

融合通道层注意力机制的 UNet 的衍射极限荧光点检测和定位

余永建¹, 王越^{2**}, 李寰², 周文超², 舒风风², 高明^{2,3}, 吴一辉^{1,2*}

¹温州医科大学眼视光学院, 浙江 温州 325035;

²中国科学院长春光学精密机械与物理研究所光学系统先进制造技术重点实验室, 吉林 长春 130033;

³中国科学院大学, 北京 100049

摘要 针对高通量荧光显微成像中高密度、低信噪比、亚衍射极限荧光斑点的自动化精准检测和定位问题, 基于 UNet 提出一种轻量级神经网络方法。该方法采用挤压和激发通道层注意力机制和残差模块优化特征信息, 构建密度图和偏移量多输出架构, 直接执行检测和亚像素定位。在公开数据集和模拟数据集进行实验, 所提方法对低信噪比和高密度的荧光点检测优于当前算法, 尤其对于达到衍射极限的高密度荧光点, 有很好的检测性能, 比如在 128×128 像素具有 1200 个荧光点并且大部分点达到衍射极限的图像下。所提算法对斑点的识别精度 F1 分数超过 97.6%, 定位误差为 0.115 pixel, 相比最新 deepBlink 方法, F1 提升 16.2 个百分点并且定位误差减小 0.63 pixel。

关键词 荧光显微镜; 数字图像处理; 模式识别; 神经网络; 医学和生物成像

中图分类号 TP751.1

文献标志码 A

DOI: 10.3788/LOP230718

Channel-Wise Attention Mechanism Relevant UNet-Based Diffraction-Limited Fluorescence Spot Detection and Localization

Yu Yongjian¹, Wang Yue^{2**}, Li Huan², Zhou Wenchao², Shu Fengfeng², Gao Ming^{2,3},
Wu Yihui^{1,2*}

¹School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou 325035, Zhejiang, China;

²Key Laboratory of Optical System Advanced Manufacturing Technology, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, Jilin, China;

³University of Chinese Academy of Sciences, Beijing 100049, China

Abstract This paper proposes a lightweight neural network method based on UNet to accurately detect and localize high-density, low signal-to-noise ratio (SNR) sub-diffraction fluorescence spots in high-throughput fluorescence microscopy imaging. This method combines a squeeze and excitation channel-wise attention mechanism with a residual module to optimize feature information. A density map and offset multioutput architecture are also constructed for direct detection and subpixel localization. The proposed method has been verified on public and simulated datasets, and outperforms current algorithms for low SNR and high-density fluorescent spot detection. Notably, the detection performance of the proposed method is excellent for high-density fluorescent spot that reaches the diffraction limit, such as in images with a resolution of 128×128 pixels having 1200 fluorescent spots. The spot detection accuracy (F1 score) of the proposed algorithm exceeds 97.6%, and the localization error is 0.115 pixel. Compared with the latest deepBlink method, the F1 of the proposed algorithm has improved by 16.2 percentage points, and the localization error has been reduced by 0.63 pixel.

Key words fluorescence microscope; digital image processing; pattern recognition; neural network; medicine and biological imaging

收稿日期: 2023-02-27; 修回日期: 2023-03-31; 录用日期: 2023-04-17; 网络首发日期: 2023-04-27

基金项目: 国家自然科学基金(U21A20395)

通信作者: *yihuiwu@ciomp.ac.cn; **wangyue@ciomp.ac.cn

1 引言

在生物医学研究中,单分子荧光显微成像在亚细胞成分动态过程可视化^[1]、基因测序^[2]和单分子荧光原位杂交(FISH)^[3]等方面是不可替代的研究手段。荧光标记的分子在荧光图像中表现为荧光点,对荧光点的检测、计数和定位是后续图像分析的基本步骤。随着单分子荧光显微镜向高通量方向发展,不仅促使成像系统视场(FOV)的扩大,也促使分辨率的提升,这反过来在准备过程中可以采用更少量的荧光分子制备更小的生物样品。因此,这产生了具有高密度、低信噪比(SNR 小于等于 5)^[4]和小尺寸的荧光斑点图像。在具有不同背景亮度水平和斑点质量的单分子荧光显微镜图像中进行准确、高通量的斑点检测和定位,尤其是高密度相互重叠达到衍射极限的斑点检测,对当前的斑点检测方法提出了挑战。

在过去二十年里,为了克服荧光显微成像中的各种挑战,许多传统的和基于深度学习的斑点识别算法被提出,并追求高检测精度。传统的斑点检测技术是基于预定义和简单外观模型通过减少噪声和增强信号执行检测的,例如点增强滤波器^[5];基于 HDome 变换的检测器^[6];基于最佳尺寸选择的方法(ATLAS)^[7],用于检测显微镜成像中的囊泡;基于马尔可夫随机场的逻辑回归方法(LR-MRF)^[8]。这些方法具有特定的应用优势,但它们鲁棒性和泛化能力受限于原始图像的固定特征提取。最近,基于深度学习的方法在检测领域获得极大的关注,因为它们提高了准确性、自动化和鲁棒性^[9-14]。

基于深度学习的斑点检测方法通过使用卷积核以数据驱动的方式自动提取目标特征,克服了传统方法的限制。斑点检测方法通常由特征金字塔网络(FPN)^[15]、基于 UNet 的网络^[16]和其他简单网络^[17-18]组成。其他简单网络采用较为简单的网络架构在局部小区域中判定单个斑点,利用滑动窗口方法完成对整个图像多个斑点的检测。这些方法适合于固定尺寸范围的荧光斑点并且使用滑动窗口,需要大量的计算量。基于 FPN 的网络融合多尺度特征得到高分辨率、强语义的特征来确定斑点^[19-20]。基于 UNet 的网络采用编码器-解码器结构获取上下文信息和位置信息来执行斑点识别,它因可扩展性和直接的图像到图像映射被广泛应用,甚至应用于三维荧光斑点数据集^[21-22]。例如,为了自动适应不同 DNA FISH 数据集,SpotLearn 网络利用基于 UNet 的算法^[23]在 DNA FISH 斑点数据集进行检测和分类任务^[24]。为了减轻网络的负担,DetNet 基于 UNet 使用残差主干优化信息流,并去除跳跃连接,减少了参数量^[25]。为了实现在可变强度和高噪声水平 RNA 斑点图像下的斑点检测,DeepSpot 网络针对小目标通过设计上下文融合模块学习全面、互补的特征,并使用残差卷积优化这些特征信息^[26]。

为了在不同背景亮度水平和斑点质量的单分子荧光显微镜图像中进行准确、高通量的斑点检测和定位,deepBlink 方法结合 UNet 和注意力机制,直接结合荧光点的概率和荧光点的偏移输出检测荧光点^[27],具有出色的检测和定位能力。为了提高 UNet 的性能,许多功能模块与 UNet 相结合,探索提高 UNet 架构网络性能的可能性。最近结合注意力机制和残差模块的 UNet 被提出并在图像分割领域得到精度验证^[28-30]。

本文基于融合通道层注意力的 UNet 提出一种用于荧光斑点检测的方法,即 SE-Res-UNet,其可以直接执行斑点识别和定位。采用挤压和激发(SE)的注意力机制和残差块优化特征信息;此外构建密度图和偏移量的多输出结构,直接精确执行亚像素斑点定位;将损失函数修改为自适应机翼(AWing)损失和加权绝对误差(AE)损失的组合,在 AWing 专注于前景的基础上,通过加权 AE 更关注荧光中心,能够更好地量化网络预测图像和标注图像之间的差异。SE-Res-UNet 将更加专注于斑点周围尤其斑点中心的训练。通过在多个数据集上对 SE-Res-UNet 与最先进的斑点检测方法进行基准测试,结果表明,SE-Res-UNet 在高密度和低信噪比下具有优秀的斑点检测和定位性能。

2 材料和方法

2.1 网络结构

设计一个基于 UNet 的卷积神经网络架构。UNet 作为生物医学图像处理的经典架构之一,由收缩(用于特征提取的编码器)和扩展(用于图像重建的解码器)路径组成,具有跳跃连接,增强了特征信息,缓解重复池化造成的空间信息损失问题。图 1(a)为针对斑点检测设计的模型结构,它能根据真实坐标生成标注图像,计算预测图像和标注图像之间的损失并通过反向传播优化网络权重。它在收缩和扩展路径中增加 SE 注意力模块,通过对特征通道进行加权,突出有效特征,并使用残差模块替代网络主干中的原始卷积块;添加 dropout 操作,防止小样本数据的过拟合;同时,在扩展路径中使用最近邻插值的反卷积操作,避免棋盘伪影^[31]。整体网络减少了提取的特征图的数量,大大减少了参数的数量。该网络只有大约 21 万参数。此外,密度图和偏移图作为网络的输出用于直接和快速的亚像素定位。

所提网络进行三通道输出,其大小与输入图像相同,如图 1(b)所示,利用多通道输出描述荧光点相关信息,以便准确检测和精确定位。第一个通道为密度图,表示斑点中心存在某像素的概率 p ,在网络训练中,这能避免高度非线性地直接预测点坐标的问题。另外两个通道描述该像素相对于最近的真实点中心坐标 (x, y) 的偏移量。为了进一步精确定位,利用局部极大值方法获取密度图中斑点的像素坐标 $(\Delta x, \Delta y)$,通过整合偏移量计算出斑点的亚像素坐标 $(\Delta x + x,$

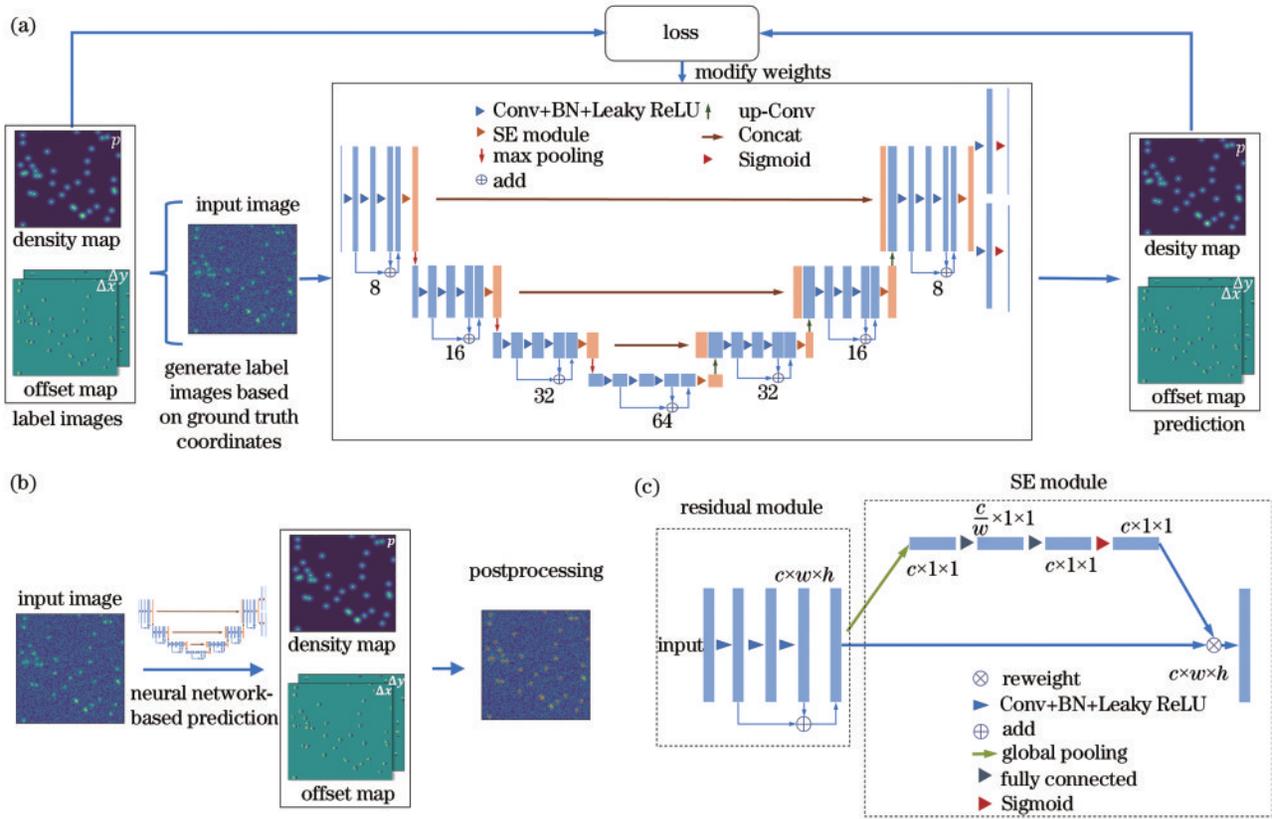


图1 所提网络的整体结构。(a)训练网络;(b)网络预测过程;(c)SE残差模块的结构

Fig. 1 Overall architecture of our network. (a) Training network; (b) network prediction process; (c) structure of the SE residual module

$\Delta y + y$)。

在低信噪比下,背景和光斑之间的边界模糊,难以检测斑点。为了加强算法在噪声数据中的有效自适应能力,采用注意力机制来筛选特征^[32]。具体来说,在收缩和扩展路径中加入通道注意力机制和残差模块,如图1(c)所示,利用残差模块优化提取的特征,随后用SE模块筛选特征信息。输入维度为 $c \times w \times h$,分别代表特征图的通道数、宽度和高度。首先,残差模块使用跳跃连接将浅层特征与网络的深层特征结合,通过重用浅层特征增强图像细节;之后执行SE模块,对SE模块进行一个改进,将通道的注意力权重限制在0.3和1.0之间。因此最低权重的通道被抑制,而不是完全被消除,网络有可能在下一层重新使用它们。

残差运算执行三次卷积,并将第一次和第三次卷积结果总和作为SE模块的输入,其维度为 $c \times w \times h$;之后执行SE操作,其由三部分组成,即挤压、激发和重新加权。挤压操作使用全局平均池化,将每个二维特征通道压缩成具有全局感受野的权重,输出维度为 $c \times 1 \times 1$ 。激发操作利用两个全连接操作将维度 $c \times 1 \times 1$ 减小到 $c/w \times 1 \times 1$,然后恢复到维度 $c \times 1 \times 1$,通过降维操作更好地描述特征通道之间的关联性。最后,采用Sigmoid变换将权重限定在0和1之间。重新加权操作将激发操作中输出的权重逐通道加权到原始特

征,这抑制了大量的无效特征信息,提高了斑点检测性能。

2.2 损失函数

AE量化预测图像和标注图像之间的差异。由于在标注图像中斑点和背景之间的像素数量严重不平衡,损失会对背景区域更加敏感。因此,对背景、斑点周围和斑点中心引入不同的权重解决这种不平衡,表达式为

$$E_A = \sum_{i=1}^I W_i |y^{gt} - y^{pred}|, \quad (1)$$

式中: I 表示类别,即背景、斑点周围和斑点中心; W_i 代表类别 i 的权重; y^{gt} 是标注图像; y^{pred} 是网络预测图像。AWing损失减少前景像素上的小误差,准确定位荧光点,同时容忍背景像素的小误差以获得更好的收敛速度,表达式为

$$f_{\text{AWing}} = \begin{cases} \omega \ln \left(1 + \frac{|y^{gt} - y^{pred}|}{\epsilon} \right), & |y^{gt} - y^{pred}| < \theta \\ A |y^{gt} - y^{pred}| - C, & |y^{gt} - y^{pred}| \geq \theta \end{cases}, \quad (2)$$

$$A = \omega \left\{ 1 / \left[1 + \left(\frac{\theta}{\epsilon} \right)^{\alpha - y^{gt}} \right] \right\} \left[(\alpha - y^{gt}) \left(\frac{\theta}{\epsilon} \right)^{\alpha - y^{gt} - 1} \right] \left(\frac{1}{\epsilon} \right), \quad (3)$$

$$C = \left\{ \theta A - \omega \ln \left[1 + \left(\frac{\theta}{\epsilon} \right)^{\alpha - y^{\text{gt}}} \right] \right\}, \quad (4)$$

式中: ω 、 θ 、 ϵ 和 α 为正数, 注意, α 必须略大于 2, θ 是阈值; ω 是权重。在本文的实验中, 使用 $\omega=14$, $\epsilon=1$, $\theta=0.5$ 。将密度图损失函数定义为 AWing 和 AE 损失的总和。AWing 对拟合密度图十分有效, 但是对中心位置像素拟合效果较差, 因此添加 AE 突出中心区域。由于偏移量的输出为 0 和 1 之间, 直接使用 AE 作为损失函数。

2.3 评价指标

算法性能由检测性能和定位性能组成。它们能衡量荧光斑识别的准确性和预测坐标与真实坐标之间的误差。在本文中, 预测坐标和真实坐标的偏差在 2 个像素以内认为是正确匹配的^[25], 否则是错误匹配。

在检测斑点时会出现三种情况: 假阳性 (FP), 预测为阳性的阴性样本; 假阴性 (FN), 预测为阴性的阳性样本; 真阳性 (TP), 预测为阳性的阳性样本。使用 F1 分数来综合评估识别精度和错误识别率, 表达式为

$$F_1 = \frac{2 \times P_{\text{precision}} \times P_{\text{recall}}}{P_{\text{precision}} + P_{\text{recall}}}, \quad (5)$$

式中: $P_{\text{precision}}$ 用来预测正确结果在预测总数中的比例, 即 $N_{\text{TP}} / (N_{\text{TP}} + N_{\text{FP}})$; P_{recall} 用来预测准确结果在整个正面样本中的比例, 即 $N_{\text{TP}} / (N_{\text{TP}} + N_{\text{FN}})$ 。采用均方根误差 (RMSE) 来衡量算法的定位误差, 表达式为

$$E_{\text{RMSE}} = \sqrt{\frac{1}{N_{\text{TP}}} \sum_{j=1}^{N_{\text{TP}}} (y_j^{\text{gt}} - y_j^{\text{pred}})^2}, \quad (6)$$

式中: y_j^{gt} 是斑点的真实坐标; y_j^{pred} 是斑点的预测坐标。

2.4 实验环境

一台配备 Intel (R) Xeon (R) Platinum 8358P 处理器 (15 核)、80 GB 内存和 RTX A5000 GPU 的在线服务器用于模型训练。具有 Intel Core i5 处理器、16 GB 内存和 RTX1050TI GPU 的台式机用于模型预测。使用支持 GPU 的 keras API 搭建模型。选择 Adam 作为本模型的优化器。设定的学习率会随着迭代次数的增加逐渐降低, 以找到全局最优值。

3 实验结果

为了评估所提算法的性能, 与两种公开先进的方

法进行比较, 即 deepBlink 和 Big-FISH^[33]。其中 Big-FISH 是一个用于分析单分子荧光原位杂交 (smFISH) 图像的 Python 软件包, 它能够进行斑点检测和细胞与细胞核的分离。deepBlink 依靠神经网络方法自动检测和定位斑点。提供不同尺寸、不同密度模拟数据及公开数据集来验证算法性能, 如 SunTag 活细胞图像^[34]、受体 (receptor) 和囊泡 (vesicle), 它们来自粒子跟踪挑战数据集 (ISBI)^[35]。

首先验证引入注意力、残差模块和优化损失函数的有效性, 之后在公开和模拟多数据集上验证所提模型的性能和泛化能力。

3.1 损失函数和相关模块的有效性

由于本文架构具有多个输出, 对于密度图和偏移图输出, 要分别计算误差, 对于偏移图 (值范围为 $-2 \sim 2$) 采用 AE 评估误差, 对于密度图采用 AE 和 AWing 结合的方式评估误差。为了验证当前损失函数的影响, 对密度图使用不同的计算方式来验证密度图损失函数的有效性, 即 AE 和 AWing 的组合、AWing、BCE。表 1 给出了 SE-Res-UNet 模型 ($C_{\text{initConvNum}}=8$, $C_{\text{initConvNum}}$ 是初始输出空间维度) 对 3 种不同损失函数的检测性能, 最佳性能用加粗标识。结果显示: 使用二元交叉熵 (BCE) 和 AWing 损失会导致次优的检测性能, 虽然 BCE 取得更高 precision 指标, 但是在总体样本下 recall 指标低, 这造成样本检测性能较差; 所提组合损失中, 重点突出点中心区域, 可以提高检测和定位性能, SE-Res-UNet 通过使用加权平方误差来优化密度图中的斑点中心区域, 进一步提升了斑点检测精度。

表 1 使用不同密度图损失函数的 SE-Res-UNet 的性能

Table 1 Performance of SE-Res-UNet using different density map loss functions

Loss	F1 / %	RMSE / pixel	Recall / %	Precision / %
Our loss	84.6	0.545±0.357	82.4	88.4
AWing loss	83.0	0.564±0.348	80.1	89.0
BCE loss	80.2	0.539±0.337	76.9	94.0

进一步全面地验证量化残差主干和 SE 模块的有效性。随机选择一部分囊泡数据集进行验证。UNet 模型的性能作为检测精度和参数数量的基准。表 2 为不同模型的检测性能, 包括 SE-Res-UNet、Res-UNet、

表 2 不同模型在囊泡数据集上的性能

Table 2 Performance of different models on vesicle dataset

Model	F1 / %	RMSE / pixel	Recall / %	Precision / %	Parameter quantity
SE-Res-UNet (16)	84.5	0.549±0.356	82.7	88.1	772539
SE-Res-UNet (8)	84.3	0.542±0.354	81.6	89.2	210271
Res-UNet (16)	83.7	0.543±0.369	83.2	84.8	743311
Res-UNet (8)	83.9	0.548±0.354	81.4	88.2	758531
SE-UNet (8)	84.0	0.557±0.351	81.1	89.3	141615
UNet (8)	83.4	0.570±0.351	81.9	86.1	138003

SE-UNet和UNet。UNet实现83.4%的F1和0.570 pixel的RMSE,参数量最少,但性能较差。本文通过增加SE模块和残差模块来修改UNet的结构,以评估网络性能。SE-UNet和Res-UNet的性能都优于UNet,F1最少提升0.5个百分点,定位误差减小0.013 pixel,这表明SE模块和残差模块能帮助提升对荧光点进行识别的轻量级网络的性能。之后,将SE模块添加到调整后的Res-UNet中,结果表明,SE-Res-UNet($C_{initConvNum}=8$)和SE-Res-UNet($C_{initConvNum}=16$)模型都明显提高了检测性能,F1分别从83.9%增加到84.3%和83.7%增加到84.5%。基于这一观察可以推断,SE和残差模块组合可有效地提高对斑点的检测精度。虽然Res-UNet($C_{initConvNum}=16$)和SE-UNet($C_{initConvNum}=8$)取得最好的precision指标和recall指标,但是荧光点识别性能要综合衡量precision和recall指标(理想情况下两者都高最佳)。相比SE-Res-UNet($C_{initConvNum}=16$),SE-Res-UNet($C_{initConvNum}=8$)的参数量更少,性能差别微弱。因此使用SE-Res-UNet($C_{initConvNum}=8$)作为本文网络。

SE-Res-UNet的高精度检测性能和较快的收敛速度可能与调整后的残差主干和SE模块的使用有关。网络主干设置初始卷积数为8,以抑制参数的指数增长,并尽量减少内存的使用。此外,它还使用残差块来改善信息和梯度流,从而使模型能够快速学习。总之,SE-Res-UNet对荧光信号有出色的检测性能和更快的运算速度。

3.2 在不同尺寸、密度和信噪比下的性能分析

在不同的荧光显微成像环境中,密集的荧光点和更少单个荧光点的像素数量是高通量的具体表现。此外,斑点的像素大小和密度也在一定范围内变化。因此,为了评估所提算法的有效性和适用性,使用模拟数据集评估算法在不同光斑尺寸和密度下的检测性能,即同一光斑密度下的不同像素数,其次是同一像素大小下的不同光斑密度。此外,在不同的图像中,信噪比的定义也不同:在稀疏斑点图像中,通常使用单个斑点的最大信噪比作为该图像的信噪比;对于高密度的斑点图像,信噪比是用整个图像的强度分布来估计的。因此,无法在模拟数据集下测量统一的信噪比,所以使用公开的数据集(ISBI)评估算法在不同信噪比下的检测能力。

本文使用相同参数生成一个模拟数据集,构建具有不同半径(点扩散函数半宽全高)和密度的荧光点图像。deepBlink、Big-FISH和所提算法对不同尺寸斑点的数据集的检测性能,如图2(a)所示。所提算法生成的斑点半径从1 pixel逐渐增加到7 pixel。对于检测到的半径为1~4像素的斑点,所有三种算法的检测性能即F1都高于99%。随着尺寸的增加,在相同的光子数下,光斑具有更低的信号(信噪比逐渐降低),所提算法的性能下降速度比deepBlink和Big-FISH慢,但是所提算法和deepBlink性能差异很微弱。deepBlink通过直接预测像素网格的荧光点概率、 x 和 y 坐标来确定坐标,所提算法通过结合密度图和偏移量来确定坐标。

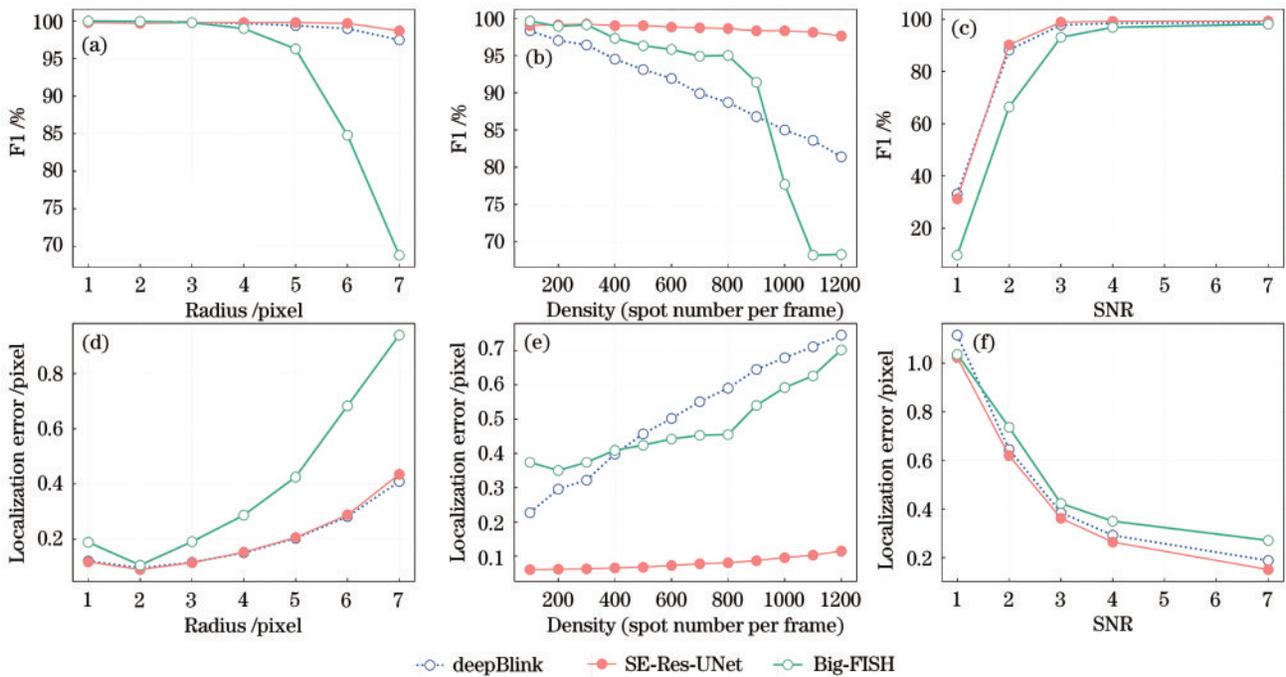


图2 在不同尺寸、密度和信噪比荧光点图像下算法的性能。(a)不同尺寸下的F1;(b)不同密度下的F1;(c)不同信噪比下的F1;(d)不同尺寸下的定位误差;(e)不同密度下的定位误差;(f)不同信噪比下的定位误差

Fig. 2 Algorithm performance under images with different sizes, densities, and SNRs. (a) F1 under different sizes; (b) F1 under different densities; (c) F1 under different SNRs; (d) localization error under different sizes; (e) localization error under different densities; (f) localization error under different SNRs

这个结果表明,两者都有很好的亚像素定位精度。

在低信噪比下,光斑和背景无法区分,边界模糊,这对算法提出更高要求。图 2(c)显示不同算法在低信噪比下的性能。在高信噪比(SNR 大于等于 4)下,这些算法的识别性能差异非常小;但是在 SNR 为 3 和 2 的情况下,如表 3 所示,所提算法的 F1 指标达 98.9% 和 90.2%,相比于 deepBlink 和 Big-FISH,相对 deepBlink F1 性能明显提高,最少提升 1.1 个百分点;

在 SNR 为 1 的情况下,所有算法识别精度较差。在定位误差方面,所有算法在不同信噪比下都具有较好的定位误差,所提算法相比 Big-FISH 最低减少 0.06 pixel(不包含 SNR 为 1 的情况,该信噪比下,所有算法识别精度过差,带来的误差均偏大),相比 deepBlink 至少减少 0.02 pixel。这进一步证明密度图和偏移量的结合对在低信噪比下斑点亚像素定位仍具有出色性能。

表 3 不同信噪比下的不同算法的检测和定位性能

Table 3 Detection and localization performances of different algorithms under different SNRs

SNR	SE-Res-UNet		deepBlink		Big-FISH	
	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel
1	31.2±13.3	1.023±0.134	33.1±11.6	1.115±0.076	9.6±3.4	1.036±0.148
2	90.2±8.1	0.620±0.075	88.2±8.1	0.645±0.083	66.4±16.1	0.736±0.108
3	98.9±0.6	0.362±0.023	97.8±1.0	0.385±0.031	93.1±2.1	0.423±0.033
4	99.2±0.6	0.264±0.028	98.5±0.9	0.292±0.039	96.9±2.1	0.350±0.069
7	99.3±0.6	0.151±0.017	98.7±0.7	0.188±0.038	98.1±1.4	0.271±0.075

进一步检测不同密度的斑点图像,在 128×128 像素图像(小图像可以产生更高密度的图像,可以使用较少算力)中以 2 像素作为斑点半径,逐渐生成 100 到 1200 的斑点数量。在检测性能方面,如图 2(b)所示,与所提算法相比,Big-FISH 和 deepBlink 在低密度(每幅图像少于 400 个斑点)下的检测精度只有微弱的差异;然而,所提算法在密度高于 400 时仍然保持较高的准确率;随着密度的增加,deepBlink 和 Big-FISH 的检

测精度迅速下降,所提算法一直保持出色的检测性能。在定位误差方面,如图 2(e)所示,所提算法的定位误差已经远远低于 deepBlink 和 Big-FISH。如表 4 所示,在最高的斑点密度下,相比最新 deepBlink 方法,所提算法的 F1 提升 16.2 个百分点并且 RMSE 减小 0.63 pixel。所提算法能在高密度的斑点图像中实现更精确的亚像素定位。

表 4 不同密度下的不同算法的检测和定位性能

Table 4 Detection and localization performances of different algorithms under different densities

Density	SE-Res-UNet		deepBlink		Big-FISH	
	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel
200	99.1±0.50	0.062±0.0041	97.0±0.98	0.296±0.0435	98.9±0.61	0.350±0.0715
400	99.0±0.30	0.066±0.0026	94.5±0.58	0.398±0.0193	97.3±1.10	0.409±0.0252
600	98.8±0.31	0.073±0.0037	91.9±0.70	0.502±0.0251	95.8±1.00	0.442±0.0185
800	98.6±0.35	0.081±0.0040	88.7±0.54	0.590±0.0209	95.0±1.50	0.455±0.0414
1000	98.3±0.24	0.096±0.0044	85.0±0.66	0.679±0.0152	77.7±7.79	0.592±0.0412
1200	97.6±0.25	0.115±0.0047	81.4±0.56	0.745±0.0104	68.3±11.81	0.702±0.0634

为了进一步探究所提算法在高密度荧光点下出色的性能,评估高密度图像性质和荧光点特性。分别获取预测图像[图 3(b)中方框区域]中相邻两个荧光点和模拟环境中达到衍射极限的两个荧光点的光强,折线图如图 3(a)所示,通过对比,测序数据中两个荧光点相邻距离达到衍射极限(通过计算,该图像中两个荧光点的衍射极限是 2.32 pixel,而预测图像中两个光点之间的距离是 2.29 pixel)。此外,当两个斑点相邻距离达到衍射极限时,荧光点重叠区域的强度高,这易将某点误判为一个斑点,这对算法检测提出极大挑战。

值,让其保持为 1,这突出荧光点的中心区域。并且在训练过程中,所提算法不仅注意对整个荧光点密度图的训练,更加聚焦对中心点的训练。在高密度检测中,虽然所提算法受到来自周围荧光点和噪声的很多干扰,但在本文的预测中,如图 3(b)所示,密度图正确地显示了荧光的峰值位置,这为后续的荧光检测和准确定位提供极大帮助。与其他算法相比,所提算法在高密度的荧光点上具有更高的检测准确性。

所提算法使用密度回归来检测荧光点及其周围的像素。在生成密度图标签数据时,增加峰值中心的数

图 3(c)显示了各算法对高通量图像的检查结果。很明显,所提算法具有较高的识别性能,即 F1 为 97.6%,而其他两种算法会遗漏部分斑点或者将两个荧光点判定为一个。此外,由于图 3(c)中斑点符合高

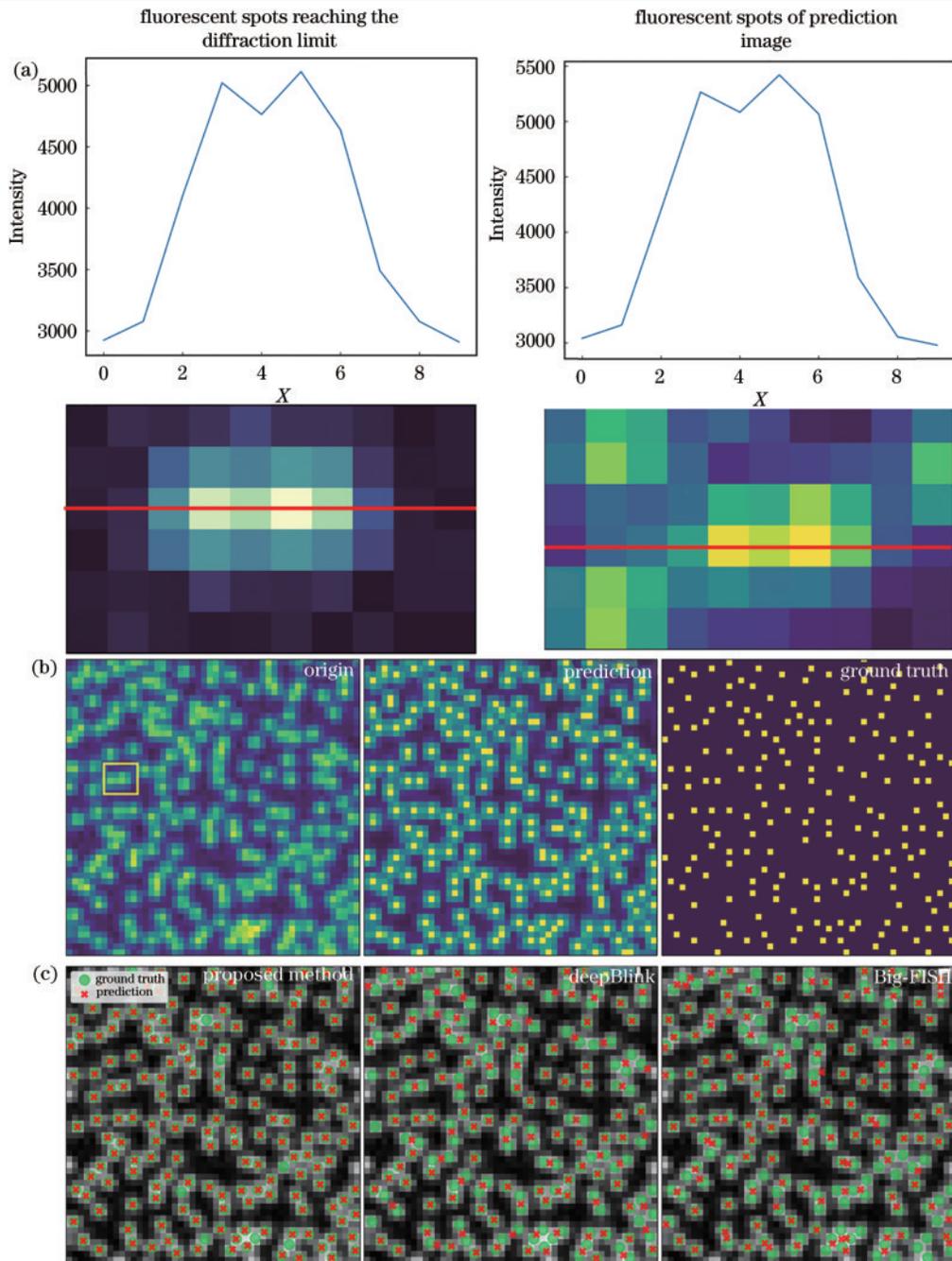


图3 高密度光斑图像的预测。(a)衍射极限下的荧光点;(b)基于密度图的预测结果;(c)在每幅图像有1200个斑点的情况下各种算法的检测结果

Fig. 3 Prediction for high-density spot images. (a) Fluorescent spot prediction reaching the diffraction limit; (b) prediction based on density image; (c) detection results of each algorithm in the case of 1200 spots per image

通量图像下的高密度和小尺寸要求,表明所提算法为高通量斑点检测提供了有效途径。

3.3 多数据集性能和预测时间结果分析

使用几个公开的数据集来验证所提算法的性能,包括真实数据集 SunTag 和模拟数据集,即囊泡、受体和高密度斑点数据,每个数据集的代表图像及其预测过程显示在图4中。所提算法对多个数据集都具有优秀的检测能力。进一步量化性能,如表5所示,显示所提算法在不同数据集下具有出色表现,在大部分数据集下检测精度和定位误差都优于 deepBlink 和 Big-

FISH。此外结果也说明所提算法能够适应多个数据集下不同光学系统的设置。

比较所有方法在一组样本图像上的执行时间, Big-FISH 的时间为 (121.1 ± 11.8) ms, 所提算法的时间为 (80.7 ± 8.4) ms, deepBlink 的时间为 (134.6 ± 7.9) ms, 发现所提算法相比 Big-FISH 和 deepBlink 具有更快的预测速度。同时,由于 Big-FISH 需要逐个图像调整参数,并对单个荧光点进行亚像素精度检测,执行速度进一步降低。

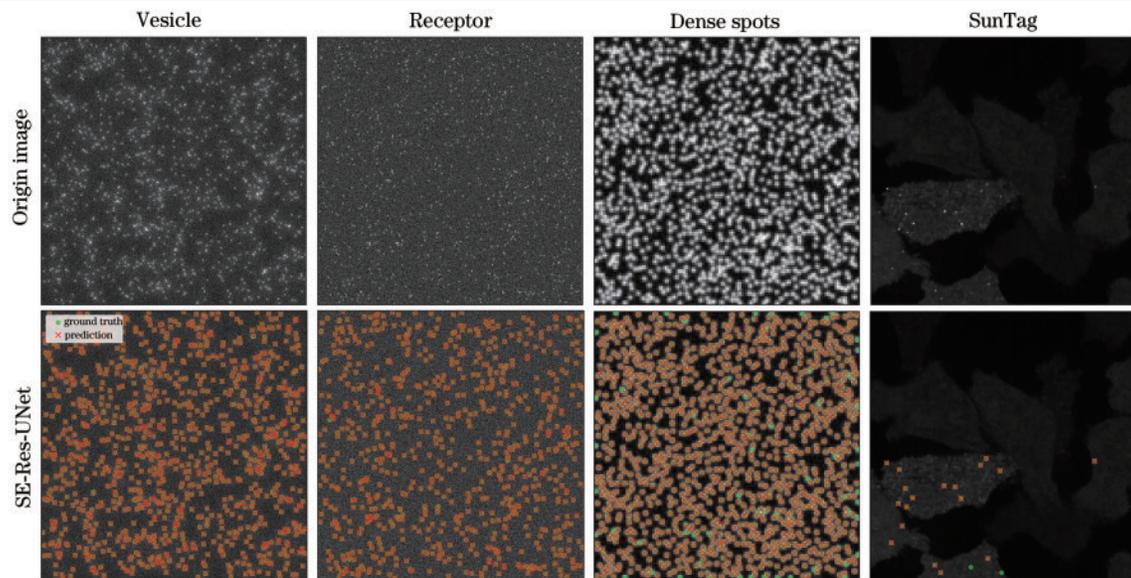


图 4 数据集的代表性图像及相应的预测

Fig. 4 Representative images of datasets and their corresponding prediction

表 5 不同算法对不同数据集的检测和定位性能

Table 5 Detection and localization performances of different algorithms on different datasets

Dataset	deepBlink		SE-Res-UNet		Big-FISH	
	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel	F1 / %	RMSE / pixel
SunTag	83.0	0.382 ± 0.289	85.4	0.595 ± 0.266	62.8	0.553 ± 0.298
Vesicle	82.9	0.577 ± 0.384	83.8	0.531 ± 0.363	67.8	0.654 ± 0.341
Receptor	80.6	0.512 ± 0.320	80.7	0.471 ± 0.305	72.8	0.519 ± 0.272
Dense spots	90.6	0.510 ± 0.167	98.6	0.079 ± 0.017	90.1	0.478 ± 0.118

4 结 论

提出一种基于融合通道层注意力机制的UNet的荧光图像斑点检测算法。该算法在荧光图像识别中受到高密度斑点和高噪声挑战,使用SE注意力机制模块和残差模块来优化特征信息。它通过构建密度图和偏移量的多输出架构直接执行亚像素定位。它可以同时进行准确的斑点检测和定位,在高密度、低信噪比斑点图像中具有出色的检测性能,尤其在高密度斑点图像中,其性能远远优于其他算法。

所提算法具有较快预测速度是因为两个方面:计算复杂度只取决于输入图像的大小,而不是图像中斑点的数量(后处理中局部极值需要一定算力);在密度图的像素坐标和偏移量之间进行简单加法运算,获得斑点亚像素坐标。所提算法可以像基于GPU的MLE拟合器一样快速预测坐标。例如,在目前的计算机配置下,所提算法只需约60 ms就能预测一幅 512×512 像素的荧光图像。在训练时间方面,通过使用Batch Normalization^[36]和减少训练参数量,所提算法只用了3 h就得到了一个高性能模型。此外,通过对现有网络进行微调^[37],可以进一步减少训练时间。

荧光斑检测和定位是定量的第一步,本文神经网络架构用于准确检测单次采集图像上的荧光斑点,并

可广泛应用于生物医学图像分析的其他领域。例如,在FISH研究中,对细胞核中染色体、结构域和特定基因组的位置和组织的荧光点进行定位和检测是不可缺少的,它可以对能量强度进行精确定位和量化^[38-39]。为了量化亚细胞和病毒结构的动态行为,由于光学系统的空间分辨率有限,这些结构往往显示为模糊的斑点^[40],需要一种优秀算法来检测荧光斑点,并进一步跟踪结构。此外,在超分辨率定位显微镜中,组织或细胞结构通常是对荧光点的稀疏发光图像进行重建获得的。精确、快速的斑点定位和准确的强度评估是获得超分辨率图像的必要过程之一^[41]。通过提高对密集和低信噪比斑点的检测能力,所提算法展示了提高通量的潜力,生物样品制备的尺寸和距离能够进一步得到减小,而且成像过程中使用的数值孔径可以得到进一步降低。

尽管有上述优点,所提算法也有一些局限性。对于超低信噪比(SNR为1),所提算法识别性能较差,无法在该高干扰情况下获取出色的识别性能。此外,所提算法具有较少参数量来处理复杂的识别任务,随机的初始参数会对模型训练有较大影响。最后,由于手动注释真实生物样品的斑点亚像素坐标成本高,使用模拟数据集用于训练,然而模拟数据和真实的荧光数据有一些偏差,这导致算法应用在真实荧光图像时造成一些性能损失。

参 考 文 献

- [1] Zhang Y D, Schroeder L K, Lessard M D, et al. Nanoscale subcellular architecture revealed by multicolor three-dimensional salvaged fluorescence imaging[J]. *Nature Methods*, 2020, 17(2): 225-231.
- [2] Chen J, Dalal R V, Petrov A N, et al. High-throughput platform for real-time monitoring of biological processes by multicolor single-molecule fluorescence[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(2): 664-669.
- [3] Kwon S. Single-molecule fluorescence *in situ* hybridization: quantitative imaging of single RNA molecules[J]. *BMB Reports*, 2013, 46(2): 65-72.
- [4] Smal I, Loog M, Niessen W, et al. Quantitative comparison of spot detection methods in fluorescence microscopy[J]. *IEEE Transactions on Medical Imaging*, 2010, 29(2): 282-301.
- [5] Sage D, Neumann F R, Hediger F, et al. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics[J]. *IEEE Transactions on Image Processing*, 2005, 14(9): 1372-1383.
- [6] Smal I, Niessen W, Meijering E. A new detection scheme for multiple object tracking in fluorescence microscopy by joint probabilistic data association filtering [C]//2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, May 14-17, 2008, Paris, France. New York: IEEE Press, 2008: 264-267.
- [7] Basset A, Boulanger J, Salamero J, et al. Adaptive spot detection with optimal scale selection in fluorescence microscopy images[J]. *IEEE Transactions on Image Processing*, 2015, 24(11): 4512-4527.
- [8] Ruusuvaari P, Manninen T, Huttunen H. Image segmentation using sparse logistic regression with spatial prior[C]//Proceedings of the 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO), August 27-31, 2012, Bucharest, Romania. New York: IEEE Press, 2012: 2253-2257.
- [9] Spilger R, Lee J Y, Chagin V O, et al. Deep probabilistic tracking of particles in fluorescence microscopy images[J]. *Medical Image Analysis*, 2021, 72: 102128.
- [10] Peng Y Y, Chen Z Y, Zhu W F, et al. ADS-Net: attention-awareness and deep supervision based network for automatic detection of retinopathy of prematurity[J]. *Biomedical Optics Express*, 2022, 13(8): 4087-4101.
- [11] O'Connor T, Anand A, Andemariam B, et al. Deep learning-based cell identification and disease diagnosis using spatio-temporal cellular dynamics in compact digital holographic microscopy[J]. *Biomedical Optics Express*, 2020, 11(8): 4491-4508.
- [12] Lin J P, Haberstroh F, Karsch S, et al. Applications of object detection networks in high-power laser systems and experiments[J]. *High Power Laser Science and Engineering*, 2023, 11(1): e7.
- [13] 高慧, 阎晓东, 张衡, 等. 基于 Res2Net 的多尺度遥感影像海陆分割方法[J]. *光学学报*, 2022, 42(18): 1828004.
- Gao H, Yan X D, Zhang H, et al. Multi-scale sea-land segmentation method for remote sensing images based on Res2Net[J]. *Acta Optica Sinica*, 2022, 42(18): 1828004.
- [14] 袁媛, 陈明惠, 柯舒婷, 等. 基于集成卷积神经网络和 Vit 的眼底图像分类研究[J]. *中国激光*, 2022, 49(20): 2007205.
- Yuan Y, Chen M H, Ke S T, et al. Fundus image classification research based on ensemble convolutional neural network and vision transformer[J]. *Chinese Journal of Lasers*, 2022, 49(20): 2007205.
- [15] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [16] 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.
- [17] Danial J S H, Shalaby R, Cosentino K, et al. DeepSinse: deep learning-based detection of single molecules[J]. *Bioinformatics*, 2021, 37(21): 3998-4000.
- [18] Newby J M, Schaefer A M, Lee P T, et al. Convolutional neural networks automate detection for tracking of submicron-scale particles in 2D and 3D[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(36): 9026-9031.
- [19] Zakrzewski F, de Back W, Weigert M, et al. Automated detection of the HER2 gene amplification status in Fluorescence *in situ* hybridization images for the diagnostics of cancer tissues[J]. *Scientific Reports*, 2019, 9(1): 8231.
- [20] Wollmann T, Rohr K. Deep Consensus Network: aggregating predictions to improve object detection in microscopy images[J]. *Medical Image Analysis*, 2021, 70: 102019.
- [21] Spilger R, Chagin V O, Bold C S, et al. Deep neural network for 3D particle detection in 3D fluorescence microscopy images via density map regression[C]//2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), March 28-31, 2022, Kolkata, India. New York: IEEE Press, 2022.
- [22] Speiser A, Müller L R, Hoess P, et al. Deep learning enables fast and dense single-molecule localization with high accuracy[J]. *Nature Methods*, 2021, 18(9): 1082-1090.
- [23] Okumura N, Yamada S, Nishikawa T, et al. U-net convolutional neural network for segmenting the corneal endothelium in a mouse model of fuchs endothelial corneal dystrophy[J]. *Cornea*, 2022, 41(7): 901-907.
- [24] Gudla P R, Nakayama K, Pegoraro G, et al. SpotLearn: convolutional neural network for detection of fluorescence *in situ* hybridization (FISH) signals in high-throughput imaging approaches[J]. *Cold Spring Harbor Symposia on Quantitative Biology*, 2017, 82: 57-70.

- [25] Wollmann T, Ritter C, Dohrke J N, et al. Detnet: deep neural network for particle detection in fluorescence microscopy images[C]//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), April 8-11, 2019, Venice, Italy. New York: IEEE Press, 2019: 517-520.
- [26] Bouilhol E, Savulescu A F, Lefevre E, et al. DeepSpot: a deep neural network for RNA spot enhancement in single-molecule fluorescence *in situ* hybridization microscopy images[J]. *Biological Imaging*, 2022, 2: e4.
- [27] Eichenberger B T, Zhan Y X, Rempfler M, et al. deepBlink: threshold-independent detection and localization of diffraction-limited spots[J]. *Nucleic Acids Research*, 2021, 49(13): 7292-7297.
- [28] Li B N, Gao J C, Chen S P, et al. POI detection of high-rise buildings using remote sensing images: a semantic segmentation method based on multitask attention res-U-net[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16.
- [29] Wang J K, Zhang X Y, Lü P Q, et al. Automatic liver segmentation using EfficientNet and attention-based residual U-net in CT[J]. *Journal of Digital Imaging*, 2022, 35(6): 1479-1493.
- [30] Ashkani Chenarlogh V, Shabanzadeh A, Ghelich Oghli M, et al. Clinical target segmentation using a novel deep neural network: double attention Res-U-Net[J]. *Scientific Reports*, 2022, 12(1): 4706916.
- [31] 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.
- [32] Liang S W, Huang Z Z, Liang M F, et al. Instance enhancement batch normalization: an adaptive regulator of batch noise[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(4): 4819-4827.
- [33] Imbert A, Ouyang W, Safieddine A, et al. FISH-quant v2: a scalable and modular tool for smFISH image analysis[J]. *RNA*, 2022, 28(6): 786-795.
- [34] Mateju D, Eichenberger B, Voigt F, et al. Single-molecule imaging reveals translation of mRNAs localized to stress granules[J]. *Cell*, 2020, 183(7): 1801-1812.
- [35] Chenouard N, Smal I, de Chaumont F, et al. Objective comparison of particle tracking methods[J]. *Nature Methods*, 2014, 11(3): 281-289.
- [36] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning, July 6-11, 2015, Lille, France. New York: ACM Press, 2015: 448-456.
- [37] Tajbakhsh N, Shin J Y, Gurudu S R, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1299-1312.
- [38] Cremer M, Grasser F, Lanctôt C, et al. Multicolor 3D fluorescence *in situ* hybridization for imaging interphase chromosomes[J]. *Methods in Molecular Biology*, 2008, 463: 205-239.
- [39] Solovei I, Cavallo A, Schermelleh L, et al. Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence *in situ* hybridization (3D-FISH) [J]. *Experimental Cell Research*, 2002, 276(1): 10-23.
- [40] Liu S L, Wang Z G, Xie H Y, et al. Single-virus tracking: from imaging methodologies to virological applications[J]. *Chemical Reviews*, 2020, 120(3): 1936-1979.
- [41] Lelek M, Gyparaki M T, Beliu G, et al. Single molecule localization microscopy[J]. *Nature Reviews Methods Primers*, 2021, 1(1): 39.