)^2] + E_A[(D_B(G_A(A)))^2]), \quad (7)$$

$$L_{D_A} = \frac{1}{2}(E_A[(D_A(A) - 1)^2] + E_B[(D_A(G_B(B)))^2]) \quad (8)$$

生成器的对抗损失则期望生成图像被判别器判为真实图像, 定义为

$$L_{GAN}^A = E_A[(D_B(G_A(A)) - 1)^2] \quad (9)$$

$$L_{GAN}^B = E_B[(D_A(G_B(B)) - 1)^2] \quad (10)$$

$$L_{GAN} = L_{GAN}^A + L_{GAN}^B \quad (11)$$

在实现中, 我们使用均方误差 (MSELoss) 作为 LSGAN 的具体形式, 并以全 1 张量作为真实样本, 全 0 张量作为生成样本的监督信号。

在无配对图像转换中, 生成器可能出现与真实生物结构无关的偏差, 如明场图像中的细胞位置、形状在转为荧光图像时可能会发生偏离。为缓解这一问题, 我们采用恒等损失, 约束当输入图像已经来自目标域时, 生成器应尽可能保持其不变。具体而言, 对于 B 域的真实图像  $B$ , 将其输入  $G_A$  应得到近似恒等输出, 对于 A 域图像同理。恒等损失函数为:

$$L_{idt}^B = \|G_A(B) - B\|_1 \quad (12)$$

$$L_{idt}^A = \|G_B(A) - A\|_1 \quad (13)$$

$$L_{idt} = L_{idt}^A + L_{idt}^B \quad (14)$$

以上各项损失按照一定权重加权求和, 构成模型的总目标函数。

## 2.5 训练

本研究在 PyTorch 框架下实现模型训练。生成器与判别器的参数均采用高斯分布进行初始化, 以加速网络收敛并提高训练稳定性。模型训练使用 Adam 优化器, 其中生成器采用联合优化策略, 同时更新  $G_A$  与  $G_B$  的参数; 两个判别器  $D_A$  与  $D_B$  分别独立优化。所有优化器的动量参数设置为  $\beta_1=0.5$ 、 $\beta_2=0.999$ , 初始学习率设为  $1 \times 10^{-4}$ 。在训练过程中, 学习率采用线性衰减策略, 在第 30 个 epoch 后逐步衰减至零, 以避免后期训练震荡并提升模型收敛性。在训练策略上, 我们采用两阶段机制。首先进行基础阶段的训练, 仅对生成器施加对抗损失、恒等损失以及掩码自监督损失, 而不引入循环一致性约束。在该阶段, 生成器

重点学习初步的跨域映射能力, 避免过早施加过强的约束而导致模式崩溃, 从而保持生成多样性。当第一阶段训练收敛后, 进入完整训练阶段, 在上述全部损失项的基础上加入循环一致性损失继续训练。此时生成器已具备一定映射能力, 再施加循环一致性可进一步确保重建保真度, 防止过度自由导致输出失真。两阶段训练策略提高了模型训练的稳定性并最终性能。

## 2.6 评估

为评估模型在不同细胞类型和成像条件下的虚拟染色性能, 本文在多种细胞类型数据上分别进行训练与测试, 并观察生成结果在图像质量和结构保真度方面的表现。通过比较不同数据集上的虚拟染色结果, 可以分析模型在多样化生物样本中的适用性。我们将本研究的方法与传统 CycleGAN 进行了对比实验, 在相同数据上分别训练和测试不含 MAE 模块的 CycleGAN, 对两种方法在视觉质量与定量指标上的差异进行系统比较。

本文从定性和定量两个方面分别对明场图像细胞核与细胞膜的虚拟染色任务进行对比分析。实验评估了 BJ-5ta 人成纤维细胞、HEK293T 人胚肾细胞以及 Neuromast 斑马鱼神经丘三种细胞形态差异显著的数据集。在定量评估中, 采用多种图像质量评价指标, 从像素误差、结构相似性、感知一致性以及整体视觉质量等多个方面, 对模型生成的虚拟荧光图像和对应的真实荧光图像进行逐像素对比, 对模型性能进行衡量。

在图像质量评价外, 本文还通过下游细胞分割任务对虚拟染色结果进行进一步验证。使用相同的分割模型与参数, 分别对明场图像与虚拟染色图像进行细胞核和细胞膜分割, 并对分割结果进行对比分析。通过观察分割结果在结构完整性、细胞边界清晰度、细胞密集区域的可分性等方面的差异, 对虚拟染色结果在实际细胞分析任务中的可用性与可靠性进行定性评估。

## 3 结果

### 3.1 不同虚拟染色方法的整体定量性能比较

图 3 比较了 MVS-CycleGAN 和传统 CycleGAN 在多种细胞类型和细胞膜与细胞核上的虚拟染色性能。为保证比较的公平性, MVS-CycleGAN 与作为基线的 CycleGAN 模型采用了完全

一致的网络架构(除 MVS-CycleGAN 增加的掩码自监督模块)、训练超参数及实验设置。具体而言,两者共享相同的生成器与判别器结构,使用相同的优化器、学习率策略、损失函数权重(除 MVS-CycleGAN 独有的  $L_{\text{mask}}$ )、训练轮次及批次大小,并在相同的数据划分与软硬件环境下进行训练。性能对比均基于同一测试集和评估指标,因此结果的差异直接反映了掩码自监督模块的贡献。可以看出,本方法生成的荧光图像在视觉上与真实染色非常接近:细胞核和细胞膜等结构清晰可见,信号背景噪声低,且不同细胞类型间的形态差异也被正确地反映出来。例如, BJ-5ta 成纤维细胞呈梭形分布, HEK293T 细胞呈上皮样扁平形态, 斑马鱼神经丘中细胞排列紧密形成器官结构, 模型都能相应地在虚拟染色中再现这些特征。

这表明模型提取到了细胞形态学的关键特征并成功迁移到荧光图像中, 而不局限于单一的细胞类型。同时我们对图像质量进行定量评估, 采用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、均方根误差(Root Mean Square Error, RMSE)、结构相似性指数(Structure Similarity Index Measure, SSIM)<sup>[27]</sup>、特征相似性指数(Feature Similarity Index Measure, FSIM)<sup>[28]</sup>、特征相似性指数(Learned Perceptual Image Patch Similarity, LPIPS)<sup>[29]</sup>和视觉保真度(Visual Information Fidelity, VIF)<sup>[30]</sup>等多种互补指标, 分别从像素误差、结构相似性以及感知一致性等角度衡量生成结果的质量, 结果见表 1 和表 2。总体而言, 本文方法在多数评价指标上均优于 CycleGAN, 有效提升了虚拟染色结果和整体视觉质量。

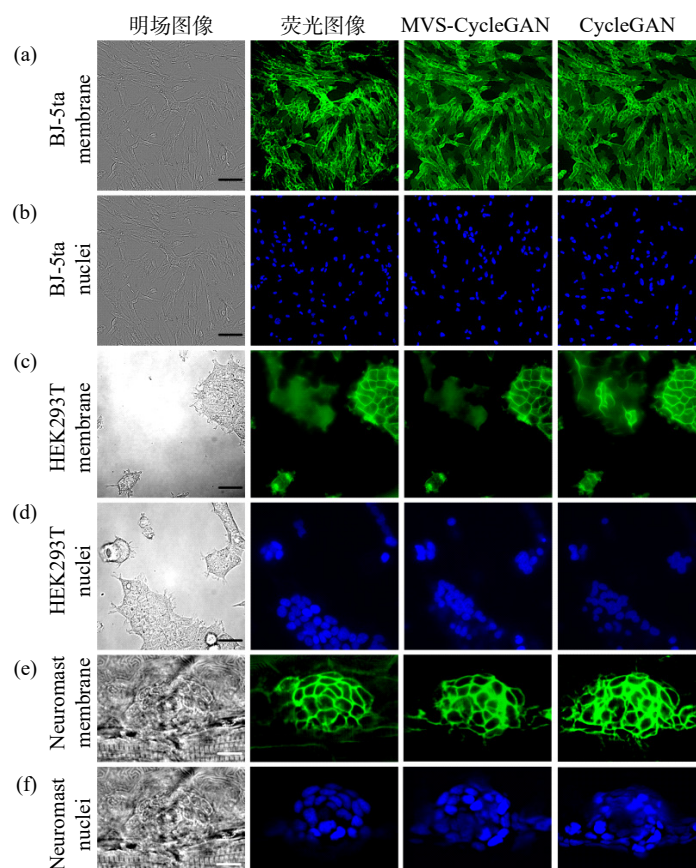


图 3 不同方法对 BJ-5ta 细胞、HEK293T 细胞和 Neuromast 细胞膜与核进行虚拟染色图像。BJ-5ta 细胞膜 (a) 和细胞核 (b) 明场图像、真实荧光图像及虚拟染色对比, 比例尺: 50  $\mu\text{m}$ ; HEK293T 细胞膜 (c) 和细胞核 (d) 明场图像、真实荧光图像及虚拟染色对比, 比例尺: 50  $\mu\text{m}$ ; Neuromast 细胞膜 (e) 和细胞核 (f) 明场图像、真实荧光图像及虚拟染色对比, 比例尺: 10  $\mu\text{m}$

Fig. 3 Virtual staining images of cell membranes and nuclei in BJ-5ta, HEK293T and Neuromast produced by different methods. Comparisons of bright-field images, real fluorescence images, and virtual staining results are shown for BJ-5ta cell membranes (a) and nuclei (b), scale bar: 50  $\mu\text{m}$ ; HEK293T cell membranes (c) and nuclei (d), scale bar: 50  $\mu\text{m}$ ; and Neuromast cell membranes (e) and nuclei (f), scale bar: 10  $\mu\text{m}$

表 1 细胞及类器官虚拟染色像素误差及结构相似性指标

Tab. 1 Cell and organoid virtual staining pixel error and structural similarity index

Models	Cell Type	PSNR	RMSE	SSIM
MVS-CycleGAN	BJ-5ta Membrane	<b>17.27</b>	<b>0.138</b>	<b>0.582</b>
CycleGAN		16.96	0.143	0.550
MVS-CycleGAN	BJ-5ta Nuclei	<b>13.40</b>	<b>0.215</b>	<b>0.724</b>
CycleGAN		12.96	0.226	0.703
MVS-CycleGAN	HEK293T Membrane	<b>20.72</b>	<b>0.094</b>	<b>0.555</b>
CycleGAN		20.60	0.095	0.485
MVS-CycleGAN	HEK293T Nuclei	<b>19.80</b>	<b>0.106</b>	0.763
CycleGAN		19.48	0.110	0.763
MVS-CycleGAN	Neuromast Membrane	<b>13.91</b>	<b>0.206</b>	<b>0.589</b>
CycleGAN		13.40	0.218	0.522
MVS-CycleGAN	Neuromast Nuclei	<b>13.07</b>	<b>0.224</b>	0.345
CycleGAN		12.48	0.240	<b>0.346</b>

表 2 细胞及类器官虚拟染色结构相似性及感知质量指标

Tab. 2 Cell and organoid virtual staining structural similarity and image perceived quality indicators

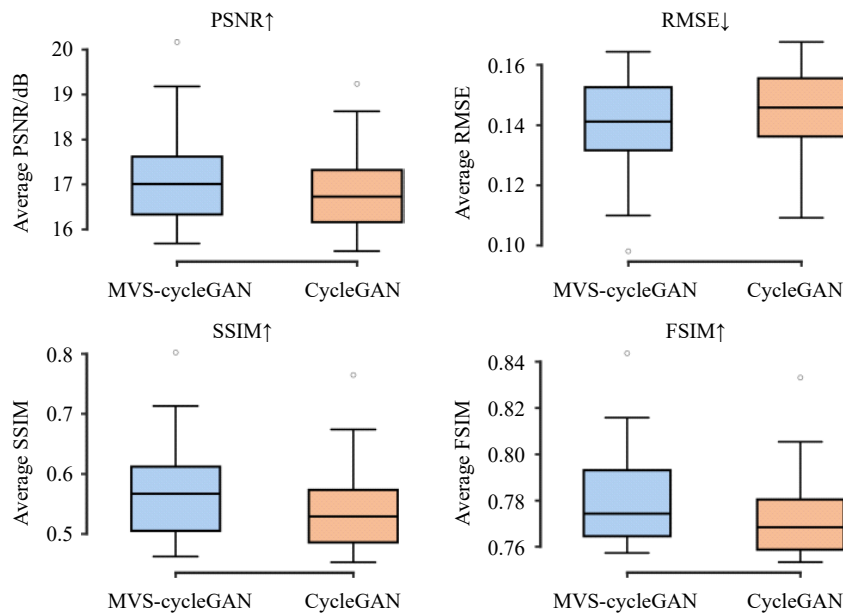
Models	Cell Type	FSIM	LPIPS	VIF
MVS-CycleGAN	BJ-5ta Membrane	<b>0.784</b>	<b>0.337</b>	0.0820
CycleGAN		0.776	0.418	<b>0.0824</b>
MVS-CycleGAN	BJ-5ta Nuclei	<b>0.565</b>	<b>0.474</b>	0.0364
CycleGAN		0.516	0.521	<b>0.0434</b>

续表 2

Models	Cell Type	FSIM	LPIPS	VIF
MVS-CycleGAN	HEK293T Membrane	<b>0.854</b>	<b>0.184</b>	<b>0.0827</b>
CycleGAN		0.845	0.191	0.0797
MVS-CycleGAN	HEK293T Nuclei	<b>0.830</b>	<b>0.171</b>	<b>0.0433</b>
CycleGAN		0.823	0.183	0.0408
MVS-CycleGAN	Neuromast Membrane	<b>0.657</b>	<b>0.401</b>	<b>0.1573</b>
CycleGAN		0.650	0.412	0.0852
MVS-CycleGAN	Neuromast Nuclei	<b>0.740</b>	<b>0.253</b>	<b>0.0906</b>
CycleGAN		0.701	0.296	0.0723

## 3.1.1 BJ-5ta 人成纤维细胞虚拟染色结果

在 BJ-5ta 细胞数据集上, 分别对细胞膜和细胞核虚拟染色任务进行了评估(图 4)。从定量结果可以看到, 本文方法相较于 CycleGAN 在 PSNR 和 RMSE 指标上有进一步提升, 细胞膜的 PSNR 达到 17.27, 同时在 SSIM 和 FSIM 等结构相似性指标上取得更高的中值和更紧密的分布, 细胞核染色结果的 SSIM 到达 0.724, 显示其生成结果在整体结构一致性方面更高。感知质量指标也证实了这一点, 在 LPIPS 上, MVS-CycleGAN 与 CycleGAN 相比表现出更低的误差分布, 细胞膜与细胞核分别为 0.337 和 0.474, 表明生成图像在感知空间中更接近真实荧光图像。从视觉结果来看, MVS-CycleGAN 生成的虚拟荧光图像在细胞边界清晰度和背景抑制方面均优于传统 CycleGAN, 但两者都出现了对细胞核空间位置恢复偏差较大的问题。



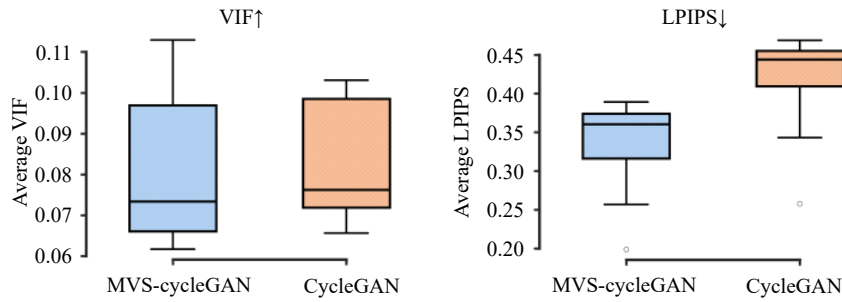


图4 BJ-5ta 人成纤维细胞细胞膜虚拟染色指标对比

Fig. 4 Comparison of virtual staining indices of BJ-5ta human fibroblast cell membranes

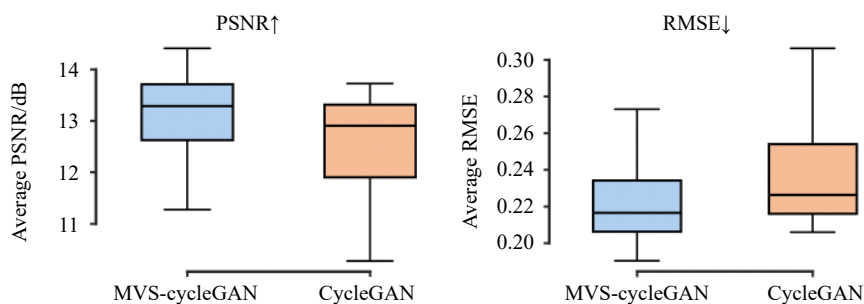
### 3.1.2 HEK293T 人胚肾细胞虚拟染色结果

在 HEK293T 数据集细胞膜与细胞核虚拟染色任务中,三个模型在虚拟染色上均实现较好的效果,本文方法在细胞膜染色上的 PSNR 为 20.72。在结构相似性相关指标上优势明显,细胞膜染色的 SSIM 达到 0.555 远高于 CycleGAN 的 0.485,显示其在保持细胞膜结构连续性和关键特征方面优于 CycleGAN。在细胞核虚拟染色任务中, MVS-CycleGAN 整体表现优于 CycleGAN。在 VIF 指标上,本方法在膜染色与核染色任务上分别取得了 0.0827 和 0.0433 的结果,均高于 CycleGAN 模型,说明 MVS-CycleGAN 在信息保持方面具有更加稳定的优势,能够更有效地恢复细胞膜轮廓与核结构细节。从视觉上看来,在细胞膜虚拟染色中, MVS-CycleGAN 避免了 CycleGAN 在部分图像生成中产生的伪影,增强了图像的真实感和可读性。对于细胞核的虚拟染色, MVS-CycleGAN 和 CycleGAN 模型都存在对少部分细胞核染色不清晰的现象。

### 3.1.3 Neuromast 斑马鱼神经丘虚拟染色结果

Neuromast 斑马鱼神经丘细胞有着更高的结构复杂度和显著不同于体外培养细胞的组织形态,

因此对于模型在复杂样本上的适应能力提出了更高要求。在 Neuromast 细胞膜和细胞核虚拟染色实验中,本文方法在绝大多数量化评估指标方面都优于 CycleGAN 或至少不会出现大幅度波动。特别是在 SSIM 和 FSIM 这些结构感知指标上, MVS-CycleGAN 具有显著优势,细胞膜染色任务中 SSIM 达到 0.589 高于 CycleGAN 的 0.522,细胞核染色任务中 MVS-CycleGAN 在 FSIM 指标上取得了 0.740 的成绩,如图 5 所示,说明 MVS-CycleGAN 能够在复杂的组织结构里更好地保持细胞之间的联系以及整体几何结构;而在 LPIPS 感知指标上, MVS-CycleGAN 表现出更小的感知距离,尤其是细胞核的 LPIPS 仅为 0.253,证明了其生成的结果更加接近真实的荧光图片。视觉对比进一步印证了定量结果。从 Neuromast 虚拟染色图像来看,使用 MVS-CycleGAN 可以更好地恢复出整个神经丘的大致形状以及其内部不同层级的结构,而 CycleGAN 则会出现局部结构失真甚至信号弥散的情况,这也与 VIF 指标所示结果一致,表明 MVS-CycleGAN 在结构保真度与视觉自然性之间取得了更好的平衡。



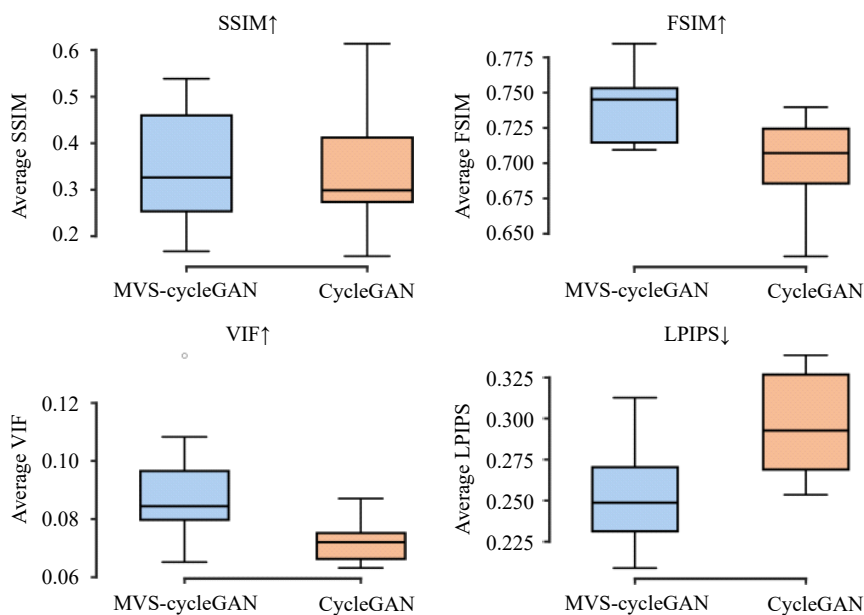


图 5 Neuromast 斑马鱼神经丘细胞核虚拟染色指标对比

Fig. 5 Comparison of virtual staining indices of Zebrafish Neuromast cell nuclei

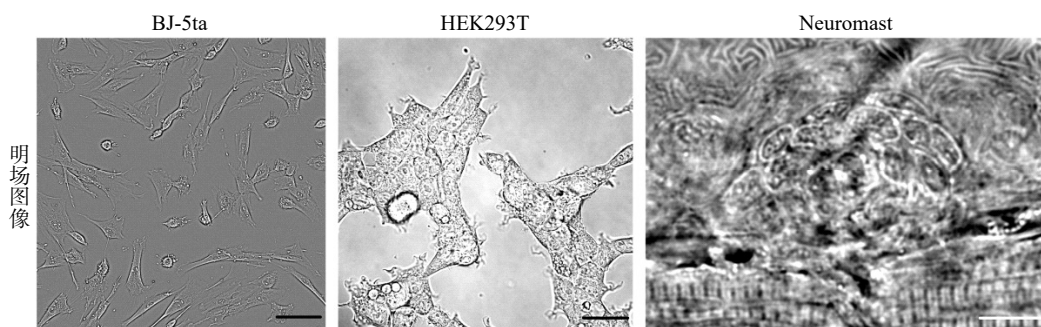
### 3.2 掩码自监督机制对虚拟染色性能的影响

综合不同细胞类型和染色目标的实验结果可以看出,所提出的掩码自监督机制在提升虚拟染色质量方面发挥了关键作用。我们在训练阶段将输入图像的一部分掩码处理,然后让生成器在少量上下文条件下填充被遮盖的部分,因此 MVS-CycleGAN 更关注细胞轮廓、边界和整体结构等形态学信息,从而在一定程度上缓解了传统 CycleGAN 在无配对转换中易出现的结构漂移问题。

总体而言, MVS-CycleGAN 在大部分指标上相比 CycleGAN 都有稳定提升,尤其是结构相关指标 (SSIM、FSIM) 和感知指标 (LPIPS),说明本文提出的基于掩码的自监督约束有利于产生更加符合真实荧光结构分布的虚拟染色结果。

### 3.3 虚拟染色对分割任务的提升

虚拟染色中的一个应用价值是帮助后续的定量分析,比如分割以及检测。本文采用 Cellpose3.0 模型作为分割模型,对比了明场直接分割与虚拟染色之后的分割效果,证明虚拟荧光图像对分割任务的有效性。图 6 为 BJ-5ta 细胞、HEK293T 细胞和 Neuromast 斑马鱼神经丘上的细胞膜和细胞核分割结果。该结果显示,虚拟染色后分割的结果能刻画出大部分细胞的膜以及核信息,因此本研究提出的 MVS-CycleGAN 框架在未配对数据条件下实现了高质量的虚拟荧光染色,不仅在可视化上再现了荧光显微成像的细节和对比,而且能够为后续分割任务提供更清晰的结构信息。同时,虚拟染色图像能够显著改善自动分割等定量任务的效果,表明了本方法对于真实生物图像处理任务中可能的价值。



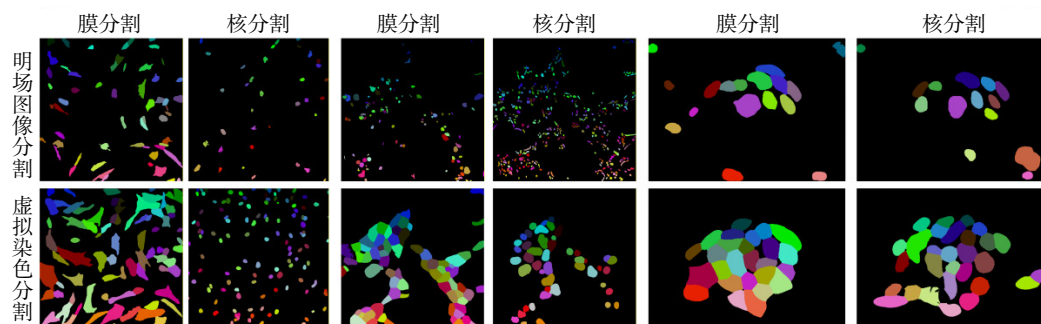


图 6 细胞明场图像及虚拟染色图像细胞分割结果。比例尺: 50  $\mu\text{m}$ (BJ-5ta 细胞, HEK293T 细胞); 10  $\mu\text{m}$ (Neuromast 细胞)  
Fig. 6 Cell segmentation results from bright-field and virtual staining images. scale bar: 50  $\mu\text{m}$ (BJ-5ta cells, HEK293T cells); 10  $\mu\text{m}$ (Zebrafish Neuromast)

## 4 结 论

本文提出了一种融合掩码自监督学习的无监督虚拟染色框架 MVS-CycleGAN, 用于实现明场图像到荧光图像的转换。该方法在 CycleGAN 的基础上引入掩码重建任务, 增强了模型对细胞形态结构信息的关注, 在无需严格配准数据的条件下提升了虚拟染色结果的结构保真度和图像质量。实验结果表明, 本文方法在多个数据集上均取得了较好的虚拟染色效果。以 FSIM 指标为例, 在 BJ-5ta 细胞膜与细胞核任务上分别达到 0.784 和 0.565, 在 HEK293T 细胞膜与细胞核任务上分别达到 0.854 和 0.830, 在 Neuromast 细胞膜与细胞核任务上分别达到 0.657 和 0.740, 整体结果优于传统 CycleGAN。与此同时, 下游分割实验表明, 虚拟染色图像能够为细胞分割等任务提供更清晰的结构信息, 说明该方法具有一定应用潜力。

虽然本文提出的 MVS-CycleGAN 在多个数据集上取得了较好的虚拟染色结果, 并在一定程度上提升了生成图像的结构保真度和图像质量,

但该方法本质上仍建立在 CycleGAN 这一无监督图像转换框架之上。因此与传统 CycleGAN 类似, 当测试样本与训练样本差异较大时, 模型的性能仍会下降。这表明, 尽管引入的掩码自监督机制加强了对细胞形态特征的关注, 但当前方法尚未从根本上克服无监督虚拟染色中模型泛化能力有限的问题。

针对上述不足, 未来的研究可以从以下几个方向寻求性能的进一步突破。首先, 可引入少样本微调或域自适应策略, 在仅使用少量目标域样本的情况下对模型进行快速调整, 从而提高其对新细胞类型或新成像条件的适应能力。其次, 可以探索引入更强的结构一致性约束, 例如结合边缘一致性、感知损失或语义结构保持机制, 以减轻无监督图像转换过程中可能出现的结构漂移。此外, 通过集成注意力机制、采用更强大的特征提取主干网络, 或结合其他自监督预训练策略, 增强模型对跨域稳定形态特征的学习能力, 从而提升其在复杂样本条件下的表现。通过上述改进, 有望进一步提高无监督虚拟染色方法在不同生物样本和成像场景中的实用性。

### 参考文献:

- [1] 高歌, 郭晓光, 吴俊楠, 等. 用于单切片双模态光学关联成像的肾脏组织样本处理方法[J]. 中国光学(中英文), 2024, 17(5): 1227-1235.  
GAO G, GUO X G, WU J N, *et al.*. Methods for processing renal tissue samples for single-slice dual-mode optical correlation imaging[J]. *Chinese Optics*, 2024, 17(5): 1227-1235. (in Chinese).
- [2] 王鹏, 周瑶, 赵宇轩, 等. 用于多尺度高分辨率三维成像的双环光片荧光显微技术[J]. 中国光学(中英文), 2022, 15(6): 1321-1331.  
WANG P, ZHOU Y, ZHAO Y X, *et al.*. Double-ring-modulated light sheet fluorescence microscopic technique for multi-scale high-resolution 3D imaging[J]. *Chinese Optics*, 2022, 15(6): 1321-1331. (in Chinese).
- [3] KUMAR A, MCNALLY K E, ZHANG Y X, *et al.*. Multispectral live-cell imaging with uncompromised spatiotemporal

- resolution[J]. *Nature Photonics*, 2025, 19(10): 1146-1156.
- [4] XIANG D, WANG ZH CH, ZHENG H W, *et al.*. Organic small-molecule NIR-II fluorophores for tumor phototheranostics[J]. *Light: Science & Applications*, 2026, 15(1): 173.
- [5] CHRISTIANSEN E M, YANG S J, ANDO D M, *et al.*. *In silico* labeling: predicting fluorescent labels in unlabeled images[J]. *Cell*, 2018, 173(3): 792-803. e19.
- [6] HOU Y W, WANG W Y, FU Y ZH, *et al.*. Multi-resolution analysis enables fidelity-ensured deconvolution for fluorescence microscopy[J]. *eLight*, 2024, 4(1): 14.
- [7] SHAKED N T, BOPPART S A, WANG L V, *et al.*. Label-free biomedical optical imaging[J]. *Nature Photonics*, 2023, 17(12): 1031-1041.
- [8] WANG Q, AKRAM A R, DORWARD D A, *et al.*. Deep learning-based virtual H& E staining from label-free autofluorescence lifetime images[J]. *npj Imaging*, 2024, 2(1): 17.
- [9] 黄宇然, 张智敏, 董婉洁, 等. 多色虚拟荧光辐射差分显微成像[J]. 中国光学(中英文), 2022, 15(6): 1332-1338.  
HUANG Y R, ZHANG ZH M, DONG W J, *et al.*. Multi-color virtual fluorescence emission difference microscopy[J]. *Chinese Optics*, 2022, 15(6): 1332-1338. (in Chinese).
- [10] COMBS C A, SHROFF H. Fluorescence microscopy: a concise guide to current imaging methods[J]. *Current Protocols in Neuroscience*, 2017, 79: 2.1. 1-2.1. 25.
- [11] SEO J, SIM Y, KIM J, *et al.*. PICASSO allows ultra-multiplexed fluorescence imaging of spatially overlapping proteins without reference spectra measurements[J]. *Nature Communications*, 2022, 13(1): 2475.
- [12] PIRONE D, BIANCO V, MICCIO L, *et al.*. Beyond fluorescence: advances in computational label-free full specificity in 3D quantitative phase microscopy[J]. *Current Opinion in Biotechnology*, 2024, 85: 103054.
- [13] YIN Z CH, HE B, YING Y ZH, *et al.*. Fast and label-free 3D virtual H&E histology via active phase modulation-assisted dynamic full-field OCT[J]. *npj Imaging*, 2025, 3(1): 12.
- [14] KREISS L, JIANG SH W, LI X, *et al.*. Digital staining in optical microscopy using deep learning - a review[J]. *Photonix*, 2023, 4(1): 34.
- [15] ZHANG Y J, HUANG L ZH, PILLAR N, *et al.*. Pixel super-resolved virtual staining of label-free tissue using diffusion models[J]. *Nature Communications*, 2025, 16(1): 5016.
- [16] ICHITA M, YAMAMICHI H, HIGAKI T. Virtual staining from bright-field microscopy for label-free quantitative analysis of plant cell structures[J]. *Plant Molecular Biology*, 2025, 115(1): 29.
- [17] KAMATH V, BHAT V G, RAJU G, *et al.*. Application of fluorescence lifetime imaging-integrated deep learning analysis for cancer research[J]. *Light: Advanced Manufacturing*, 2025, 6(3): 49.
- [18] OUNKOMOL C, SESHAMANI S, MALECKAR M M, *et al.*. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy[J]. *Nature Methods*, 2018, 15(11): 917-920.
- [19] PARK E, MISRA S, HWANG D G, *et al.*. Unsupervised inter-domain transformation for virtually stained high-resolution mid-infrared photoacoustic microscopy using explainable deep learning[J]. *Nature Communications*, 2024, 15(1): 10892.
- [20] DAI W X, WONG I H M, WONG T T W. Exceeding the limit for microscopic image translation with a deep learning-based unified framework[J]. *PNAS Nexus*, 2024, 3(4): 133.
- [21] MA J B, LI W Q, LI J B, *et al.*. Generative AI for misalignment-resistant virtual staining to accelerate histopathology workflows[J]. arXiv: 2509.14119, 2025. (查阅网上资料, 请核对文献类型及格式是否正确).
- [22] ZHU J Y, PARK T, ISOLA P, *et al.*. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, 2017: 2242-2251.
- [23] LI X Y, ZHANG G X, QIAO H, *et al.*. Unsupervised content-preserving transformation for optical microscopy[J]. *Light: Science & Applications*, 2021, 10(1): 44.
- [24] LIU Z W, HIRATA-MIYASAKI E, PRADEEP S, *et al.*. Robust virtual staining of landmark organelles with Cytoland[J]. *Nature Machine Intelligence*, 2025, 7(6): 901-915.
- [25] HE K M, CHEN X L, XIE S N, *et al.*. Masked autoencoders are scalable vision learners[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2022: 15979-15988.
- [26] PANG SH Y, XIANG J W, ZUO ZH Q, *et al.*. Contrastive masked feature modeling for self-supervised representation learning of high-resolution remote sensing images[J]. *Remote Sensing*, 2026, 18(4): 626.

- [27] WANG ZH, BOVIK A C, SHEIKH H R, *et al.*. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [28] ZHANG L, ZHANG L, MOU X Q, *et al.*. FSIM: a feature similarity index for image quality assessment[J]. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378-2386.
- [29] ZHANG R, ISOLA P, EFROS A A, *et al.*. The unreasonable effectiveness of deep features as a perceptual metric[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2018: 586-595.
- [30] SHEIKH H R, BOVIK A C, DE VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. *IEEE Transactions on Image Processing*, 2005, 14(12): 2117-2128.

作者简介:

林俊豪(2000—),男,安徽铜陵人,硕士研究生,2022年于福建农林大学获得学士学位,主要从事计算显微成像方面的研究。E-mail: [202312490675@nuist.edu.cn](mailto:202312490675@nuist.edu.cn)

张国勋(1997—),男,河南新乡人,博士,2024年于清华大学获得博士学位,主要从事智能增强显微成像方面的研究。E-mail: [zgx31415@gmail.com](mailto:zgx31415@gmail.com)

谢浩(1988—),男,浙江杭州人,博士,副研究员,博士生导师,2017年于北京大学获得博士学位,主要从事光学显微成像方面的研究。E-mail: [xiehao@iphy.ac.cn](mailto:xiehao@iphy.ac.cn)